



ISSN: 2434-8821

***JAAL IN JACET
PROCEEDINGS***

VOL. 6

ISSN 2434-8821

JAAL in JACET Proceedings

Volume 6

The 6th JAAL in JACET Conference (JAAL in JACET 2023)

December 2, 2023, Ochanomizu University, Tokyo, Japan



JAAL in JACET Proceedings, Volume 6

Published by the Japan Association of College English Teachers (JACET)

First Division (JAAL in JACET) Committee

Director-in-Charge

SATO, Takehiro (Nagoya University of Foreign Studies)

Advisory Board

ODA, Masaki (Tamagawa University)

WATANABE, Atsuko (Bunkyo University)

Editorial Board

KATO, Yoshitaka (Chubu University)

SHIMOMURA, Fuyuhiko (Ritsumeikan University)

YAMANAKA, Tsukasa (Ritsumeikan University)

Reviewers

HATAKEYAMA, Yukako (Ochanomizu University)

KANAMARU, Toshiyuki (Kyoto University)

KAWAI, Yasushi (Hokkaido University)

KUROSHIMA, Satomi (Tamagawa University)

NAKAYAMA, Natsue (Bunkyo University)

NIGO, Yoshihiro (Yamaguchi Gakugei University)

YAMANAKA, Tsukasa (Ritsumeikan University)

YOSHIMUTA, Satomi (Kwassui Women's University)

Editorial Office

The JACET Office

Address: 55 Yokotera-machi, Shinjuku-ku, Tokyo 162-0831 JAPAN

Phone: +81-3-3268-9686 / Fax : +81-3-3268-9695

E-mail: jacet@zb3.so-net.ne.jp

Japan Association of Applied Linguistics in Japan Association of College English Teachers (JAAL in JACET) Proceedings, Volume 6

The 6th JAAL in JACET Conference (JAAL in JACET 2023), December 2, 2023, Ochanomizu University, Tokyo, Japan

Published by the Japan Association of College English Teachers (JACET)

First published on 31st March 2024

ISSN: 2434-8821

Cite as: *JAAL in JACET Proceedings, 6*

Copyright © 2024 by JACET

All rights reserved. No part of this publication may be reproduced or republished in any form without permission in writing from JACET.

The articles published herein do not reflect the opinions of JACET.

Contents

Articles in English

[Research Article]

Translanguaging and Pedagogical Issues

DANTE, Laurence 1

Reconsidering Vocabulary Acquisition: With a Focus on Polysemous Words

IIJIMA, Hisanori and FUJIWARA, Takafumi 5

What Does Writing an English Paragraph Mean to Learners? A Qualitative Exploration of Their Reflections

SAKAMOTO, Kiyoko 12

Applying Chomskyan Perspectives to English Language Teaching

YASUNAGA, Akie 20

Choosing Statistical Analyses for Small Samples

JONES, Marc 28

Articles in Japanese

(研究論文)

日本語の主題卓越型構造からの転移に関する一考察—日本人初級・中級英語学習者の文産出データの比較を通して—

橋尾 晋平 33

日本語母語英語学習者の時制誤用の変化—使用（用法）基盤モデルと事態把握の観点から—

長 加奈子 41

日本人大学生の会話に対する前提認識が英語会話に与える影響—英語会話データとフォローアップインタビューからの考察—

大谷 麻美 49

混合型研究法における妥当性検証の枠組み—教室におけるパフォーマンステストの形成的評価での事例研究—

松村 香奈 57

(実践報告)

必修英語授業における ChatGPT 活用の試み

大賀まゆみ・豊島知穂・山中司 65

語用論に着目した TOEIC Part 2 応用活動—ChatGPT を利用して—

藏屋伸子 69

Translanguaging and Pedagogical Issues

DANTE, Laurence*

**Shujitsu University
ldante@shujitsu.ac.jp*

Abstract

In recent years, translanguaging has gained much global recognition and has challenged several established concepts regarding linguistic theory and language pedagogy. It represents a potential paradigm shift in thinking concerning the use of language in classes and other academic settings. Translanguaging pedagogy challenges the conceptual framework that advocates the separation of languages in bilingual or multilingual educational environments. Moreover, translanguaging theory promotes the idea that educators should create environments in which learners can freely, flexibly, and dynamically use all of their linguistic repertoire to form meaning. This stance has the potential to augment and enhance deeper understanding of subject matter and also facilitate home-school cooperation, by supporting home language use and community involvement in education. Another important aspect of translanguaging theory is to challenge, and hopefully rectify, unfair and imbalanced power dynamics, the de-evaluation of languages and language practices, and discriminatory behavior toward minoritized individuals and groups. The purpose of this research is to explain and critically analyze the fundamental tenets of translanguaging in order to reveal some potential pedagogical issues that should be addressed for effective implementation of translanguaging pedagogical practices.

Keywords: translanguaging, language education, linguistic theory

1. Introduction

Translanguaging represents a potential paradigm shift in language learning theory and practice, particularly regarding bilingual and multilingual education. In modern times, translanguaging was first codified by Williams (1994), a Welsh researcher and educator who was attempting to revitalize the Welsh language and ensure that his students got a well-rounded education while using their Welsh and English language skills. He emphasized the use of both English and Welsh, all of his students' linguistic repertoire, not only for the acquisition of linguistic factors and features but also for the negotiation of meaning leading to deeper understanding.

Garcia and Wei (2014) further developed translanguaging theory and practice, particularly in the North American context, re-emphasizing flexible use of students' full linguistic repertoire and the need to challenge power dynamics that often alienate minority groups. The authors also explain that minority groups and individuals are often regarded to be "defective" monolinguals rather than valuable academic resources worthy of attention and respect.

Although a thorough discussion of the theoretical basis of translanguaging is beyond the scope of this essay, an explanation of the basic tenets and underpinnings of translanguaging is given below. It must, however, be noted that although translanguaging practices have begun to spread around the world, and some research has shown their beneficial effects, many questions remain regarding their effectiveness and appropriate methods of assessment.

2. Translanguaging

Translanguaging theory advances the idea that

educators should create environments in which learners can freely, flexibly, and dynamically use all of their linguistic (semiotic) repertoire to form meaning. Another basic tenet of translanguaging is that people do not learn languages but *do* languages, which includes the idea that social interaction and context are key elements in the development of linguistic features and cultural knowledge. This dynamic process, which includes various forms of interaction, emphasizes the flexible use of people's complex linguistic resources to make meaning in and of their lives (García, 2014).

Otheguy, García and Reid (2015) have defined translanguaging as "the deployment of a speaker's full linguistic repertoire without regard for watchful adherence to the socially and politically defined boundaries of named (and usually national and state) languages" (p. 283). As briefly mentioned above, national or "official" (named) languages often involve power dynamics that isolate and devalue other languages or dialects. Translanguaging involves the idea that theories about "language" must begin at the individual level and that interaction is a primary basis for the formation of meaning. Thus, translanguaging theory is based on heteroglossia and emphasizes idiolects over nationally supported or sanctioned languages. Furthermore, bilingualism or multilingualism should privilege speakers' own dynamic linguistic and semiotic practices above those of named languages.

Translanguaging also calls for more social justice in schools and respect for and recognition of all cultures and linguistic practices. This is a way to transform society so that the home languages and cultures of historically marginalized groups are not only respected but also recognized as resources

for higher levels of self-efficacy and academic success for all.

Translanguaging theory further posits that bilingual and multilingual speakers are not the amalgamation of two or more monolingual individuals, but “their language practices are seen as the deployment of different features from a unitary language repertoire for diverse social interactions with many social actors” (Kleyn & García, 2019, p. 72). More traditional bilingual or multilingual theory sees learners as an amalgamation of two or more monolinguals ($L1 + L2 = B$), with perhaps some overlapping aspects of language. However, the two named languages maintain their distinct linguistic features and are separate sources of vocabulary, concepts, etc. According to translanguaging theory, students have one unified linguistic repertoire which is similar to the image of a tapestry, with many linguistic and cultural features and experiences interwoven together. This repertoire serves as a resource or wellspring for the formation of meaning in life, and for the development of higher-level linguistic elements.

3. Translanguaging and Pedagogy

Regarding translanguaging and education, García, Johnson and Seltzer (2017) have elucidated three main components of translanguaging educational practices: stance, design, and shifts. A teacher’s stance encompasses their beliefs and ideologies in regard to emergent bilingual or multilingual individuals. In order for translanguaging to be fully effective, teachers should regard each student, and that individual’s linguistic and cultural background, as a resource for further learning; including academic achievement. This stance can help to transform and disrupt socially and culturally embedded structures of power and privilege.

Design refers to the idea that a teacher’s planning and methods should adhere to translanguaging theoretical guidelines. This includes creating time for pair and group work not only among speakers of similar home languages but also among a diverse combination of individuals. This also includes the creation of tasks that provide students with opportunities to use their full linguistic repertoire. In educational settings, resources should be bilingual or multilingual and multimodal, so that students have many opportunities to create meaning out of and internalize new features of language and culture. This includes interaction and engagement with bilingual or multilingual resources and individuals, including local community members and other individuals throughout the world via technology.

“Shifts” refer to a teacher’s flexible response to the pedagogical or learning flow among students in each classroom. Translanguaging should be a dynamic, interactive process that places students at the center of pedagogy and learning. Teachers must plan and design appropriate class content and tasks, but must also make unplanned changes to facilitate student learning and the construction of meaning. Teachers should monitor students and carefully listen to their interactions and queries, and then allow for, or facilitate, natural, purposeful interaction that leads to deeper understanding.

Several potential benefits of translanguaging pedagogy have been revealed through research. When students use

elements from their whole linguistic repertoire in various practices and activities, benefits include obtaining a deeper understanding of academic content, access to more information, the development of metalinguistic awareness, and higher levels of bilingual or multilingual proficiency (Cenoz & Gorter, 2022). Cummins (2021) has explained that well-planned translanguaging educational activities can scaffold learning and help students form connections between the school curriculum and their lives and community. Translanguaging practices can also help students form positive self-identities. Furthermore, these practices can help to develop and reinforce knowledge of how language works as a systematic form of communication.

In regard to benefits for English language learning in Taiwan, Ke and Lin (2017) have stated that translanguaging techniques can help students deepen their understanding of instructional content and help them develop a positive attitude toward English as an effective means of communication that is viable and useful in many contexts. Translanguaging practices also help learners develop a “positive identity” as a successful user of English.

Some examples of translanguaging pedagogical techniques and activities may help to clarify the somewhat abstract theory explained above. As a basic example, the current author used a translanguaging technique in an English language class for second-year English majors at a small, private university in Japan. The goal of the class was for students to carry out a 10-minute group discussion in English about eating meat and a healthy diet using research results and other evidence to support relevant points.

Despite the introduction of appropriate vocabulary, a review of reading assignments and other preparatory activities, the author noticed that students were still struggling to carry out an effective discussion. The author also noticed that some groups were negotiating meaning and asking for opinions in Japanese. Subsequently, the author gave students time to gather their thoughts and then had the students make groups and carry out the discussion in Japanese. Then, after another review of key vocabulary, phrases and concepts in English, with time for students to prepare some more notes, students were reformed into new groups and they carried out a much more effective, in-depth discussion of the topic in English. This is a simple example of allowing students to use their entire linguistic repertoire to understand and reconfirm concepts, so that they are able to effectively express themselves in at least a bimodal manner.

Cenoz and Santos (2020) provide another example of translanguaging pedagogy. They explain about a secondary school class in the Basque Country in which students used Basque, Spanish, and English to negotiate meaning. Students analyzed the structure of news articles about the same topic in all three languages. They then listened to the same content and discussed the content and structure of the news items. The students also composed a news item in the minority language, Basque. These activities helped students to develop more metalinguistic awareness of differences between oral and written language and differences in the structure of language and vocabulary among the three languages. Students also

developed their oral and written competencies, and at the same time, they gained a deeper understanding of content.

One final example involves an English language class in China. Students were asked to form a request in both Chinese and English from female employees to their supervisor regarding taking an afternoon off for personal reasons. Students were further informed that the employees were located respectively in China and Australia. The students were then asked to discuss their requests, paying particular attention to levels of directness, the use of politeness strategies, and other relevant cultural factors. This helped students to more clearly understand not only differences in linguistic elements but also sociolinguistic and cultural disparities. More effective intercultural communication skills could also be a benefit of such activities (Cenoz & Gorter, 2022).

Translanguaging activities such as those described above could be beneficial for students, however, they must be theoretically sound and thoroughly planned and executed or some negative effects may arise.

4. Translanguaging, Pedagogy and Cautionary Remarks

As mentioned above, translanguaging planning and methods, a teacher's stance, should be firmly based on translanguaging theories and guidelines. However, if teachers do not fully grasp and adhere to the spirit and content of translanguaging theory, some negative effects may arise. As an example of misunderstanding on the part of teachers, Aleksić and García (2022) describe research that they conducted in Luxembourgian preschools. The three teachers who participated in the research were white Luxembourgian females with more than 10 years of work experience. The teachers participated in more than 17 hours of professional development, which was carried out over six months in seven approximately two-and-a-half-hour sessions. It should be noted that this program was supported by the Luxembourg Ministry of Education.

The three teachers then planned a class based on their understanding of translanguaging principles. The class involved four children of Portuguese and Serbian heritage. The students were asked to pick a flag representing the language that they use at home. Three out of the four children wanted to choose the Luxembourg flag, explaining proudly that they speak Luxembourgish at home. However, due to preconceived notions about immigrants, the teachers urged and even coerced the students to choose the flag corresponding to their heritage language. During and after the lesson, the students seemed confused, disheartened, and demotivated. Although the teachers wanted to celebrate diversity, due to their prejudices about immigrants and linguistic practices, they ended up shaming the students and disempowering them as speakers of Luxembourgish.

As mentioned above, according to translanguaging theory, all students are worthy of respect. They are seen as unique individuals who should be allowed to use all of their linguistic and cultural capital to form meaning of the world around them, to enrich their lives and the lives of others. The teachers were trying to acknowledge or even promote

diversity, but they did not accept the children as full-fledged members of Luxembourg society and culture, and did not listen to what students were saying. Instead of making a "shift" they reflected majority cultural views in the classroom, "othering" the students, which had a negative effect on their sense of self-efficacy and motivation.

The following research provides an example of how well-meaning teachers could help to produce unsatisfactory results if translanguaging theory and practices are not accepted and promoted in educational settings as a whole, and in surrounding communities. Hamman (2018) conducted research concerning translanguaging practices at Rockland Elementary School, a kindergarten to second grade institution in a small Midwestern city in the United States. The research was conducted regarding a Spanish-English dual language program that consisted of 14 students divided evenly between their respective primary home languages. The Spanish students were mostly from low-income families, while the English speakers came from predominantly middle-class families. The program followed the 90/10 model; thus, kindergarten students were supposed to receive 90% of instruction in Spanish and 10% in the majority language, English. According to this model, the percentage of class content conducted in English increases until the fourth grade, when both languages are commonly used in even distribution.

In the second-grade class that Hamman (2018) studied, the language of instruction was supposed to be 70% Spanish and 30% English. However, due to resource limitations, pressure to prepare students for mandated tests in third grade, and various socioeconomic inequalities, the proper amount of language input was not maintained in the classroom. To further clarify, the two languages were separated in the classroom by subject. Reading, writing and mathematics were taught in Spanish, and science or social studies, depending on the day, were taught in English. However, since other classes, music, physical education and art, were taught in English, the actual ratio was 60/40. Furthermore, students could choose which language to use during lunch, recess, and in their free time activities. English remained the dominant language in school as well as in the community surrounding the school.

In the second-grade class, the bilingual teacher had tried to strictly separate the two languages, according to more traditional bilingual theory, but had become familiar with translanguaging practices. She, therefore, started to allow children to choose either English or Spanish for various tasks. Unfortunately, the English speakers began to show their dominance, and the amount of Spanish used in class markedly diminished.

Regarding the class explained above, although one main goal of the curriculum was the development of high-functioning bilingual and biliterate individuals, unfortunately, because of the dominance of English in the school and in the surrounding community, translanguaging opportunities actually increased the dominance of English. Free and flexible use of a student's full linguistic repertoire is a main tenet of translanguaging theory; however, classes and activities must be meticulously planned in order for the benefits of translanguaging to manifest themselves.

In summary, translanguaging for educational purposes “means that we start from a place that leverages all the features of the children’s repertoire, while also showing them when, with whom, where, and why to use some features of their repertoire and not others, enabling them to also perform according to the social norms of named languages as used in schools” (García & Kleyn, 2016, p. 15). This seems to acknowledge the idea that not only is the acquisition of linguistic and cultural knowledge important but that educators and students should also be aware of sociopolitical factors and realities that influence education. These factors and realities should be carefully considered when planning curricula and activities based on translanguaging theory.

5. Conclusion

For translanguaging, and other bilingual or multilingual strategies and pedagogical practices, to be effective, students should be welcomed into the classroom as full members of a given educational environment. For linguistic and academic development, students’ cultural, ethnic, and linguistic backgrounds, as well as those of their family members, should be respected and leveraged as vital educational resources. Furthermore, educators and staff members of educational institutions must be educated about translanguaging theory and methodology. Research-based, proven theory and practice should be embraced and serve as the basis for school-wide programs, teaching, and learning. Moreover, for benefits to manifest themselves, translanguaging activities must be thoroughly planned and executed, but at the same time, teachers must maintain a flexible stance in educational settings and facilitate spontaneous and free usage of multimodal means of communication and interaction.

Although translanguaging can be an effective method for enhancing multimodal communication and deeper understanding of academic concepts, more research must be conducted to answer many still unanswered questions. With that in mind, the following questions need to be answered in the future to unequivocally, or at least reasonably, prove that translanguaging theory and practice are viable and effective. What are effective ways to educate teachers, school staff, and local community members about translanguaging theory and pedagogical methods? Is the separation of languages needed for effective education? What is the proper ratio of language use in instruction? Is there a suggested mix of students from different backgrounds? Is simply creating a free and flexible environment enough to promote the acquisition of linguistic features, academic knowledge and analytical skills, etc.? How can the acquisition of high-level academic knowledge and achievement be promoted? What are some concrete guidelines for assignments, tasks and assessment? What are some clear guidelines for teachers who do not speak or understand their students’ languages and ethnic or cultural backgrounds? Is there any influence or what are some pedagogical concerns when students translanguage between languages that are relatively “distant” from each other, such as Japanese and English?

Teachers should educate themselves about translanguaging theory and should experiment with, and

conduct research about, translanguaging practices in their classes; while hopefully avoiding some of the pitfalls explained above. However, at the same time, for translanguaging to really generate a paradigm shift in bilingual or multilingual theory and education, advocates of translanguaging must answer the questions above, and provide concrete models of effective translanguaging curricula and practices.

References

- Aleksić, G., & García O. (2022). Language beyond flags: Teachers misunderstanding of translanguaging in preschools. *International Journal of Bilingual Education and Bilingualism*, 25(10), 3835–3848. <https://doi.org/10.1080/13670050.2022.2085029>
- Cenoz, J., & Gorter, D. (2022). Pedagogical Translanguaging and Its Application to Language Classes. *RELC Journal*, 53(2), 342–354. <https://doi.org/10.1177/00336882221082751>
- Cenoz, J., & Santos, A. (2020). Implementing pedagogical translanguaging in trilingual schools. *System*, 92, 102273. <https://doi.org/10.1016/j.system.2020.102273>
- Cummins, J. (2021). *Rethinking the education of multilingual learners*. Multilingual Matters.
- García, O. (2014). TESOL translanguaged in NYS: Alternative perspectives. *NYS TESOL Journal*, 1(1), 2–10.
- García, O., Johnson S., & Seltzer, K. (2017). *The translanguaging classroom: Leveraging student bilingualism for learning*. Brookes Publishing.
- García, O., & Kleyn, T. (Eds.). (2016). *Translanguaging with multilingual students*. Routledge.
- García, O., & Wei, L. (2014). *Translanguaging: Language, bilingualism, and education*. Palgrave MacMillan.
- Hamman, L. (2018). Translanguaging and positioning in dual language immersion: A case for criticality. *Language and Education*, 32(1), 21–42. <http://dx.doi.org/10.1080/09500782.2017.1384006>
- Ke, I. C., & Lin, S. (2017). A translanguaging approach to TESOL in Taiwan. *English Teaching & Learning*, 41(1), 33–61. <https://doi.org/10.6330/ETL.2017.41.1.02>
- Kleyn, T., & García O. (2019). Translanguaging as an act of transformation: Restructuring teaching and learning for emergent bilingual students. *The Handbook of TESOL in K-12*, 69–82. <https://doi.org/10.1002/9781119421702.ch6>
- Otheguy, R., García, O., & Reid, W. (2015). Clarifying translanguaging and deconstructing named languages: A perspective from linguistics. *Applied Linguistics Review*, 6(3), 281–307. <https://doi.org/10.1515/applirev-2015-0014>
- Vogel, S., & García, O. (2017). *Translanguaging*. In G. Noblit & L. Moll (Eds.), *Oxford research encyclopedia of education* (pp. 1–21). Oxford University Press.
- Williams, C. (1994). *Arfarniad o dduliau dysgu ac addysgu yng nghyd-destun addysg uwchradd ddwyieithog* (Unpublished doctoral dissertation). University of Wales, Bangor.

Reconsidering Vocabulary Acquisition: With a Focus on Polysemous Words

IJIMA, Hisanori* and FUJIWARA, Takafumi**

**Graduate School, Keio University*

hisanori@keio.jp

***Matsumoto University*

takafumi.fujiwara@t.matsu.ac.jp

Abstract

This research offers a novel perspective in the domain of vocabulary acquisition, with a particular emphasis on polysemy and its repercussions for learners of English. In cognitive linguistics, polysemy is characterized as the phenomenon whereby a single word form exhibits a spectrum of interrelated meanings. Although this aspect of vocabulary acquisition represents a burgeoning field of study, there is a notable dearth of comprehensive literature reviews. An experiment of literature search (Experiment 1) utilizing databases such as MLA International Bibliography and Linguistics and Language Behavior Abstracts (LLBA) collected approximately 1300 research articles. Subsequent analysis through the KJ method (Kawakita, 1967) facilitated the organization of these articles into three developmental categories: semantic classification, classification convergence, and the emergence of new theoretical frameworks. It was observed that most research during the new theory stage predominantly addressed polysemy from the vantage point of native language speakers, thereby overlooking its complexity in the context of English as a second or foreign language. A further examination of 249 selected articles, employing the KJ method, revealed a tendency to interpret language acquisition in terms of contextual factors, particularly pragmatics, over semantics. Considering these findings, we offer a vocabulary acquisition model based on usage-based perspective by Langacker (1990).

Keywords: polysemy, vocabulary acquisition, usage-based model

1. Introduction

In this section, firstly, we have shown what polysemy is, along with specific examples. Secondly, we also explain the importance of knowledge on polysemous words in the context of learning vocabulary. Finally, that necessity and significance of this research in the context of English language education were mentioned.

1.1 What is Polysemy? - The Definition of Polysemy -

Polysemy refers to the phenomenon whereby a single word, such as “run,” can have a variety of interrelated meanings (Ravin & Leacock, 2000). For example, “run” can refer to the physical act of running, as someone might do in a marathon, such as “run in a marathon”. Alternatively, it can mean to oversee or manage something, such as “run a business.” Despite these different applications, the meanings are considered related and part of the same word, rather than being separate words.

Linguists, or researchers, believe that these meanings are not random but semantically related as its definition is utilized as the premise to the linguistic research. That is, there is a perceived relationship between all the different uses of a polysemous word, and they are grouped together under the same “umbrella” term in our mental lexicon. The passage also cites several scholars who have contributed to the understanding of polysemy, indicating that this is a well-

researched and supported idea in linguistic studies.

The fact that many references are cited (Gries, 2006; Taylor, 2003; Goddard, 2000) suggests a strong academic interest and a substantial number of papers on the subject. Finally, the paper notes that historically the concept of polysemy has been considered more of a theoretical idea within academic discussions.

This emphasizes the fact that while the theories are well established to some degree, there are not as many applications to the teaching of English vocabulary, or how scholars have approached the theoretical diversity of interconnected meanings within a single word form.

1.2 Importance of Polysemous Words' Knowledge

Vocabulary knowledge is widely recognized as a crucial component in the acquisition of a foreign language due to its fundamental role in the development of listening, reading, speaking, and writing skills and its significant influence on English usage (Read, 2000; Saville-Troike, 1984). In addition, vocabulary serves as a key tool for effective communication in English (Newton & Nation, 2020; Schmitt, 2010).

Discussions about the aspects of vocabulary that language learners need to cultivate have recently focused on the breadth of vocabulary (vocabulary breadth), the depth of word knowledge (vocabulary depth), and the speed of word

recognition (cognitive speed) (Daller, Milton, & Treffers-Daller, 2007; Henriksen, 1999).

Vocabulary breadth involves the recognition of many words, including those that are less commonly used, whereas vocabulary depth involves the understanding of detailed aspects of words, such as their different meanings, uses, and associations with other words. Also, cognitive speed refers to the speed with which a person can recognize and understand a word. It means the quick retrieval and use of word meanings, and those who are proficient in this can use words quickly, indicating a robust vocabulary. Thus, the challenge in vocabulary instruction is to increase the number of words known (breadth), the depth of understanding of those words (depth), and the ability to recognize (or recall) them quickly (cognitive speed).

This study underscores the importance of learning high-frequency words, especially those with multiple meanings, because they are often polysemous. Nation (1990) suggested that learning about 3,000 high-frequency words would enable learners to comprehend 95% of typical texts. Since many basic polysemous words occur frequently, focusing on them is essential in English vocabulary instruction.

In terms of vocabulary depth, it is crucial to fully utilize the knowledge of polysemy. It's not enough to know just one meaning of a common word, the understanding of its multiple meanings. Vocabulary enhancement involves not only increasing the number of known words, but also deepening the understanding of the multiple meanings of each word (Nation, 2001; Tanaka, 2012).

1.3 Necessity of This Study in English Language Education

There exists a growing need within contemporary English language education for learners to obtain a broad vocabulary to capably respond to the diverse contexts and situations encountered in authentic communication. Specifically, comprehending polysemous words, which may adopt distinct meanings based on contextual factors, constitutes a vital component underlying the communicative competence. The ability of English language learners to aptly understand and apply the multiple senses of these terms is so critical for successful language acquisition.

However, the theoretical framework for the study of polysemy in English language teaching has not been well discussed. Therefore, the purpose of this study is to re-examine the theoretical framework of polysemy and to obtain suggestions for the teaching of polysemy.

2. Literature Review

2.1. The Reason for the Choice of Gries (2006)

In this section, firstly, we explained the overview of Gries (2006)'s research to elucidate the history of polysemy research within cognitive linguistics. But first, let us explain why the authors chose Gries (2006) as our target paper.

While the seminal work of Gries (2006) provided crucial insights into the development of polysemy research, an updated analysis was not found to characterize progress in the field since its publication. There appears to be a

dependence within the academic community on Gries' historical conceptualization of the developmental phases of cognitive linguistics, without systematic evaluation of subsequent theoretical and empirical progress. This dependence signals deficiencies in review surveys that capture contemporary trajectories in polysemy research in linguistics.

Consequently, few studies have endeavored to build upon Gries' (2006) foundational tripartite schema encompassing the "extreme splitting", "discussion and revision", and "newer developments" epochs. There persists an opportunity to scrutinize innovations in both methodological approaches and theoretical perspectives that postdate these designated eras.

As we distance further temporally from the time of Gries' (2006) writing, an updated retrospective synthesis becomes imperative to accurately depict the state of advancement within this rapidly evolving field. The integration of modern computational tools such as BERT and Word2Vec and so on also warrants an analytical reevaluation of theoretical framework of the polysemy research.

Therefore, the current academic landscape lacks holistic overviews contextualizing contemporary advancements and reimagining historic developmental frameworks to recent theoretical progression and its application. An urgent scholarly need exists to address this gap.

2.2. Gries (2006)'s Three Phases of Polysemy Research

Gries (2006) describes the three phases of polysemy research that have been done so far as follows. They are described in detail below.

2.2.1 The Period of "Extreme Splitting"

Initially, cognitive linguistics centered inherent in linguistic expressions, diverging from the primary focus on "meaning as language" that prevailed during its initial stage. A quintessential example of this is Lakoff's (1987) extensive analysis of the preposition "over", where he assigns differentiated significations to its usage, reflecting even the minutest semantic variations. Such as when 'over' is employed to depict a person vaulting a fence or clouds enveloping a mountain. During this era, Lakoff's (1987) approach led to what he described as a quasi-infinite proliferation of meanings, each distinguished by its unique nuances—a groundbreaking notion at the time.

2.2.2 The Period of "Discussion and Revision"

This era witnessed a paradigm shift towards the consolidation of meanings. Theoreticians like Lakoff (1987) acknowledged the potential for generating a limitless spectrum of meanings; however, this approach neglected considerations of lexicon and pedagogy. Consequently, there emerged an impetus to amalgamate terms sharing semantic affinities. With respect to 'over,' despite the fragmentation of its meaning by nuanced delineations, efforts were made to coalesce those with akin nuances.

2.2.3 The Period of “Newer Developments”

This phase signifies an endeavor to formulate a novel theoretical framework for polysemy. The prior two phases, namely, the “Extreme Splitting” and “Discussion and Revision,” were predominantly concerned with the identification and examination of specific polysemous words. Subsequently, research shifted towards a more holistic perspective, as exemplified by Taylor (2003), focusing on the overarching patterns and theories concerning polysemy.

2.3. Summary of Gries (2006)

These phases demarcate the tripartite schema of polysemy's historical development in cognitive linguistics as posited by Gries (2006), delineating a gradual progression from semantic granularity to theoretical consolidation. The passage discusses how there's not a lot of comprehensive overviews on the topic of polysemy, which is when a single word has several related meanings. He made an important contribution to this topic by bringing together a lot of research on it. However, his work is part of a bigger book or article and reflects his own take on the research.

Gries (2006) believes that the study of polysemy will start to rely more on data and technology, like large collections of text (corpora) and computer-based language analysis, instead of just looking at specific examples in a more subjective, less data-driven way.

3. Objective and Research Questions

Gries (2006) prognosticated an impending shift towards more quantitative and data-driven methodologies within polysemy research in cognitive linguistics. However, in the 17 years since its publication, there remains a paucity of empirical evidence that substantiates this prediction.

Therefore, the main purpose of this study is to conduct a systematic review of the literature, providing a synoptic appraisal of scholarly activity and theoretical progress in this domain. Through meticulous analysis, we endeavor to elucidate contemporary trends while identifying any lacunae within existing frameworks. To fulfill this objective, our inquiry shall be guided by two key research questions (RQs) written below.

RQ1) What is the trajectory of quantitative research in the domain of polysemy postdating the seminal work of Gries (2006)?

There is an implied assumption within the academic community that as we approach the present day, research within cognitive linguistics has progressively adopted a more quantitative methodology. This study seeks to evaluate the veracity of this assumption. Is the shift towards quantitative analysis substantiated by empirical evidence, or is it merely a perception colored by scholarly subjectivity?

RQ2) Has there been a paradigm shift in the theoretical framework of polysemy since 2006?

While there has been an acknowledgment of efforts to

formulate a comprehensive theory of polysemy, the question remains whether these endeavors have crystallized into a novel theoretical edifice post-2006. Have we witnessed the emergence of groundbreaking theories that redefine our understanding of polysemy?

The research will meticulously address these questions, aiming to illuminate the recent scholarly discourse and theoretical advancements in the study of polysemy.

4. Method

The experiment was divided into two parts: one for Research Question 1 (RQ1) and one for Research Question 2 (RQ2).

4.1 Experiment 1

To assemble a comprehensive dataset for our investigation into polysemy, we employed a systematic search strategy across two primary databases: Linguistics and Language Behavior Abstracts (LLBA) and Modern Language Association International Bibliography (MLA).

Our search methodology was meticulously designed to capture the breadth of research on polysemy. For LLBA, we initiated our inquiry with the term “Polysemy” to encapsulate the entirety of the field's research, setting our temporal parameters from the year 1990 up to 2023. The choice of the initial year corresponds with the founding of the International Cognitive Linguistics Association (ICLA) and the official publication of a dedicated cognitive linguistics journal, marking the formal establishment of the field. The meta-review search strategy employed in the LLBA and MLA databases was meticulously crafted with specific parameters to filter scholarly discourse on polysemy.

Scholarly journals were the exclusive source type considered, ensuring a focus on academic rigor. The temporal scope of the search was set from January 1, 1990, through August 31, 2023, providing a comprehensive overview of over three decades of research. The search query was carefully constructed to exclude documents such as review articles, case studies, reports, undefined documents, and articles pertaining to evidence-based medicine (EBM), thereby narrowing the focus specifically to original research articles on polysemy. The search was targeted to retrieve documents written in English, with 'polysemy' as the central theme, and limited to journal articles, streamlining the results towards high-quality academic contributions to the topic.

Similarly, for MLA, the same keyword “Polysemy” was utilized with an identical time frame. Post-exclusion of non-linguistic materials and uncategorized texts, we identified 482 references. Again, we narrowed our focus to those studies conducted in English. In the final phase of our data selection, we excluded any scholarly works not authored in English, such as those in Italian or Spanish, to align with the linguistic capabilities of the researchers involved in this study. Consequently, the language of the research corpus was standardized to English, culminating in a collection of 759 references from LLBA and 429 from MLA. This methodology ensures a robust and focused compilation of

literature for our meta-review; that means we collected more than 1300 papers as the first experiment.

Figure 1

Formula of LLBA

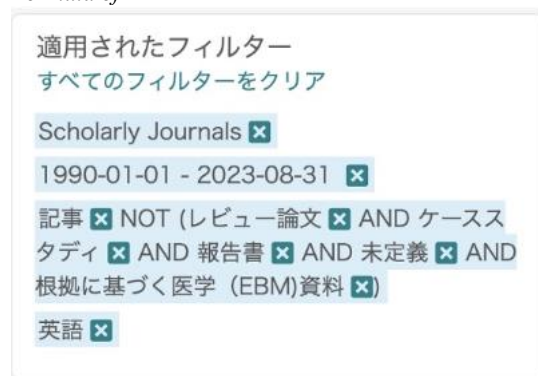


Figure 2

Formula of MLA



4.1.1 Analysis - The Use of KJ Method

4.1.1.1. The reason for the Choice of KJ Method.

The KJ Method, also known as the Affinity Diagram, was developed by Jiro Kawakita in 1960s (Kawakita,1967).

It is a tool used in the field of knowledge management and is particularly useful for organizing large amounts of data into meaningful groups, often in the context of brainstorming or problem-solving sessions. The KJ Method is particularly effective for synthesizing qualitative data, revealing hidden patterns, and fostering a shared understanding among participants.

In detail, the KJ method is an effective qualitative data analysis technique, particularly for categorizing large volumes of data. In this study, over 1300 articles were subjected to analysis. Consequently, the KJ method was deemed optimal for processing this substantial amount of data. Moreover, the KJ method is frequently utilized in the brainstorming and problem-solving processes of researchers. Therefore, it was judged to be a suitable technique for examining article trends, as required by this study.

Furthermore, as the KJ method involves collaborative classification based on manipulating visual maps, it is adept at organizing intricate information into logical,

comprehensible categories. Compared to other qualitative approaches such as thematic analysis, the KJ method excels at discerning patterns across substantial datasets. Additionally, it represents a data-driven approach, thereby reducing authors' subjectivity.

For these reasons, the KJ method was determined to be the most appropriate for this study and we wrote down based on the following four criteria in a card.

1) Perspective on Meaning

Since the study of polysemy is the study of "words," it is necessary to clarify one's stance on "words." Therefore, we considered it necessary to clarify our stance on "words" in this experiment. According to Tanaka (1983), there are two predominant perspectives: pragmatic and semantic. Based on what Tanaka (1983) claims, we identified each study's theoretical stance towards the stances of word meanings. The pragmatic perspective argues that the meaning of a word is a context dependent, whereas the semantic perspective suggests that word meaning is context independent. Our classification noted which of these perspectives was embraced in the studies of polysemous words.

2) Experimental Methodologies

We delineated the various experimental approaches employed across the studies, categorizing them as corpus-based analyses, psychological experiments, computational natural language processing techniques, or the compilation of illustrative examples. This classification enabled us to capture the methodological diversity within the research on polysemy.

3) Developmental Phases

Taking cue from the tripartite division posited by Gries (2006), we synthesized the categorization of each paper according to these phases. In other words, we investigated where the experimental articles fall into the three levels of categorization in his polysemous language research. This historical contextualization allowed us to trace the evolution of the field and understand the progression of theoretical and empirical approaches.

4) Lexical Focus

Early research (From 1990s to mid-2000s) in polysemy often centered on the analysis of specific words to ascertain their polysemous words. We collated a list of such words from the studies to identify patterns in lexical selection and to observe the range of polysemy across different lexemes. By organizing the data on these cardstock sheets, we facilitated a visual and thematic analysis of the accumulated research, allowing for a coherent synthesis of the field's current state, methodological trends, and theoretical developments in the study of polysemy.

4.1.2. Results

Findings was described based on the four categories as aforementioned. Perspective on Meaning, Experimental Methodologies, Developmental Phases, and Lexical Focus were explained. Each section shows the overall trend.

1) Perspective on Meaning

The analysis indicated a paradigmatic transition from a semantic to a pragmatic orientation in the study of meaning within polysemy research. A significant corpus of contemporary studies, exemplified by Evans (2015), delineates ‘semantic concepts’ from ‘meaning,’ concurrently advocating for the interpretation of meaning as contextually contingent. This shift underscored a growing consensus that word meanings were not fixed entities but were dynamically shaped by the situational context in which they were deployed.

2) Experimental Methodologies

A historical review of methodological approaches revealed a clear evolution. Initially, during the nascent and mid-phase 2 periods as identified by Gries (2006), research predominantly relied on author-generated examples. The advent of the 21st century heralded a methodological shift concomitant with the development of linguistic corpora and computational tools, steering research away from the reliance on constructed examples towards empirical, data-driven analysis termed as “quantitative turn” by Janda (2013).

3) Developmental Phases

The developmental period of cognitive linguistics, from its inception in the 1990s through the 2000s, was primarily characterized by studies falling within what Gries (2006) denotes as “Phase 2.” This period was marked by a concerted effort to consolidate meanings and was reflective of the field's maturation. The detail is shown in the following graph.

Figure 3

Phase, Percentile, the Number of Papers

Phase 1	21.80%	284
Phase 2	59.10%	770
Phase 3	19.10%	249
Total	100.0%	1303

4) Lexical Focus

Lexical focus was divided into the following five categories.

1. Prepositions and adverbs with high semantic frequency that have been studied since the early days, such as Lakoff (1987)
- 2 Basic verbs (From initial stage)
- 3.Modal auxiliary (e.g., Sweetser,1990)
- 4.Before cognitive semantics, the relationship between homonymy and polysemy was discussed. The most frequently discussed topics were “bank” and “ear.”
- 5 (Especially, In English language learning) there is a practical problem of the polysemy of higher-level words, such as those in academic vocabulary.
- 6 Examples of cross-synonymity: book

4.1.3 Discussion

The initial experimental findings exhibit considerable alignment with the theoretical prognostications of Gries (2006). Nevertheless, it is noteworthy to acknowledge the substantial advancements in computational linguistics tools that have exceeded his predictions.

In particular, the recent emergence of corpus tools such as the BNC (British National Corpus) and COCA (Corpus of Contemporary American English), as well as natural language processing tools such as BERT and Word2Vec, were not mentioned in Gries (2006), so his research in linguistics through these studies has been what he expected in the paper. These tools, with their profound capabilities for understanding the nuances of human language, have provided unprecedented avenues for the empirical exploration of polysemy. Their algorithmic intricacies enable a more nuanced detection and interpretation of semantic variations than was previously anticipated.

The predictive nature of Gries’ (2006) work did not fully encapsulate the rapid evolution and the profound impact of these technologies on the field. Thus, while the foundational aspects of his hypotheses remain pertinent, the advent and integration of advanced NLP systems have enriched the methodological repertoire available to researchers, thereby expanding the horizons of polysemy research beyond the initial predictions. This technological progression underscores the necessity for continuous adaptation of theoretical frameworks to incorporate the capabilities and insights afforded by cutting-edge computational tools.

4.2. Two Problems

The first experiment analyzed how the study of polysemy has evolved according to the four criteria. The results are identical to those that Gries (2006) has written semi-subjectively. Under the circumstances, we decided to conduct the second experiment for the following reasons.

Experiment 1 sought to validate the seminal hypotheses posed by Gries (2006) regarding developmental trajectories within polysemy research. Through meticulous analysis, categorized across four categorical dimensions, the corroboration of his theoretical model was established.

However, given the temporal distance since the original publication, the need to formulate an updated framework was deemed imperative. Thus, Experiment 2 sought to address this need by exclusively examining the subset of articles characterized by Gries (2006) as manifesting “Newer Developments,” constituting his designated Phase 3. By re-implementing the KJ methodology, an attempt was made to identify emergent patterns and elucidate novel theoretical models being put forth in contemporary research. Consequently, the linkage between the experiments focused on validating the foundational tripartite model proposed by Gries (2006) in Experiment 1, followed by an evaluation of modern advances within the pivotal Phase 3 epoch in Experiment 2. This dual approach facilitated the incremental formulation of an updated perspective.

4.3 Experiment 2

The ‘Newer Developments’ within polysemy research, as delineated by Gries (2006), represents a pivotal era of linguistic exploration and theoretical advancement. Prior investigations, such as our initial experiment, have touched upon the emergent concepts and methodologies characterizing this period. However, a comprehensive understanding of the depth and breadth of the theoretical frameworks developed post-Gries (2006), and the methodological underpinnings that facilitated such advancements, has yet to be fully elucidated. For these reasons, “Experiment 2” was conducted.

4.3.1 Analysis tool - KJ Method -

To address this scholarly need, we proposed to implement the KJ method, as in Experiment 1, a robust qualitative analysis technique, to dissect and interpret a great much research conducted during this “Newer Developments”.

This analysis was to fulfill a dual purpose: firstly, to critically evaluate the introduction and integration of new theoretical models in polysemy studies postdating Gries’ (2006) influential publication, and secondly, to methodically appraise the research methods that have been pivotal in propelling theoretical progress in this field.

4.3.2 Analysis

As mentioned above, 249 papers were classified as phase 3. Therefore, they were included in the analysis.

Through the application of the KJ method as in Experiment 1, we anticipated the assembly of a synthesized dataset that will unveil discernible patterns, emerging trends, and notable lacunae in the contemporary polysemy research landscape.

4.3.3 Result

The findings of this study suggested three principal insights.

1) Transition from Corpus to Natural Language Processing

The investigators initially hypothesized that, despite the prevalence of quantitative analyses, most of the research would predominantly employ corpus-based methods. Contrary to these assumptions, it emerged that studies leveraging natural language processing (NLP) tools surpassed those reliant on corpus analyses in number.

2) Advancements through Natural Language Processing

Building upon the previous point, it has been proposed that the clarification and understanding of linguistic phenomena via NLP are advancing. However, these advancements are not confined to the realm of cognitive linguistics; rather, as the field of NLP research intensifies, the exploration of linguistic phenomena is transcending disciplinary boundaries, indicating a broader interdisciplinary engagement.

3) The Scarcity of Educational Applications

Notwithstanding the emphasis on the elucidation of

linguistic phenomena, it has been observed that investigations aimed at applying these insights within the context of English language education remain exceedingly scarce. This highlights a significant gap in the research, emphasizing the need for a focused exploration of how linguistic phenomena, elucidated through NLP, can be effectively integrated into educational practices.

4.3.4. Discussion

In this second experiment, a comprehensive examination of the contemporary landscape of polysemous words’ research within the domain of cognitive linguistics was conducted by utilizing the KJ method. The outcomes revealed a notable progression from analyses based on corpora to the adoption of Natural Language Processing (NLP) methodologies. This observation indicates that NLP has the potential to offer novel perspectives and methodologies for data analysis in cognitive linguistics research.

Nevertheless, the present underutilization of linguistic insights in the NLP domain highlights a communication gap between the two disciplines, underscoring the necessity for enhanced interdisciplinary collaboration in subsequent studies.

Additionally, it was ascertained that the integration of NLP research findings into educational practices necessitates several preparatory steps. This underscores the imperative to devise actionable guidelines and frameworks to facilitate the effective implementation of NLP techniques in actual educational contexts.

5. Conclusion and Next Step

5.1 Conclusion

The objective of this investigation was to re-evaluate the strategies for vocabulary acquisition, with a specific emphasis on polysemy, by offering a comprehensive review of the scholarly inquiries into polysemy after Gries (2006).

Through the synthesis of extant research, it has become apparent that the post-Gries (2006) era has witnessed a pronounced shift towards “quantitative research” over “qualitative research.” Notably, the research conducted during the phase 3 period, as delineated by Gries (2006), was identified as critically significant and possessed the capacity to radically alter the conceptual framework of polysemy, thereby leading to the formulation of a novel theoretical construct. This development has been corroborated through the application of the KJ method. Subsequently, the conclusions drawn from this study will be articulated in alignment with the two research questions.

RQ1) What is the trajectory of quantitative research in the domain of polysemy postdating the seminal work of Gries (2006)?

As our Experiment 1 suggested that this query could be succinctly addressed by affirming that “quantitative research” has gained ascendancy. However, a more nuanced explanation is warranted. Two principal factors have

catalyzed the predominance of quantitative research: firstly, the advent of corpora has precipitated a shift towards more quantitative methodologies in research, and secondly, the ensuing application of natural language processing (NLP) tools is effectuating a profound transformation in the paradigm of cognitive linguistics.

RQ2) Has there been a paradigm shift in the theoretical framework of polysemy since 2006?

There has been no definitive evidence to suggest a complete transformation in the theoretical understanding of polysemy. Historically, polysemy has predominantly been examined within the realm of semantics; however, there is a growing trend towards analyzing it through the perspective of pragmatics.

5.2 Next Step

Subsequent research ought to meticulously scrutinize the array of methodologies employed throughout the “Newer Developments,” alongside the outcomes they yielded. Furthermore, despite the experimental corroboration that Gries (2006) suggested, the pursuit of more sophisticated investigative approaches, including meta-analytical studies, is imperative for the continued progression and refinement of the field.

Structured around Gries (2006) as an overview of previous research on polysemy in language, this study draws several conclusions regarding beneficial practices for English language teaching. Given the pragmatic nature of polysemy, a method to learn polysemous meanings in context appears feasible. This aligns with Langacker’s (1990) usage-based model which posits that “one’s ability to use language is shaped by the way words are used.” In pedagogy, it would be advisable to develop teaching methods that promote the contextual disambiguation of polysemous words and instruction accordingly. For instance, language activities that simulate real-life situations could boost learning motivation concerning specific polysemous words and their application to authentic contexts. Such measures are expected to enrich teaching materials and bolster the overall development of students’ language abilities.

In other words, in the conclusion of this paper, we would like to propose vocabulary learning through context, utilizing the usage-based model (Langacker, 1990).

Acknowledgments

We would like to thank Professor Toshio Ohori of the Faculty of Environment and Information Studies at Keio University for his invaluable advice in carrying out this research. We would like to express our deepest gratitude to him. Also, we would like to appreciate the invaluable comments from the two anonymous reviewers. We would like to express our gratitude to both.

Reference

Daller, H., Milton, J., & Treffers-Daller, J. (Eds.). (2007). *Modelling and assessing vocabulary knowledge*. Cambridge University Press.

Evans, V. (2015). A unified account of polysemy within LCCM

theory. *Lingua*, 157, 100–123.
<https://doi.org/10.1016/j.lingua.2014.12.002>

Goddard, C. (2000). Polysemy: A problem of definition. In Y. Ravin & C. Leacock (Eds.), *Polysemy: Theoretical and computational approaches* (pp. 129–151). Oxford University Press.

Gries, S. T. (2006). Chapter 2: Polysemy. In E. Dąbrowska & D. Divjak (Eds.), *Cognitive Linguistics – Key topics* (pp. 23–43). De Gruyter Mouton.

Henriksen, B. (1999). Three dimensions of vocabulary development. *Studies in Second Language Acquisition*, 21(2), 303–317. <http://dx.doi.org/10.1017/S02722263199002089>

Janda, L. (2013). *Cognitive linguistics – The quantitative turn: The essential reader*. De Gruyter Mouton.

Kawakita, J. (1967). Hassou-hou [KJ Method]. Chuko shinsho.

Lakoff, G. (1987). *Women, fire, and dangerous things: What categories reveal about the mind*. University of Chicago Press.

Langacker, R. W. (1990). *Concept, image, and symbol: The cognitive basis of grammar*. De Gruyter Mouton.

Nation, I. S. P. (1990) *Teaching and learning vocabulary*. Cengage Learning.

Nation, I. S. P. (2001). *Learning vocabulary in another language*. Cambridge University Press

Newton, J. M., & Nation, I. S. P. (2020). *Teaching ESL/EFL listening and speaking* (2nd ed.). Routledge.

Ravin, Y., & Leacock, C. (2000). Polysemy: An overview. In Y. Ravin & C. Leacock (Eds.), *Polysemy* (pp. 1-29). Oxford University Press.

Read, J. A. (2000). *Assessing vocabulary*. Cambridge University Press.

Saville-Troike, M. (1984). What really matters in second language learning for academic achievement? *TESOL Quarterly*, 18(2), 199–219. <https://doi.org/10.2307/3586690>

Schmitt, N. (2010). Vocabulary use and acquisition. In N. Schmitt (Ed.), *Researching vocabulary. Research and practice in applied linguistics* (pp. 3–43). Palgrave Macmillan UK.

Sweetser, E. (1990). *From etymology to pragmatics: Metaphorical and cultural aspects of semantic structure* (Vol. 54). Cambridge University Press.

Tanaka, S. (1983). *Language transfer as a constraint on lexico-semantic development in adults learning a second language in acquisition-poor environments* [Doctoral dissertation, Teachers College, Columbia University]. University Microfilms International.

Tanaka, S. (2012). New directions in L2 lexical development. *Vocabulary Learning and Instruction*, 1(1), 1–9.
<https://doi.org/10.7820/vli.v01.1.tanaka>

Taylor, J. R. (2003). Polysemy’s paradoxes. *Language Sciences*, 25(6), 637–655.
[https://doi.org/10.1016/S0388-0001\(03\)00031-7](https://doi.org/10.1016/S0388-0001(03)00031-7)

What Does Writing an English Paragraph Mean to Learners? A Qualitative Exploration of Their Reflections

SAKAMOTO, Kiyo*

**The University of Shiga Prefecture
sakamoto.ki@office.usp.ac.jp*

Abstract

This study reports on how 94 Japanese-speaking English learners (first- and second-year students in three different departments at a public university) experienced the activity of writing a paragraph in English. After writing an English paragraph and rewriting it with teacher feedback, the study participants answered open-ended questionnaires about how they viewed the topic of the paragraph they were asked to write and how they felt about the writing experience itself. The responses were sectioned, continuously compared, and grouped into categories for qualitative analysis. It was found that study participants had different reactions to the given paragraph topic, and their preferred topics were also quite different. As for the experience of writing an English paragraph, some participants found it useful or enjoyable while about twice as many other participants found it difficult for various reasons. The final section examines what the experience of writing English may mean to the participants in this study. Pedagogical implications are also discussed on the basis of these findings.

Keywords: qualitative analysis, open-ended questionnaires, writing topics, EFL writing experience, Japanese university students

1. Introduction

This study is a qualitative exploration of how the activity of writing a paragraph in English was perceived by students in English classes of a medium-sized public university. As Casanave (2018) puts it, the goal of qualitative research on English as a second language writing is “to uncover and understand the local and contextualized particulars of what it means for a person to learn to write in a second language and simultaneously to continue second language development” (p. 2466). With this in mind, this study aims not to generalize the results of the analysis to learners of English as a foreign language (EFL) or to Japanese university students in general, but to gain insight into the experiences of the learners who participated in this study from their perspective. Furthermore, it is hoped that understanding of their experiences may lead to ways for teachers to make the activity of writing in English more meaningful for learners.

The first research question of this study is related to the choice of topics in writing. There has been a body of research on the effects of writing topics in EFL writing learning. Both Winfield and Barnes-Felfeli (1982) and Tedick (1990) found that when the writing topic was related to learners (because of their culture or field of study), the learners gained better evaluation on their writing, not only in content knowledge but also in the quality of the writing product, such as fewer grammatical errors. He and Shi (2012) further investigated the impact of topic familiarity on learners' writing, asking international students studying at a Canadian university to write about topics connected to university life (familiar to the students) and topics related to Canadian politics (unfamiliar to them). Regardless of the learners' level

of English proficiency, writings on the familiar topic were rated higher in terms of both content and organization. Other studies measured the complexity, accuracy, and fluency of learners' writing when they write about different topics; Yang and Kim (2020) found that a less familiar topic had substantially negative effect on lexical complexity of the study participants' writing, and Abdi Tabari and Wang (2022) found a considerably positive effect in syntactic complexity with a familiar topic. From a different perspective, Li and Yu (2023) focused on how topic familiarity affects the participant's writing behavior rather than the written products. By recording the keystrokes of learners working on writing assignments, they found that when writing about less familiar topics, they wrote slower, wrote less, and spent more time planning, while spending less time revising and rewriting. While such a body of research exists, not many studies have addressed the question of how writing topics are perceived by the learners themselves. Lo and Hyland (2007) is one of the few examples, reporting on an action research study of a new program in a fifth-grade ESL writing class in Hong Kong. The writing topics in the new program were more relevant to the learners' real-life experiences and were thought to better motivate them to write. They assessed learners' writing as they implemented the old and new programs, while interviewing learners on an ongoing basis. The learners enthusiastically welcomed the new tasks, but their ratings of writing accuracy declined somewhat. Also, the learners did not necessarily have a negative reception of topics from the old program, which seemed far removed from their experience. Thus, the impact of writing topics on learners and their writing is multidimensional and yet warrants further

study. In particular, it seems necessary to explore how and why learners in various context respond to the topics, based on a qualitative analysis.

This study further aims to investigate what it means for the learners in this study to write in English, and to explore ways in which the findings can be used to help them learn English writing better. Learners who participated in the current study were asked to complete an open-ended questionnaire, reflecting on their experience after completing the task of writing a paragraph in English. The resulting learners' own words were used as data, and continuous comparisons were used to analyze them. The intercepted data were repeatedly read, compared, categorized, and reread to organize the most plausible categories. This ensured that a wide variety of responses could be understood and described, and that it is possible to return to the original data if necessary. Note that this study was conducted in the context of regular educational practice, and as such, no experimental or control group was established. However, as a result of inviting students from four classes belonging to three different departments to participate in the study, it was possible not only to analyze the responses of the participants as a whole, but also to do natural group comparisons between the different departments.

Research questions of this study were as follows.

RQ1: How did three groups of Japanese-speaking English learners (first- and second-year students of different departments in a public university) reflect on the topics of paragraph they were assigned to write?

RQ2: How did the same three groups of Japanese-speaking English learners reflect on the experience of writing a paragraph in English?

RQ3: Based on the above findings, what are the possible meanings of learning English writing for the learners who participated in this study?

2. Method

2.1 Study Participants

The researcher of this study was a native Japanese speaker with approximately 25 years of university English teaching experience. During that time, she had taught writing-intensive classes for several years and had incorporated writing instruction into other classes when possible.

The participants in this study were students in the classes taught by the researcher, all of whom were first- or second-year university students who were able to use Japanese as their native language. At the time the data were collected, weekly 90-minute face-to-face classes had been held for about two and a half months, so a certain degree of friendly rapport had been established between the researcher and the participants. In addition, each class was composed of students from the same department, who were often taking other courses together and sharing similar interests in specialized subjects. As a result, it was observed that the relationship between the students was rather good, especially in the second-year classes. All classes were English courses that were required by the common education program. Class participation was good, and with the exception of a few

students, assignments were submitted punctually.

The departments and breakdown of the study participants are shown in Table 1. Group A was one class of sophomores in Biological Resources Management, who were to major in the fields of agriculture, fisheries, livestock, and their relationship to the environment. Group B was one class of sophomores in Human Relations Studies, which offers majors in psychology, sociology, and education. Group C was two classes of freshmen in Intercultural Communications. English language proficiency was higher in Group C, Group B, and Group A, in that order.¹⁾ Many of the students in Group C planned to study abroad while in college, so their motivation or urgency to learn English was presumably higher than that of the other groups. The difference between the number of participants in the study and the number of participants who actually provided data was due to the exclusion of participants from whom all data could not be collected because of absence or other reasons.

Table 1
Study Participants

	Group A	Group B	Group C
Department	Biological Resources Management	Human Relations Studies	Intercultural Communications
Year	2nd	2nd	1st
TOEIC score ²⁾	430	(no data)	570
Class enrollment	32	31	52 (in two classes)
Study participants	30	25	51
Participants whose data were used	25	18	51

Recruitment of research participants was done carefully, given that the researcher was teaching these classes. First, the research plan was submitted to the university's ethics committee for approval before the research began. Next, a research cooperation request form, which was also approved by the university, was distributed to each student in the class. Participants were recruited after the researcher emphasizing that the evaluation of the learners would not change at all depending on whether they participated in the research or not. The research cooperation request form included a contact person within the university in case students believed they had suffered human rights violations in participating or not participating in the study. The research participation consent form included a form for students to withdraw their consent after giving it as well.

2.2 Data Collection

This study was conducted in the first semester of 2023 (April-July, 2023). Group A and Group B classes used the same textbook (Pavlik, 2015), while the two Group C classes used the same series as Group A&B, but with a slightly higher-level textbook (Pavlik, 2014). Both textbooks

consisted mainly of two reading materials for each unit, with exercises for content comprehension, idiomatic expressions, vocabulary building and so on. Each unit also has a section called “Idea Exchange” in which the students were to express their thoughts in English about a topic related to the unit. Students prepared their responses to the “Idea Exchange” in English and worked in pairs to talk about their opinions. Each student also made an English presentation on a topic of their choice, once during the semester.

Paragraph writing was introduced in Week 10 for Group A and B, and in Week 12 for Group C. The researcher explained the structure of an English paragraph and asked the students to write a paragraph for homework. Group A and B were to write on any topic related to “Work” and Group C on any topic related to “Marriage.” The learners were allowed to choose their own topics freely. The theme of Group A and B, “Work,” was the theme of Unit 4 of their textbook, and the students had already studied it in the fourth and fifth weeks of class. Therefore, they had already answered questions such as “What is your dream job in the future? Why?” in the Idea Exchange section of Unit 4. Similarly, Group C had studied the “Marriage” unit of their textbook and thought about the questions in the Idea Exchange, such as “Do you think people will stop marrying in the future?” and “When you look for a partner, what does and does not matter?” The learners were allowed to get some hints from these questions when thinking about their topics.

After the paragraphs were submitted, the researcher used an annotation tool called TIARA (Putra et al., 2021) to prepare a graphical representation of the paragraph structure and the problems with the connections between sentences for each student. This feedback was returned to learners in Week 14 for all classes, and homework for that week was to rewrite and submit their paragraphs based on the feedback. Then, the learners were asked to respond to a questionnaire about their paragraph writing (see Figure 1).

Figure 1

Questionnaire

Please answer the five questions below. Feel free to write what you think in Japanese.

Q1: What did you think of the structure of your paragraph as presented in TIARA?

Q2: How did you feel when you rewrote your paragraphs based on TIARA's analysis?

Q3: What did you think of the theme of the paragraph? Was it easy to write? Was it difficult to write? What would you like to write about?

Q4: How do you feel about writing paragraphs in English?

Q5: If you have any questions about how to write paragraphs, use TIARA, etc., please write them down.

Note. Only the responses to Q3 and Q4 were used in this study.

2.3 Data Analysis

As mentioned earlier, basic qualitative analysis (continuous comparison) was used to analyze the data. It is defined by Merriam and Tisdell (2015) as follows:

The analysis of the data involves identifying recurring patterns that characterize the data. Findings are these

recurring patterns or themes supported by the data from which they were derived. The overall interpretation will be the researcher's understanding of the participants' understanding of the phenomenon of interest. (p. 25)

In order to identify the “recurring patterns,” this study used “Ueno’s qualitative analysis” (Ueno, 2018), based on the KJ method (Kawakita, 1970). First, the answers to Q3 and Q4 of the above questionnaire were transcribed one by one into a single Excel square. After attaching the respondent's code (group name A, B, or C and a randomly assigned number 1, 2, 3... for each participant), the sheet was printed and cut into small cards, one per square. By reading and comparing the cards with each other, the researcher gradually created groups by distinguishing between similar and different content. Cards that were gathered as a group were re-read and commonalities were written down as “nameplates.” Cards with the same “nameplate” were collected together and placed on a piece of large paper, according to the relationship between the “nameplates.” The cards under one nameplate were spread out around the nameplate to check the contents again. In this way, each card was checked repeatedly about its connection with the nameplate, and some cards were moved to the group of another nameplate, if necessary. Finally, when it was determined that this was an appropriate grouping, the nameplates and all cards were fixed on top of the paper so that they could be viewed at a glance. In the Findings below, the nameplates generated as described above are listed and the study participants' words (translated into English by the researcher) are given as examples.

3. Findings and Discussion

3.1 Learners' views on the designated theme of the paragraph

This section and the following section 3.2 intend to answer Research Question 1. The responses to the first half of Q3, “What did you think of the theme of the paragraph? Was it easy to write? Was it difficult to write?” were first classified into “It was easy” and “It was difficult” groups. About 70% of the total respondents answered that it was easy, and the percentage was higher in Group C, Group B, and Group A, in that order. On the other hand, 27% of all respondents said it was difficult, and the percentage was higher in Group A at 40% (see Table 2).

Table 2

Learners' Views on the Theme

	Group A	Group B	Group C	Total
It was easy to write about.	15 (60%)	13 (72%)	38 (75%)	66 (70%)
It was difficult to write about	10 (40%)	4 (22%)	11 (22%)	25 (27%)
No answer	0 (0%)	1 (6%)	2 (4%)	3 (3%)
Total	25	18	51	94

The responses that gave reasons were further categorized using Ueno's (2018) qualitative analysis described above. As a result, four categories of reasons were

created as to why it was easy, and two categories of reasons why it was difficult.

The reasons why some participants thought the topic easy to write about were as follows:

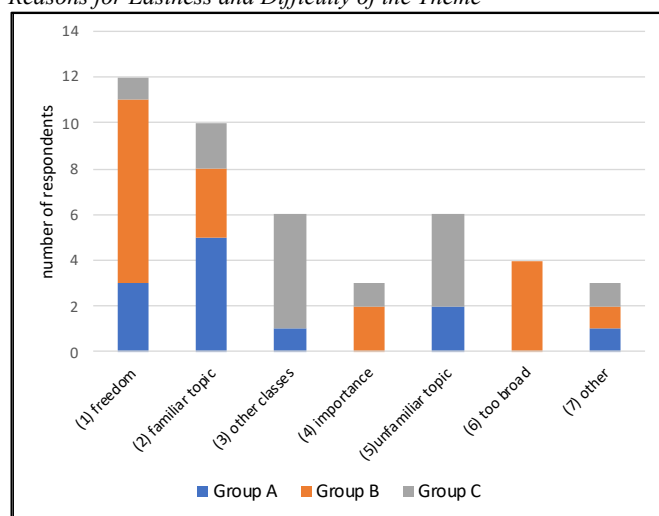
- (1) It allowed me to choose freely from a broadly set theme.
- (2) It was a familiar topic that was relevant to me.
- (3) It was connected to the content I had studied.
- (4) It was an important topic.

Others thought the topic was difficult to write about because of the following reasons:

- (5) I did not know much about the topic.
- (6) It was too broad a theme to choose from.
- (7) Other (responses that did not fit into above categories): “I found the theme difficult because there are so many different opinions on how marriage should be and should be perceived” (C22); “It was rather embarrassing to write my innermost feelings” (A27); and “I am poor at writing in English, so I didn’t think anything special about the theme” (B23).

The distribution of these reasons by group is graphed in Figure 2.

Figure 2
Reasons for Easiness and Difficulty of the Theme



Note. Please refer to the classification above about the reasons (1) through (7). Multiple responses were possible.

As this graph shows, the reasons for the ease and difficulty of the topics varied from group to group. For Group A students the topics were easy primarily because they were relevant to them. On the other hand, many students of Group B thought that the ease of writing was due to the fact that they could freely choose their own topics, and conversely, only in Group B did respondents say that the difficulty was that the topics were too broad and so they could not decide on topics easily. In other words, for many of the Group B students, there seems to be a high level of interest in “choosing the topic,”

i.e., what they write about. In Group C, the most common reason given for the ease of the topics was that the given theme was related to what the learners had studied in other classes. For both Group A and Group C, the reason given for difficulty was that the topic was unfamiliar to them. These points will be discussed in more detail in the Discussion section.

3.2 Topics that learners want to write about

The content of the responses to the second half of Q3, “What would you like to write about?” were categorized in the same way. The topics were divided into the following seven categories. Below are the categories with examples of learner’s voice.

- (1) Reflections on myself:
“I would like to write about myself, not about social issues, etc., as I would like to do some self-analysis” (C15).
- (2) Topics of my interest:
“I think I can write endlessly about my hobby” (A10).
- (3) Cultural issues:
“I would like to write if it were something like giving my opinion on a movie or painting” (C24).
- (4) Topics familiar to us (not about a specific topic of their choice):
“I would like to write about a familiar topic that I can give my opinion on” (C43).
- (5) Social issues:
“I would like to write about poverty and other topics that interest me” (C37).
- (6) Anything but social issues:
“I have written a lot about environmental and social issues, etc. in the past in preparation for entrance exams, so it would be interesting if the topic is something other than those” (B4).
- (7) For-or-against type theme:
“It is easy for me to write something that is either for or against” (C50).

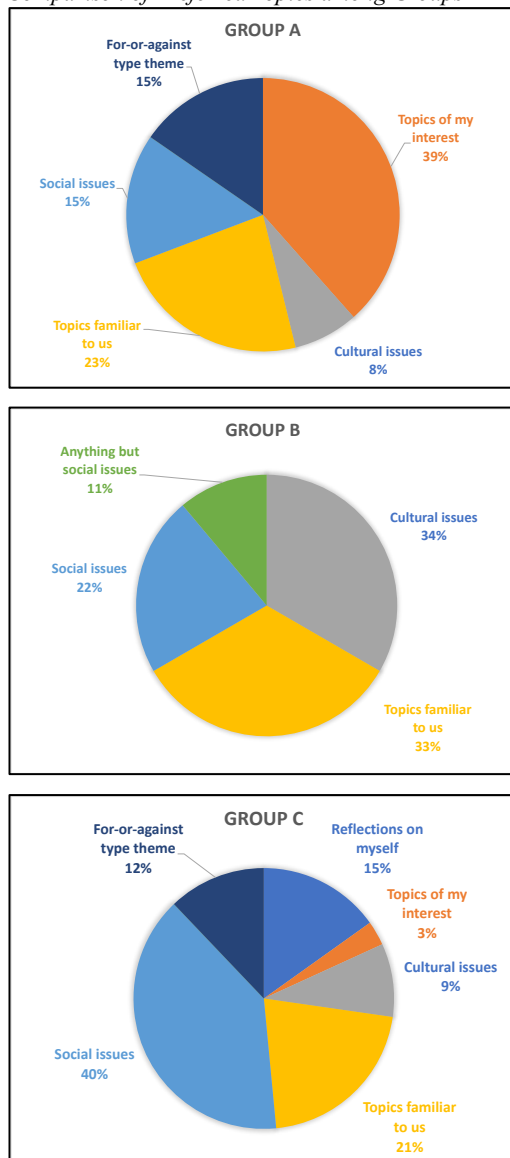
Table 3 summarizes the number of occurrences of these reasons, and Figure 3 shows the differences between the groups in pie charts.

Table 3
Topics that Learners Wanted to Write about

	Group A	Group B	Group C	Total
Reflections on myself	0	0	5	5
Topics of my interest	5	0	1	8
Cultural issues	1	3	3	6
Topics familiar to us	3	3	7	13
Social issues	2	2	13	17
Anything but social issues	0	1	0	1
For-or-against type theme	2	0	4	6

Note. Multiple responses were possible.

Figure 3
Comparison of Preferred Topics among Groups



In Group A, a much larger number of respondents wanted to write about “Topics of my interest” than in the other groups. In Group B, on the other hand, there were no responses that fell into the “Topics of my interest” category, and the majority of the responses wanted to write about “Cultural issues” or “Topics familiar to us.” It is also interesting to note that one respondent clearly stated they wanted to write about “anything but social issues,” while their major was social science. In Group C, conversely, the largest number of respondents wanted to write about social issues. By their departmental requirement, Group C students were required to take courses that discuss such themes as “issues in Japan as a multicultural society,” “the discrepancy between cultural relativism and globalization,” and “the ideals and realities of the SDGs” in their first semester. Group C students’ interest in social issues may have been caused by the content of these classes, as this is the first time that many of them have

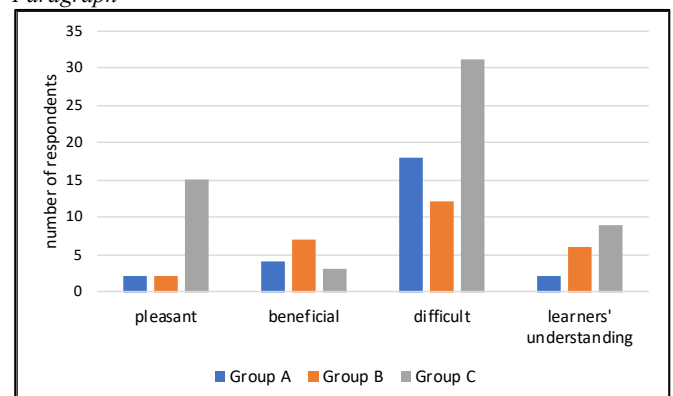
undertaken such a study in earnest after entering university.

The answer to RQ1: Although each group had its own characteristics, as a whole, many learners found the assigned theme easy to write about. The reasons for this were that they were able to choose topics freely to some extent and were able to make connections to their own experiences, values, and studies. On the other hand, there were also learners who said that the theme was difficult to write about, because they were unfamiliar with the theme or the theme was too broad.

3.3 Reflection on the experience of writing a paragraph

This section tries to answer Research Question 2. The content of the responses to Q4 of the questionnaire, “How do you feel about writing paragraphs in English?” were categorized by Ueno’s (2018) qualitative analysis. The intent of this question was to find out how learners felt about learning English writing: in other words, their affective experience. Three major categories that emerged were “pleasant,” “beneficial,” and “difficult,” but another unexpected category was also obtained. It was the “learners’ understanding” of what English paragraphing writing is, which was not directly related to the learners’ feelings. The results are summarized by group in Figure 4.

Figure 4
Four Categories of Reflection on the Experience of Writing a Paragraph



The first category of responses, “pleasant,” was accompanied by a variety of reasons. They include: “It is fun” (A20); “It is quite enjoyable. When I organize my opinions, I feel like my mind is getting organized, too” (C30); “It’s fun. It is interesting to make persuasive sentences when I write” (C36); “It stimulates parts of my brain that are not normally used much” (C29); “It is difficult for me to think of sentences and construct them from scratch, but I feel a sense of accomplishment when I finish writing them!” (C20). Group C was by far the largest group of learners in this category, with 15 out of 51 (29%) in Group C reporting that English paragraph writing was a “pleasant” experience.

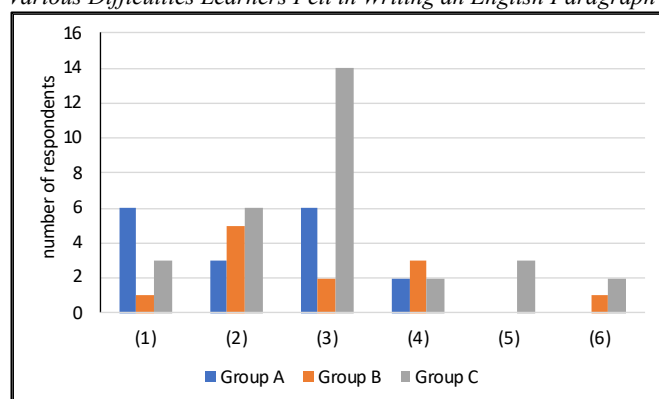
On the other hand, the second category, “beneficial,” was the most common in Group B, with seven out of 18 learners (39%) answered the merits of the activity of writing English paragraphs. Learners found English writing beneficial to them for a variety of reasons: “I think writing

paragraphs is necessary for me to remember how to make basic English sentences” (B6); “It is important. It helps me speak in English, too” (C37); “It is a good learning experience because I use dictionaries for all kinds of expressions when writing” (B25); “I also believe that the more I write, the more I can read” (B4); “I feel that it is beneficial for both thinking about the theme and developing English skills” (B18).

The third category, “difficult,” had the most responses across all groups. Further analysis of the reasons for perceived difficulty generated the following six subcategories.

- 1) It is difficult (with no specific reasons mentioned).
- 2) It is difficult to find the right words and expressions: “I often give up writing all I want to say because it is difficult to translate what I want to say into English” (A3).
- 3) It is difficult to make a proper structure: “I feel a bit depressed when I think that I have to write a paragraph with a proper structure” (C22).
- 4) It is difficult to think about the content: “It is difficult to write reasons and detailed explanations” (C12).
- 5) It is difficult to avoid one-pattern expressions: “My writing style is always the same, and I want to write in other ways, but I don't know how” (C50).
- 6) Other: “I am not good at it. This is because when I write in English, I feel that I have to write sentences that can be understood by a larger number of people (including people from overseas) than those written in Japanese” (B11); “It is very difficult to make it simple. I tend to use difficult words when I think of sentences in Japanese, so it's hard to think of something simple that I can write” (C24); I feel it is not so difficult if I write sentences in Japanese first. If it is in English from the beginning, it is very difficult (C25).

Figure 5
Various Difficulties Learners Felt in Writing an English Paragraph



Note. (1) No specific reason mentioned. (2) It is difficult to find the right words and expressions. (3) It is difficult to make a proper paragraph structure. (4) It is difficult to think about the content. (5) It is difficult to avoid one-pattern expressions. (6) Other.

The distribution of these subcategories for each group is as follows. The most common responses for Group A were in “for no specific reason” and “to make a proper paragraph

structure.” On the other hand, “to find the right words and expressions” was the most common response in Group B, while “to make a proper paragraph structure” was the most common response in Group C. Also, the subcategory, “to avoid one-patterned expressions,” was found only in Group C.

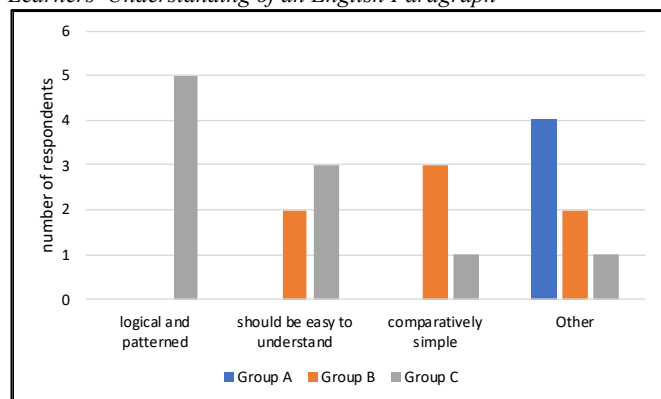
These results could reflect the differences in the English proficiency level of each group. There seems to be a progression from a level where they simply feel it is difficult to write in English, through a stage where they can write in English but sometimes cannot find the right words they want to express, to a stage where they realize that they can speak but need to think about how to structure their writing. Another category observed only in Group C was “To avoid one-patterned expressions,” which may mean that the students in Group C understand and are able to write a paragraph with “proper” structures and expressions according to the textbook, but as they continue to do so, they begin to be bothered by the fact that they can only write in the same pattern all the time. It is probably natural that, as the level of vocabulary, structure, and style increases, the students will find their next challenge.

Finally, some responses expressed the learners' understanding of English paragraph writing. These were again analyzed and grouped into the following four subcategories.

- 1) English paragraphs should be written logically and following a pattern: “I consider it to be written in the sentence pattern of presenting a claim, stating reasons, and concluding with a conclusion” (C7).
- 2) It is important to write sentences that are easy to understand: I think about how to express myself in a way that is easy to understand and easy to write” (C15).
- 3) Writing an English paragraph is comparatively simple: “I feel that writing is easier than writing in Japanese once I get used to the format, because I don't have to go into such depth on a single topic” (C39).
- 4) Other: “I think the only thing different from Japanese writing is the vocabulary I have” (A8); “In this day and age, learning writing is unnecessary because there are many advanced translation applications available” (A27); “I am a bit intimidated by the idea of writing a paragraph in English all by myself, but I feel that if I can use DeepL and other resources, I can enjoy writing” (B22).

As some of the responses in this category provided interesting insights into students' views of English paragraphs, they are discussed in more detail in the next section.

Figure 6
Learners' Understanding of an English Paragraph



The answer to RQ2: In all groups, the most common response to English paragraph writing was that it was “difficult.” The learners felt frustrated because they could not write English sentences the way they wanted to, and they were not yet confident in using the paragraph structure they had learned just before the questionnaire. On the other hand, some learners described English paragraph writing as “pleasant” or “beneficial.” There were differences in the distribution of reasons between the groups.

3.4 Possible meaning of learning English writing for the learners

As response to Research Question 3, this section as a whole discusses the meanings of writing in English for the participants in this study that can be gleaned from the above survey results. The diverse experiences they disclosed can be summarized from several perspectives. The possible pedagogical implications of each meaning/experience will also be discussed.

1) Writing English as an enjoyable and satisfying activity for the learner

Some learners seem to be able to find meaning in writing in a foreign language because they feel the activity gives them learning, the feeling of accomplishment, or sheer fun of writing. To be able to obtain such pleasure, one may need to reach a certain level of English proficiency, but having such role models among peers would be very encouraging for English learners. It would be desirable for teachers/researchers to learn more closely what these learners are getting out of writing English, so that the knowledge could be utilized in teaching and learning of writing.

2) Writing English as a chance to think and express something meaningful for the learner

Other learners may find meaning in writing if they are allowed to write about the topic they want to think about and share with others. Giving students more freedom to choose what they really want to write about might be one of the options for teachers to consider more. As Van Lier (2010) states, learning occurs in the eco-system of the classroom; in order for each learner’s agency to be emerged and exercised,

teachers should provide an environment that “allows and instigates a diversity of manifestations of agency at different levels” (p. 5). It goes without saying that learners need to feel safe in the classroom in order to reveal what they really want to write, so establishing a good atmosphere in class would be quite important.

3) Writing English as a difficult and overwhelming experience

There are certainly not a few learners who do not have confidence in their ability to write in English. For some of them, the experience of English writing may have only negative meaning, as one of the study participants wrote: “I have always been bad at writing in English, so the first feeling I get is that it’s hard” (B23). It seems quite important for teachers to realize that some learners may be affected by such self-negative feelings when they are assigned to write in English. To help them experience success in expressing themselves in English, it may be possible to carefully introduce the guided use of machine translation along with the reinforcement of learning vocabulary, grammar, etc.

4) Writing English as a simplistic task with simplified vocabulary

As we saw in the learners' understanding of English writing (at the end of section 3.3 above), quite a few learners seemed to think that the important thing about writing in English was to follow the pattern taught, demonstrate their knowledge of the rules, and use “easy” vocabulary and structure to avoid making mistakes. They describe it as follows: “It is important to make sure that the grammar is correct and that the sentences say what you want to say in a way that is easy to understand. It is clearly different from a report in Japanese” (B4); “I don’t find it difficult to write a paragraph because I can use my own vocabulary and syntax, and I don’t have to try to understand difficult things like when I am reading a given text” (B3); and “I feel that once you get used to the format, it is easier to write than when writing in Japanese, because you don’t have to go into such depth on a single topic” (C39).

For some students, however, it appears that confusion has arisen because of the (mis)conception that “simple expression” is what they must pursue above all else. One learner wondered, “Is the important thing in English writing to write simple sentences?” (C13). Another learner saw the difficulty of writing in English as follows: “It is very difficult to make the expressions simple. When I think of sentences in Japanese, I tend to use difficult words. It is hard for me to think of something simple and easy to write” (C24).

These statements and sentiments may not be surprising if we understand why such feelings are shared among college students. The following comment explains it well:

When I was in junior high school and high school, I was concerned about how to meet the word count required by the question without making grammatical mistakes. But after entering university, I became more concerned about the structure of the entire paragraph and so on. In addition, since I can use a dictionary at the university, I

have become a little more concerned about choosing more appropriate words (B16).

Her words suggest that the learners have been writing English only for tests where they were not allowed to use a dictionary; in order to get good marks on these tests, it was crucial to use “simple” expressions and vocabulary where they would never make mistakes and have points deducted. As a result, they may be dumbing themselves down by rephrasing complex thoughts and feelings that they could express in their native language into simple ones, rather than looking for the equivalent in a dictionary. Teachers may need to be more aware of the impact that learning English writing for exams has on learners.

4. Summary

This study analyzed the comments of 94 EFL learners at a Japanese university to explore how they perceived the activity of learning English paragraph writing through their responses to an open-ended questionnaire. The results suggest that the learners had widely different perceptions of the preferred writing topic and the writing activity itself, even though they were of a similar age group, belonging to the same university. Another intriguing finding of this study was some learners might share the (probably misguided) notion that what is required in English writing is only to follow the proper format and not make grammatical or vocabulary errors. They believe it is important to use simple syntax and vocabulary for fear of making mistakes. Both the diversity of the learners' reflections and the surprising notions that some of them held underscore the importance of better understanding individual learners through a qualitative approach.

Implications of the results for writing instruction include a renewed awareness of the importance of considering writing topics from various viewpoints. In addition, it may be necessary to rethink the teaching of writing at the university level in order to help students free themselves from writing for tests and, if possible, to awaken the joy of writing.

A limitation of this study is that it is only a questionnaire survey without follow-up interviews. Although participants freely expressed their thoughts in their responses to the questionnaire, some of the interpretations are merely speculation by the researcher. It is necessary to confirm the correctness of the interpretation by interviewing several participants. For future research, the researcher plans to interview some of the participants in the study to confirm or modify the interpretation and to further analyze the learners' reflections.

Notes

¹ Although there is no objective indicator such as standardized test scores for the English proficiency of the students in the three groups, the class teacher (the researcher) felt that there were clear differences based on their responses (percentage of correct answers, time required to answer, number of words written within a time limit, etc.) when using the same material. As for Group C, all students took the TOEIC test just after they entered the university, and their average score is shown in Table 1. About half of the students in Group A took the TOEIC test offered by the university in the winter of their first

year, and their average score is also shown in the table. Group B had no students who took the TOEIC test, partly due to departmental policy.

² See above note.

References

- Abdi Tabari, M., & Wang, Y. (2022). Assessing linguistic complexity features in L2 writing: Understanding effects of topic familiarity and strategic planning within the realm of task readiness. *Assessing Writing*, 52, 100605.
- Casanave, C. P. (2018). Qualitative approaches to studying second language writing. In J. Lontas (Ed.), *The TESOL encyclopedia of English language teaching* (pp. 2461–2467). Wiley Blackwell.
- He, L., & Shi, L. (2012). Topical knowledge and ESL writing. *Language Testing*, 29(3), 443–464. <https://doi.org/10.1177/0265532212436659>
- Kawakita, J. (1970). *Zoku hassou hou - KJ-hou no kaihatsu to ohyo* [The KJ method: Its development and application]. Chuokoron-sha.
- Li, S., & Yu, H. (2023). Effects of topic familiarity on L2 writing processes and behaviors. *International Journal of Applied Linguistics*, 34(1), 348–366. <https://doi.org/10.1111/ijal.12497>
- Lo, J., & Hyland, F. (2007). Enhancing students' engagement and motivation in writing: The case of primary students in Hong Kong. *Journal of Second Language Writing*, 16(4), 219–237. <https://doi.org/10.1016/j.jslw.2007.06.002>
- Merriam, S. B., & Tisdell, E. J. (2015). *Qualitative Research: A Guide to Design and Implementation*. John Wiley & Sons.
- Pavlik, C. (2014). *Burning issues: Advanced level* (T. Shizuka (ed.)). Shohakusha / Cengage Learning.
- Pavlik, C. (2015). *Burning issues: Intermediate level* (T. Shizuka (ed.)). Shohakusha / Cengage Learning.
- Putra, J. W. G., Matsumura, K., Teufel, S., & Tokunaga, T. (2021). TIARA 2.0: An interactive tool for annotating discourse structure and text improvement. *Language Resources and Evaluation*, 57, 5–29. <https://doi.org/10.1007/s10579-021-09566-0>
- Tedick, D. J. (1990). ESL writing assessment: Subject-matter knowledge and its impact on performance. *English for Specific Purposes*, 9(2), 123–143. [https://doi.org/10.1016/0889-4906\(90\)90003-U](https://doi.org/10.1016/0889-4906(90)90003-U)
- Ueno, C. (2018). *Joho seisansh ni naru*. [To Become an Information Producer]. Chikuma Shobo.
- Van Lier, L. (2010). The ecology of language learning: Practice to theory, theory to practice. *Procedia, Social and Behavioral Sciences*, 3, 2–6. <https://doi.org/10.1016/j.sbspro.2010.07.005>
- Winfield, F. E., & Barnes-Felfeli, P. (1982). The effects of familiar and unfamiliar cultural context on foreign language composition. *The Modern Language Journal*, 66(4), 373–378. <https://doi.org/10.1111/j.1540-4781.1982.tb01033.x>
- Yang, W., & Kim, Y. (2020). The effect of topic familiarity on the complexity, accuracy, and fluency of second language writing. *Applied Linguistics Review*, 11(1), 79–108. <https://doi.org/10.1515/applirev-2017-0017>

Applying Chomskyan Perspectives to English Language Teaching

YASUNAGA, Akie*

**Tokyo Denki University
naga15580@gmail.com*

Abstract

Noam Chomsky proposed a language acquisition theory in the 1960s and hypothesized an innate ‘human organ’, the language faculty (Chomsky, 1976, p. 36). He hypothesized that the particular language organ enables all children to acquire effortlessly the syntax of a language without breaching the system. The Principles and Parameter theory or phrase structures have been used to analyze the linguistic elements based on the innatism theory (Chomsky, 1981). The innatism theory has transformed the language teaching methodology into a naturalistic approach from the rule-based teaching proposed by Bloomfield’s descriptive grammar (1933/2001). Cook and Newson (2007) proposed three ways to build basic syntax structures from Chomsky’s perspectives: 1) exposure to primary linguistic evidence to activate available Language Acquisition Device (LAD), 2) employing a learning strategy of direct teaching, and 3) engaging in social exchange attached to interactional routines between a novice and caregivers. This paper presents language instructions, approximating the conditions of the three instructional strategies. Though not definite, the results imply the classroom instructions likely have contributed to transferring external language into internal language in mind..

Keywords: structural linguistics, generative grammar, the language acquisition device, interactions, dictogloss

1. Introduction

Scientific research on SLA and the methodology of TESOL, the fields of applied linguistics, do not have a very long history; they began only around seventy years ago. In the early time of history, foreign language teaching aimed to develop oral conversational proficiency, which started from a practical need to train foreigners serving in the army at times. The need for conversational proficiency, for the first time, led to the application of a scientific methodology grounded upon the Structurist perspectives, incorporating a behaviorist approach, so-called the audio-lingual method. In the 1960s, Noam Chomsky, the Father of Modern Linguistics, proposed a language acquisition theory in his book, *Language and Mind* (1972). He rejected descriptive grammar teaching and hypothesized that language acquisition is concerned with the human brain independent of general learning faculty. The proposition sparked a discussion regarding the dichotomy between theoretical concepts of whether a language is learned or acquired.

Chomsky’s language acquisition theory was sharply contrasted with the structuralist approach. He pointed out that children acquire a language effortlessly without breaching its system, the syntax of a language, in the absence of explicit instruction. The central claim of Chomsky’s hypothesis (1976) was that all humans are born with an innate capacity, called a ‘human organ’ (p. 36), the Language Acquisition Device (LAD) that enables humans to acquire language. He questioned the descriptive grammarian approach, for example, explicit teaching of particular structural patterns by presenting them sentence by sentence. In a nutshell, descriptive grammar teaching based on the Structurist

perspective hypothesized descriptive grammar could be acquired by explicit habit formation, instructing learners to use them accurately. By contrast, Chomsky argued there must be a particular organ that enables children to produce a language without conscious effort. After Chomsky’s proposition, the major approach in TESOL shifted away from exclusive teaching of grammar and error correction, and towards enhancing the exposure to target language, for example, adopting the Natural Approach through comprehensible input (Terrell & Krashen, 1998).

Chomsky’s “[l]inguistic theory is concerned with an ideal speaker-listener in a completely homogeneous speech community” (1965, p. 4); however, he also acknowledges the rise of multilingual speakers. Native speakers’ learning the linguistic syntax, or Universal Grammar (UG) is different from that of learning an additional language. “[A] second language (L2) is effectively an extra tacked on to the first language (L1), like an extension to the back of the house. ... But such an impure state cannot form the core subject-matter of linguistics” (Cook & Newson, 2007, p. 222).

Cook (2002) posits a slightly different perspective from Chomsky’s proposition. He argues that almost all people have multiple grammars in their minds; they use L1 and/or L2 depending on the context, and L2 learners possess the knowledge of more than one language in the same mind, in that they are multi-competence users (Cook, 2002). The theory views second language users as a whole person having multiple grammars, not as one learning an abstraction of the monolingual ideal grammar.

In EFL settings, there is not much research that connects classroom instruction based on Chomsky’s

hypothesis of the LAD or UG. This paper examines how Chomskyan perspectives can be applied to EFL contexts.

Cook and Newson (2007) proposed three types of language teaching strategies, which are likely associated with the activation of the learners' LAD, which include: 1) presenting primary linguistic data as available evidence for utilizing the learner's LAD; 2) giving ample opportunities to connect sound with meaning, for example, imitation and direct teaching, and 3) setting up well-designed social exchange attached to interactional routines. In this paper, I will present why these models likely activate the learner's UG and how these can be applied to language teaching. The following research questions are formed:

- 1) What is the core theoretical ground of LAD and UG?
- 2) How can the language teaching strategies be applied to classroom teaching?

In the following sections, I will first review Chomsky's language acquisition theory, contrasting it with Structural Linguistics. Next, standing on the multi-competence theory (Cook, 2002), I will elaborate on possible language instructions.

2. From Structural Linguistics to Universal Grammar

2.1 The Structuralist Approach

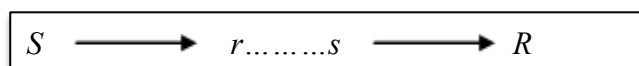
The founding father of modern linguistics, Saussure (1916/1983) posited that linguists set "the study of linguistic structure as his [the linguist] primary concern" (p. 25). He denotes that linguists should connect sound patterns with linguistic signs and the concepts people speak in a community. He stated, "[T]he linguistic signal, being auditory, has a temporal aspect and hence certain temporal characteristics" (p. 103, quoted from Zhou, 2020). Following Saussure's position, Leonard Bloomfield (1933/2001) established the field of descriptive structural linguistics with a behaviorism approach. The teaching approach describes recurrent structural units while connecting the structures with speech utterances (sounds) and practical speech events.

The structural linguistics view of teaching considers the environment as a critical factor in eliciting plausible behaviors. An American psychologist, Skinner (1976), for example, investigated the external conditions, viewing them as possible antecedent causes of subsequent behaviors, that is, when learners were provided with a stimulus from the environment, e.g., receiving rewards and praise, it was considered to accelerate learning. Scientists accumulated empirical data and induced general principles that may lead to the relationship between causes and subsequent consequences, as seen in operant conditioning. Bloomfield invented the famous stimulus-response models for teaching descriptive grammar, (Delprato & Midgley, 1992). Figure 1 describes the behaviorist's method situated in speech events.

The method considers adult intervention as critical for eliciting positive reactions. For example, the parent or caretaker should give praise to the child when they produce plausible reactions because this likely serves as a positive stimulus to reinforce the intended reactions. Conversely,

negative reactions serve to minimize implausible inaccurate utterances. Bloomfield's language teaching is completed with the sequence of stimulus, response, and reinforcement. The approach was partly redressed with the present-practice-produce cycle. Chomsky, however, severely criticized the overreliance of language learning on the external environment.

Figure 1
The Behaviorist's Method of Language Teaching



(S= practical stimulus, r= linguistic substitute reaction, s= linguistic substitute stimulus, R= practical reaction)
(Bloomfield, 1933/2001, p. 25, quoted from Duan, 2017, p. 344)

2.2 Innatism

In *Syntactic Structures* (1957), Chomsky demonstrated that even though a rewritten sentence from the kernel sentence could be grammatical, it could represent different meanings, in other words, "syntax is independent of semantics" (Cook & Newson, 2007, p. 2). On the surface level, the structure is grammatical, but at a deeper level, the structure represents different meanings, as shown in the example sentences below:

- 1) John is eager to please.
- 2) John is easy to please.

According to Chomsky, humans build abstract knowledge of language (syntax), in the individual mind by creating meaning through sounds or texts in the external physical world. The distinction between externalized and internalized languages is important to understand innatism. Chomsky (1986) presented the terms: The externalized language, which is termed E-language, is those in the environment such as sounds, sentences, copulas, and the descriptions of such data. Many studies on sociolinguistics and discourse analysis fall into this category. By contrast, the internalized language, I-language, is an internal property of the human mind, "a system represented in the mind/brain of a particular individual" (Chomsky, 1988, p. 36, quoted from Cook & Newson, 2007, p. 13). I-language is whatever is in the individual mind, namely "a computational procedure and a lexicon" (Chomsky, 2000, p. 119). Chomsky attempted to explore the mind rather than the environment; "[l]inguistics is the study of I-language and the basis for attaining this knowledge" (Cook & Newson, 2007, p. 13).

According to Chomsky, competency refers to "the speaker/hearer's knowledge of his language", and performance "the actual use of language in concrete situations" (Chomsky, 1965, p. 4). Grammatical competence is "the cognitive state that encompasses all aspects of form and meaning and their relations" (Chomsky, 1980, p. 59, quoted from Cook & Newson, 2007, p. 15). According to Chomsky, the competence of grammar which humans are capable of

producing sentences is independent of the situation and the intentions of the participants. The core part of syntax is intact even though the external environment influences the mode and attitudes of the speaker. The ideal speaker-hearer's language is "unaffected by such grammatically irrelevant conditions as memory limitations, distractions, shifts of attention and interest, and errors applying his knowledge of the language in actual performance" (Chomsky, 1965, p. 3, quoted from Cook & Newson, 2007, p. 19). Therefore, even though humans use language for communicating, which was asserted by Hymes (1972), innatism theory concerns the core part of generating language, that is, syntax.

On the other hand, performance refers to the actual use of language, utilizing competence. Errors may occur in actual performance, for example, occasional erroneous utterances of irregular verbs of the past tense, like go-went and eat-ate; however, this is not indicative of not holding knowledge of the language. It is only a performance error (Cook & Newson, 2007).

Chomsky (1993) believes the human mind is capable of transforming externalized physical representations and sounds into mental representations of meaning through the computational system. The quintessence of the computational system, according to Cook and Newson (2007), is a rich network of semantic lexicons the learners have accumulated in their linguistic experience. To put the concept into context, I describe the sentence, "The moon shone brightly in the sky." In the sentence, the word, moon, is stored in the mind as "the moon", and a matching verb, "shone." Figure 1 illustrates the converting process from E-language and I-language through the interface of the computational system. The figure is adapted from Cook and Newson (2007).

Figure 2
The Interface Between E-Language and I-Language

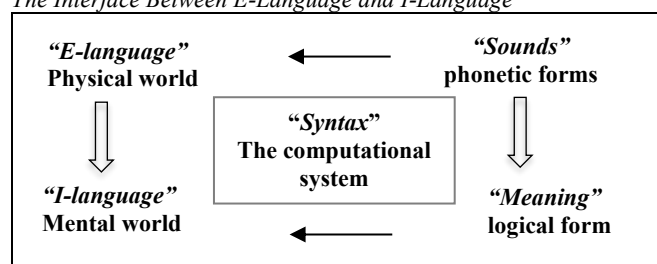


Figure 2 shows the interface between E-language in the physical world, i.e., texts and sounds, and I-language, logical meanings in the mind. The computational system is the core to convert E-language to meaning in the mental world. "[E]ach language can be regarded as a particular relationship between sounds and meaning" (Chomsky, 1972, p. 17).

Bloomfield's descriptive grammar approach was concerned with collecting patterns of structures and describing their regularities. In other words, "a grammar is a collection of descriptive statements concerning the E-language" (Chomsky, 1986, p. 20, quoted from Cook & Newson, p. 13). Chomsky (1993) argued two points: 1) children do not learn language by describing the regularities of the sentence through external stimuli; 2) children have the

innate capacity to connect sounds and texts with meaning and generate new meanings by deploying sounds and texts. He termed this capacity as the human language faculty, the Language Acquisition Device.

2.3 Universal Grammar

Innatism theorists, for example, I-language researchers, posit that language is not governed by descriptive rules or regularities of the structure, but by principles and parameters (P&P model) that every language possesses and every ideal speaker never breaches. Particular principles and parameters are present in every language, such as Korean, Japanese, or French and those are checkable. "The basic assumption of the P&P model claims that human languages consist of principles without construction of specific rules (Cook & Newson, 2007). The occurrences of universals that the speakers of that language never violate are seen in many languages, and they are termed Universal Grammar (UG).

2.4 Is UG learned or acquired?

UG researchers are concerned with principles and parameters that exist in every language, and among which the movement principles that particular languages abide by have been studied in the past. They are principles of how particular linguistic elements move within the sentence. They are called the Locality Principle. I will illustrate them from examples described in Cook and Newson (2007). For example, to form a question, subject-auxiliary inversion occurs, like the question below:

3) Why is he making a mistake?

Past research found that some movement is restricted by the locality principle. If there are two or more auxiliaries within the sentence, the movement is determined by the distance of the moving linguistic elements. For example, sentence 4) has two auxiliaries. To make a question, the shorter distance of the movement can only move, but the longer distance cannot, as seen in question 5).

- 4) The manager could have fired Beth.
5) a. Could the manager have fired Beth?
b. *Has the manager could fire Beth?

At the performative level, the question is whether this principle could be learned by explanation and practice or not, for example, by explaining the rules like sentence 6). Chomskyan perspectives deny the possibility of learning through abstract generalization (Cook & Newson, 2007). It is because most children across the world grow to speak without making a conscious effort to apply rules every time they speak a language.

6) Move a wh-element to the front to make a question.

Another question of acquiring language knowledge is concerned with its use. To Chomsky (1979), "language serves essentially for the expression of thought" (p. 88, quoted from

Cook & Newson, 2007, p. 16). Though researchers may counter that the primary purpose of language is communication, Chomsky considers the property of language lies in expressing inner thought. For example, we may observe everyday humans create new sentences to express unique meanings, like the sentences below 7) and 8). A core characteristic of human language then is to be able to produce an infinite number of sentences through language-specific UG (Cook & Newson, 2007).

- 7) There is grey in your hair.
- 8) Her facial expression was neutral.

Some justifiable evidence, though difficult to validate, is reported about activating LAD. Cromer (1987) reported his study of providing short modular instruction, in which children were given ten example sentence constructions seen in the book, *The Wolf Is Easy to Bite*, for three-month intervals over one year. In the cross-sectional study with students who received no modular instructions, Cromer (1987) found those who received the sentence constructions learned to use a structure faster than those who were not given this slim exposure. In the study, children were given no feedback on the interpretation of the structure, but the modular exposure to the constructions might have helped build rudimental principles.

In this section, I described the dawn of the SLA and TESOL that incorporated a scientific approach. The descriptive grammar method applied a behavioral approach that relied on describing recurrent structural units to make habits of producing accurate language. On the other hand, Chomsky posits that humans have the innate capacity to acquire UG without receiving descriptive instructions on rule-based structures and patterns.

3. Application to English Language Teaching

3.1 The Innatism Approach

Cook and Newson (2007) proposed three strategic approaches. From the innatism theory, learning is about constructing UG of the language, that is the principles and parameters, the movement principles, binding theories, as well as pro-drop parameters. Although precise descriptions of these are beyond the scope of this paper, researchers suggest providing exposure to the linguistic elements is critical, as “the principles and parameters are built in to the human mind” (Cook, 2008, p. 215). Chomsky (1972) points out “[i]t must be recognized that one does not learn the grammatical structure of a second language through ‘explanation and instruction’ beyond the most elementary rudiments, for the simple reason that no one has enough explicit knowledge about this structure to provide explanation and instruction” (pp. 174–175, quotation marks in original, quoted from Cook & Newson, 2007, p. 197)

The essential condition for constructing the knowledge of the language is exposure to positive “syntactic evidence”. (Cook & Newson, 2007, p. 193). Children must hear a language such as the SVO sequence, question forms, or movement principles to construct the basic principles and

parameters (Cromer, 1987; Cook & Newson, 2007). For example, parents and caregivers asking orally, “What have you done today?”, or “We’ll have dinner soon”, may serve as positive evidence in the case of L1. On the other hand, negative evidence refers to explanations or corrections on ungrammatical constructions. Overall, one of the most effective ways to enhance exposure to basic principles is what and how the teachers talk, as Cook and Newson state, “evidence other than positive evidence, by and large, cannot play a critical role; the child must learn primarily from positive examples of what people *do* actually say rather than negative examples of what they *don’t* say (2007, p. 190, Italics in original).

The second strategy is direct teaching, that is, giving on-set feedback on grammatical points. In a natural setting, a mother corrects when they hear their child’s erroneous utterances. Such a condition is realized only when the child produces their ungrammatical utterances first and then a mother or the teacher corrects their sentences. The on-set feedback can also serve as positive evidence for the child (Cook & Newson, 2007). The instructions likely take place when a teacher pushes students’ output and then points out the appropriate syntax or principles that are unnoticed by the students.

The third application is social interactions. Social interactions between the child and the parents are considered to be beneficial for language learning. Even correction and imitation can shape different forms of exchange, and such interactions when attached to social routines can provide recurrent structural patterns (Bruner, 1983, pp. 120–121, quoted from Cook & Newson, 2007), an exemplar of which was adapted from peek-a-boo games, which is described below.

Father: Are you a mucky pup?
 Child: No.
 Father: Yes, you are.
 Child: No.
 Father: Yes, you are.

The child is exposed to the essential structure of question forms through the routinized exchange. Bruner (1983) terms it as “a standardized initially microcosmic interaction pattern between an adult and an infant that contains demarcated roles that eventually become reversible” (pp. 120–121).

These approaches can be applied in foreign language teaching settings, the topic of which I will turn to in the next section.

4. This Study

I adapted the aforementioned approaches and implemented them in general English courses in tertiary education through 2022 and 2023.

4.1 The Participating Students

The participating students were second-year students, majoring in natural science courses with English levels

around A2 and B1 on The Common European Framework of Reference for Languages (CEFR). The students had a relatively good knowledge of grammar but were not very good at listening and speaking. The number of students was 20 students in 2022 and 29 in 2023. The goals of the courses included developing general English knowledge focused on reading and writing abilities. World English 3 (Lee et al., 2020) was used as the textbook.

4.2 Procedures

I will introduce two different forms of classroom instruction used: One is engaging in literate talk while reading the texts together with the students. This section illustrates the instructions given in the English course in 2023. The other is the dictogloss activity, having students reconstruct sentences after listening to short texts. This activity was done in 2022. The former aimed to increase exposure to positive evidence through talking and reading together. The latter was implemented for applying the strategy of direct teaching.

4.2.1 Literate Talk

Literate or literature talk refers to any literacy event in which the teacher talks about the interpretation and comprehension of the written texts, sharing personal experiences and ideas, and bringing participating students' voices to the front (Galda et al., 1995; Blackburn, 2019). Shanahan (2005) reported evidence-based practical reading instruction. In the report, beyond vocabulary teaching and reading fluency, literate talk was the strategy essential to connecting oral language with literacy development. It is talking about the text together with the students during reading.

I will describe a unit that used a TED Talk video in the textbook, titled "*Can You Really Tell If a Kid is Lying?*". In the video a researcher, Mr. Kang, presented his invention, transdermal video images that can monitor people's blood flow, thereby detecting the mood and stress level of the individual. His research investigated how difficult to detect children's lies. However, he found a subtle blood change occurs when they are telling a lie. He described that mind-reading abilities and self-control are important for good lying.

During the talk, I shared the video subtitles putting on the slides. By doing so they could receive both aural, oral, and written language simultaneously. This helps them refer to the slides whenever confused. The literate talk also aimed for the students to reflect on what they have learned and describe it in their reflection sheet, summarizing what they heard during the talk. Below I show the part of the transcript the key concept I focused on during reading and the excerpts of the teacher talk. Figure 3 shows the image.

And good lying requires two key ingredients. The first key ingredient is theory of mind, or the mind-reading ability. The second key ingredient for good lying is self-control. As it turns out, these two abilities are also essential for all of us to function well in our society. In fact, deficits in mind-reading and self-control abilities are associated with serious developmental problems,

such as ADHD and autism (quoted from the transcript in the TED Talk).

Excerpt 1

-
- | | |
|---|--|
| 1 | T: Reading the mind, what is it? |
| 2 | S: (Silence) |
| 3 | T: Mind reading ability is to understand how other person feels, and interpret their feelings okay? ... She feels not good if I ask her to speak English in the class? |
| 4 | T: So if you talk with other persons and if he is in a bad mood, you might be careful about what to say. ..., right? Do you understand? |
| 5 | S: (shook their head) |
| 6 | T: Okay, also children may know if their mother is easy to trick or not. She will believe me if I say (unheard)... |
| 7 | T: Or my teacher will believe me if I say the train has been delayed because of an accident? The teacher will believe me. ...So, to understand if I say something, he or she thinks accordingly, okay? |
| 8 | T: Imagine if you cannot read minds if you cannot read other people's minds, it is very serious. It is associated with 統語失調症、or something like that. |
-

*T: Teacher, S: Students

In the above instances, the teacher tries to pin down the phrase, "mind reading ability," by giving concrete examples of children telling a lie or a student telling a lie to make the concept easier to discern. The subject, "you" is repeatedly used in the discourse with if-clauses. The discourse is directed towards the students to enhance their engagement.

Excerpt 2

-
- | | |
|----|---|
| 9 | T: What is self-control? Why is it important? It means that you regulate your emotions, you don't show your emotions. |
| 10 | T: ... I don't like her, I don't want to talk with her, but you hide, 隠す, your emotions. So she will not feel bad. 他の人の心が害しないように, you hide your emotions. Right. |
| 11 | T: So, I'm very tired, but I hide my emotions. Good adults can do that. So these are important, So why are these, self-control and mind-reading abilities, important? |
| 12 | T: What is the association? Why is self-control important? So, If you tell lies, you try not to be detected by others. You may be worried if she may notice my lie or not. You are worried, but you should control your worries and emotions. So you pretend as if you are telling the truth, and confidently say Right? Okay? |
-

In the above excerpt, some Japanese words are inserted; it is because the teacher checks students' comprehension

during the flow of the talk, and she sometimes tries to make the concept explicit by using Japanese. In line 12, the teacher gives concrete examples and uses minimal sentence constructions, like SVO and SVC sequences with if-clauses to make the talk comprehensible.

Figure 3

The Image of the Classroom Interactions



4.2.2 Dictogloss

Dictogloss is an activity in which students need to interpret the meaning of what they hear and construct the texts while focusing on form. Quite different from dictation activities, which require students to write minimal words in the blank, such as nouns and verbs, dictogloss requires students to delve into their existing language knowledge and lexicons and test their hypotheses. They check whether they express the meaning accurately, coherently, and appropriately while working together with their peers (Swain, 2001). It was reported that during the collaborative talk in the text construction phase, the students shifted their attention to forms and vocabulary, and it was evidenced in the language-related episodes (LRE) (Swain & Lapkin, 2001).

I will describe two kinds of topics done for the dictogloss from the textbook: One is describing the concept of “cognitive bias” presented in a dialogic style. In the talk, the expert explains the bias is based on our previous assumptions. The second topic was about extreme weather events, in which a professor described four different types of extreme events, informing a list of facts and information about the location, and durations of the events. Both are expository talks; however, the former describes human assumptions; by contrast, the latter describes facts on environmental issues. Part of the transcripts are listed below.

...That’s right. For example, *the halo effect* is one cognitive bias. When we think a person or thing is attractive, our minds believe that he, she, or it has other skills or benefits. This feeling can have an effect on what we do. For example, think about advertisements on TV. The actors are

usually attractive people, right?... We find the actors attractive, so our minds tell us that the products they’re advertising are probably good quality or useful. (quoted from audio transcripts in Unit 2, World English 3).

When the weather is much worse than usual, we call it an extreme weather event or EWE. Often, these events cause a lot of damage. There are several ways the weather could be worse than usual. For example, the weather event may be out of season. Snow in Canada in January would probably not be an EWE, but a lot of snow in July probably would be. Or the weather event could be something that usually doesn’t happen in a place. (quoted from audio transcripts in Unit 3, World English 3)

Dictogloss was done with a group of four students. After they listened to the texts, they clarified the meaning in Japanese and constructed the appropriate sentences. The first listening is for concept building, and the second is for collaborative sentence construction, and the third is for checking the accuracy of their texts. Finally, the groups wrote their sentences on the whiteboard and got feedback from the teacher. The feedback was given on the forms and this was done as a whole class activity.

4.3 Data Collection

The students submitted their reflection papers after finishing the unit above in 4.2.1. One focused question is shown below.

What is the mind-reading ability and self-control?

To answer, the students must summarize what they heard during the teacher talk because the original transcript does not explain the question directly. The deadline for the submission was set from 1 week to 10 days after the unit.

Out of 29 students who registered in 2023, 27 students submitted the paper. Fifteen answers contained some part of the talk during the class and twelve answers copied the part of the original transcripts listed in 4.2.1. Because the students were given a week or so to formulate their answers, it seemed that they included some information from online sites and enhanced their text quality. Table 1 shows the collected linguistic data from the reflection sheets.

The results show that many students simply copied part of the original transcripts without paraphrasing or summarizing. However, more than half of the students either summarized or created the meaning from what they heard during the teacher-student literate talk. Though some contained erroneous sentences, the text well summarized the meaning during the classroom interactions. This illustrates that those students succeeded in constructing the meaning in the mind. It was also found that some students researched online sites and enhanced their text qualitatively. I researched online however exact words students used were not found,

which implies that they seemingly included some information to enhance their writing.

Table 1
The Students' Linguistic Data

Features	Cases	Excerpts of the Text
The transcript is copied.	12	"The first key ingredient is theory of mind, or the mind-reading ability. The second key ingredient for good lying is self-control. As it turns out, these two abilities are also essential ..."
The meaning is discerned.	10	"... the ability to understand other people's feelings and thoughts. Self-control controls behavior according to the other person's expectations and helps hide appropriate information when deceiving...."
The text is enhanced by online research	5	"mind-reading ability refers to the capacity to understand and predict others' thoughts, intentions, beliefs, and desires. It's like being able to read what's going on in someone's mind. This ability is crucial ..."

As for dictogloss, out of 20 students, with some absent students, three groups constructed the sentences. The results, shown in Table 2, reveal, though both texts have the same difficulty level of English, that the text told in a dialogical style was easier to discern the meaning. On the other hand, fact-loaded texts made it difficult for the students to create mental images of the meaning. This indicates particular styles were difficult for students to create meaning, which is the main force to trigger language production.

Table 2
The Results of the Dictogloss

Topics	Average word counts per group	Example sentences
Cognitive biases	15	"The way we assume and understand things just by our own knowledge."
Extreme weather events	none	none

5. Discussion

Chomsky's UG theory substantially changed EFL pedagogy. The core characteristic of UG is that language is not learned by conscious practice of the patterns and sentence structures, but humans can discern logical meaning from linguistic data and sounds from the environment. The meaning in the mind is arranged in accordance with the principles and parameters specific to a particular language. Children require little effort to connect linguistic data to logical meaning. The meaning is then stored in their mind. Therefore, Chomsky hypothesized humans have a biological organ, called LAD. These are the core features of LAD.

Next, when it is assumed that humans are equipped with innate biological organs, the primary condition to learn language is to increase exposure to positive linguistic experience. Imitation and direct teaching are other strategies, and having routinized social exchanges, which entails dialogic interactions is equally important.

The study found that 15 out of 27 students could apprehend the major part that they heard and observed in the literate section, and they could construct linguistic data to express the meaning. Chomsky states that humans can transform from E-language to I-language through their computational system. The meaning is kept in the mind.

In the study; the teacher used dialogic reading, that is, talking, asking, and reading together, to ensure a social place in which the teacher provided increased exposure qualitatively to positive linguistic evidence. The basic structures were repetitively used. The teacher modified the language to be comprehensible, and the use of gestural, and visual (facial expressions) seemingly helped comprehend the meaning.

The study implies that possibly, dialogic talk made it easier for the students to construct the mental image. In the dictogloss, the students could reconstruct the meaning for the topic of cognitive biases only. As seen in the transcript, the text has directives, such as "you" and inclusive subject, "we." In a similar vein, during the literate talk, the teacher used the same strategy of directly addressing the students, with a pronoun subject, "you." These engagement markers are commonly used in personal talk. Hyland (2012) identified these discourse markers show the writer's attitudinal dimension in academic discourse. They show the ways writers present themselves and convey their judgments, opinions, and commitments. In addition, the teacher's questioning during the talk may also had an impact on increasing engagement and the students temporally focused on the topic asked.

By contrast, the texts about extreme weather events are loaded with unrelated information, such as the names of the places, Canada and Australia, and the figures describing the durations of the events. The speaker does not demonstrate a personal stance, and she just reports the facts and terminologies. Possibly the kind of discourse can be one-way talk loaded with bits of facts, and this made it difficult for students to construct the mental image of the talk.

The study indicates that the association between E-language and I-language and their transformation can have a significant influence on second language learning contexts. To conclude, the teacher's talk, providing ample linguistic evidence, appears critical to ensure such conditions, including the way we talk to our students. Chomsky (1981) considers three types of evidence are important; "positive evidence (SVO order, fixing a parameter of core grammar, irregular verbs, adding a marked periphery)" (pp. 8-9). Cook and Newson (2007) state while hearing, "the occurrence of particular sentences in the speech children hear tells them which sort of language they are encountering and so how to set the parameters" (p. 190). It seems that one such place to provide powerful positive evidence can be the whole class

interactive sessions initiated by the teacher. During the session, the teacher affords essential parameters and SVO constructions. This might be more powerful than engaging in common classroom activities such as paired communication tasks between students, who need linguistic support, and solitary reading and writing tasks in search of answers to the comprehension questions. In the dictogloss activities, the student groups shared their sentences with the whole class and received feedback from the teacher. The students paid close attention to their sentence constructions, such as SVO sequences. Giving the on-set feedback is a strategy of direct teaching, and this also provides the students with additional positive evidence.

6. Conclusion

The paper described the core theoretical ground of Chomsky's language acquisition theory, the transformational process from E-language to I-language through LAD, by contrasting the theory with descriptive grammar. Cook and Newson (2007) suggested three ways of applying UG perspectives to second language teaching, a) providing ample positive evidence; b) applying direct teaching; and c) setting up well-designed interactions attached to social routines.

The study incorporated the literate talk for the whole class interactions and collaborative dictogloss activities. The study found that the teacher talk during the literate talk possibly helped students internalize the concepts and likely promoted the activation of the transformational process from E-language to I-language.

References

- Blackburn, M. V. (2019). Literacy teaching and learning in school as polyphonic. In D. Bloome, M. L. Castanheira, C. Leung, & J. Rowsell (Eds.), *Re-theorizing literacy practices: Complex social and cultural contexts*. (pp. 115–125). Routledge.
- Bloomfield, L. (1933/2001). *Language*. Foreign Language Teaching and Research Press.
- Bruner, J. (1983). *Child's talk*. Oxford University Press.
- Chomsky, N. (1957). *Syntactic structures*. Mouton.
- Chomsky, N. (1965). *Aspects of the theory of syntax*. MIT Press.
- Chomsky, N. (1972). *Language and mind* (Enlarged edition). Harcourt Brace Jovanovich.
- Chomsky, N. (1976). *Reflections on language*. Temple Smith.
- Chomsky, N. (1979). *Language and responsibility*. Harvester Press.
- Chomsky, N. (1980). *Rules and representations*. Blackwell.
- Chomsky, N. (1981). Principles and parameters in syntactic theory. In N. Hornstein & D. Lightfoot (Eds.), *Explanation in Linguistics* (pp. 32–75). Longman.
- Chomsky, N. (1986). *Knowledge of language: Its nature, origin and use*. Praeger.
- Chomsky, N. (1988). *Language and problems of knowledge. The Managua lectures*. MIT Press.
- Chomsky, N. (1993). A minimalist program for linguistic theory. In K. Hale & S. J. Keyser (Eds.), *The view from Building 20: Essays in honor of Sylvain Bromberger* (pp. 1–52). MIT Press.
- Chomsky, N. (2000). *The architecture of language*. Oxford University Press.
- Cook, V. J. (2002). Language teaching methodology and the L2 user perspective. In V. J. Cook (Ed.), *Portraits of the L2 user* (pp. 325–344). Multilingual Matters.
- Cook, V. J. (2008). *Second language learning and language teaching* (4th ed.). Hodder Education.
- Cook, V. J., & Newson, M. (2007). *Chomsky's universal grammar: An introduction*. Blackwell Publishing.
- Cromer, R. F. (1987). Language growth with experience without feedback. *Journal of Psycholinguistics Research*, 16(3), 223–231.
- Delprato, D. J., & Midgley, B. D. (1992). Some fundamentals of B. F. Skinner's behaviorism. *American Psychologist*, 47(11), 1507–1520.
- Duan, S. (2017). Bloomfield's concept of meaning. *Journal of Language Teaching and Research*, 8(2), 343–348.
- Galda, L., Shockley, B., & Pellegrini, A. D. (1995). *Talking to read and write: Opportunities for literate talk in one primary classroom* (Instructional Resource no. 12). National Reading Research Center.
- Hyland, K. (2012). *Disciplinary identities: Individuality and community in academic discourse*. Cambridge University Press.
- Hymes, D. (1972). Competence and performance in linguistic theory. In R. Huxley & E. Ingram (Eds.), *Language acquisition: Models and methods* (pp. 269–293). Academic Press.
- Lee, C., Johannsen, K. L., & Chase, R. T. (2020). *World English 3* (3rd ed.). National Geographic Learning.
- Saussure, F. D. (1916/1983). *Cours de linguistique générale. English version: Saussure, F. (1983). Course in general linguistics* (R. Harris, Trans.). Duckworth.
- Shanahan, T. (2005). *The national reading panel report: Practical advice for teachers*. The North Central Regional Educational Laboratory, Learning Point Associates.
<https://files.eric.ed.gov/fulltext/ED489535.pdf>
- Skinner, B. F. (1976). *About behaviorism*. Vintage Books.
- Swain, S. (2001). Integrating language and content through collaborative tasks. *Canadian Modern Language Review* September, 58(1), 44–63.
- Swain, M., & Lapkin, S. (2001). Focus on form through collaborative dialogue: Exploring task effects. In M. Bygate, P., Skehan, & M. Swain (Eds.), *Researching pedagogic tasks: Applied linguistics and language study* (pp. 99–118). Routledge.
- Terrell, D. T., & Krashen, D. S. (1998). *The natural approach: Language acquisition in the classroom*. Prentice Hall Europe.
- Zhou, F. (2020). *Models of the human in twentieth-century linguistic theories: System, order, creativity*. Springer.

Choosing Statistical Analyses for Small Samples

JONES, Marc*

**Toyo University
jones056@toyo.jp*

Abstract

Small samples are frequent in language teaching research due to participant recruitment and attrition problems. Additionally, some types of quantitative methods require sample sizes beyond the size of a usual intact class, making frequentist analysis of work with single classes problematic due to underpowered analyses. However, if little or no research exists that answers a particular research question, is it logical to prolong the existence of such a gap in the literature, or should research be carried out with small samples? This presentation provides a deeper rationale for the analysis techniques used in an already existing study (Jones & Blume, 2022) which had a very small number of participants, but which necessitated a quantitative analysis due to the logic of the research questions. The rationale for a use of Bayesian methods is discussed, along with limitations of Bayesian analysis, such as researcher and reviewer unfamiliarity. Furthermore, use of different linear regressions such as ANOVA, GLM and GLMM are discussed with regard to controlling for individual differences and environmental factors in studies.

Keywords: Bayesian analysis, GLMM, null hypothesis statistical testing

1. Introduction

Applied linguistics has many studies with small samples, perhaps because researchers are less well funded than those in the hard sciences and therefore cannot always afford to compensate participants at the level of, for example, clinical trials. However, small samples can lead to underpowered research, particularly with frequentist statistics. There are several choices facing researchers regarding the analysis of data in quantitative studies (and the quantitative data in mixed methods studies). In this article, the choices made in regard to the analyses in Jones and Blume (2022) are examined.

2. Small samples

One of the main problems in research in applied linguistics may be participant recruitment, especially when one considers that commonly recommended sample sizes for t-tests are $N > 32$ and for chi-squared tests are $N > 50$. These are not the types of sample that correspond to typical practical language classes, and therefore recruiting from one's own students can still result in coming up short. Several issues related to participant recruitment and attrition are noted in McKinley and Rose (2017). There are several contributing factors surrounding recruitment difficulty. If recruitment of volunteer participants from outside one's own classes is necessary, the fact is that rapport has not been established and therefore there is no trust-based relationship. This means that informed consent is not straightforward for potential participants because they do not know whether the researcher is likely to keep the promises made in the information sheet accompanying any consent forms, particularly in the case when ethics board approval is not required by the institution. In addition, students are frequently busy with other classes

when researchers have the time to conduct their studies, and lunch times are also far from ideal. There is also the fact that if researchers ask students to give up their time, they are potentially giving up the potential to earn money in part-time work. In sum, there is usually no inherent benefit to students in participating in research unless there is some kind of outcome that they desire, such as, for example, test practice for an upcoming standardised test.

The problems of recruitment listed above are compounded by those of participant attrition. In-person attendance problems can also contribute to participant attrition, because illness can be as difficult to predict as it is to prevent. Inconvenience for participants is also a contributing factor. Participants may agree to research, providing 'informed consent' without actually paying sufficient attention to what is required of them. Furthermore other unforeseen problems may lead to attrition.

However, consideration needs to be paid to whether research should be carried out at all with small samples if no research exists on the topic. Questions researchers may need to ask themselves are whether there is a possibility that a reasonable conclusion can be reached. If not, can comparison be made between groups or samples? Can data be estimated or simulated the data? If none of these questions can be answered with affirmatively, is recruitment of further participants reasonable in the time available? Is a suitable analysis for the data possible?

Plonsky and Oswald (2014) assert that "quantitative L2 research produces substantially larger effects than those in many other fields" (p. 890) but this could be due to L2 research's small samples, large effects required with them, and publication "bias toward studies with statistically significant findings" (Plonsky & Oswald, 2014, p. 891). One

way of avoiding overly small samples in applied linguistics research is by using power analysis tools such as G Power (Faul et al., 2009). This software ensures that a sample size for sufficient detection of given minimum effect size is known prior to commencing data collection and therefore, as mentioned above, whether further recruitment is required after the first round of data collection.

Below, the reasons for existing problems with small samples due to frequentist statistics are explained and alternatives in the form of Bayesian analyses provided.

3. Misuse of frequentist statistical analysis

Null hypothesis statistical testing (NHST) is the process of analysing data to test whether the researchers' hypothesis (often referred to, somewhat confusingly for those new to quantitative research, the *alternative hypothesis*) or the *null hypothesis* (the hypothesis that the intervention has no effect) is supported. One danger here is that hypothesis 'support' is different to a hypothesis being 'correct'. When researchers' alternative hypotheses are supported by their analyses, bold claims are sometimes made on the basis of a value that is arbitrarily selected, as explained in this section. NHST is the norm in applied linguistics research yet there are issues with them relating to both power and also the way that interpretation of statistical values are interpreted.

The way that p values tend to be used is as a measure of significance, in that they should be lower than a chosen alpha. Alpha can be simplified here as the acceptable probability of a false positive result, but it actually means the acceptable probability that a more extreme test statistic in the direction of the alternative hypothesis may be observed if the null hypothesis is true. A p value of ≤ 0.05 being a typically accepted value for significance in language teaching research. NHST is problematic because p values decrease as the sample size increases. Thus, it is viable for research projects with sufficient resources to recruit volunteers to participate in a study until an acceptable p value for significance is achieved, which is known as p-hacking. Trafimow et al., (2018) argue that despite using ever smaller p values in the psychological sciences, that this does not solve the problem of the use of an alpha for NHST, and that decisions about evidence supporting the theoretical accuracy of a hypothesis should be made over a number of studies carried out by different independent groups of researchers instead of on the basis of a single study.

Alternatively, applied linguistics researchers can diverge from the status quo and forgo NHST. The dangers of NHST are outlined in Trafimow (2024) in that alpha values are entirely arbitrary (although it is possible to make meaningful decisions regarding alpha levels regarding sample size). If one is conducting analyses using frequentist statistics, knowing what the p value represents is important. It is not simply 'significance' but the probability that the result found is due to chance. Therefore, 0.05 is perhaps too high for medical research NHST, but is conversely too low for an exploratory analysis of a small sample. That is, that a drug found to be safe having a 5 per cent probability of being unsafe is obviously far too high, but restricting research on a teaching intervention's efficacy to a 5% chance of a false

positive test seems overly cautious. Although calculating the sample required to detect a desired effect size at the p-level of one's choice is possible in GPower (Faul et al., 2009), researchers need to actually reflect upon these values, and in particular, if they use frequentist statistics, to consider the use of p values and apply common sense to the size of their chosen alpha.

Furthermore, due to the binary nature of NHST, studies are either 'successful' in confirming or negating hypotheses. Such an approach is misguided when studies are novel or are investigating cutting-edge research questions. The position put forward in this article is that NHST has no place unless exploratory studies have already been conducted. Given that p values are used to provide a benchmark value for confirming null hypotheses, it can be stated that instead of using p values as a decision trigger parameter, a nuanced view should be taken, particularly when considering effect and sample sizes. However, Bayes Factor Analysis (BFA) may be even more effective in exploratory analysis, because a probability index of how well data fits one of two models can be obtained.

4. Bayesian analysis

Bayesian data analysis is philosophically different to the frequentist statistical analysis used to conduct NHST. Frequentist statistical analysis assumes that the probability of an occurrence is based upon frequency. Conversely, Bayesian statistics is not based upon simple frequency but upon prior evidence collected to support one model in comparison to another. Previous data collection (even in other studies) can inform Bayesian priors, a coefficient that informs the model comparison (Zondervan-Zwijenburg et al., 2017). Mackey and Ross (2015) advocate use of Bayesian techniques, which they state are "optional when researchers are testing hypotheses predicated on grounded theoretical arguments, and in the present illustration, for testing framework-driven operationalizations of proficiency" (p. 329). However, they do not mention the strengths of Bayesian analysis for exploratory work in the early stages of theory development, which, arguably constitutes a large amount of SLA and speech learning work.

5. Case study: Jones and Blume (2022)

In this section the data analysis decisions taken in Jones and Blume (2022) are described from an emic point of view. The article in question describes the quantitative part of a study exploring whether there is a reason to maintain the status quo in English language teaching of using the prestige varieties of English in listening for perceptual phonology acquisition. By the term 'prestige varieties' we are describing the prestige that has been traditionally accorded to them, essentially the varieties prevalent in the colonial mercantile cities of predominantly Anglophone societies (Rubdy, 2015). In contrast, the set of inputs we used as an alternative were by speakers of Global Englishes, the varieties of English used outside the countries in Kachru's (1984) inner circle, and frequently by non-white people.

The rationale behind the study was that while there is

rhetoric behind use of prestige varieties as the 'default' listening input, there is a shortage of empirical work. Additionally, prior work on L2+ much of the vowel acquisition with Japanese learners of English has been laboratory based (Nishi & Kewley-Port, 2007, 2008), rather than using students in intact classes, thus potentially lacking ecological validity.

The study was undertaken with a grouped pretest-posttest design, which allowed for gains (or in fact, losses) in perceptual ability to be measured easily. Learners of English with L1/dominant language Japanese (N=16) undertook a pretest of the vowels the vowels /æ/, /ʌ/, /ɔ:/ and /ɜ:/, after which they were paired by score and each pair was split, placing them into opposing groups. Both groups took 5 online lessons using TED talks, training the vowels /æ/, /ʌ/, /ɔ:/ and /ɜ:/ in consonant-vowel-consonant word contexts, the words in utterance contexts, then the utterances in the context of a video edited for length. One group, (n=8) received input from prestige varieties, and the other (n=8) received input from Global Englishes speakers. Due to the exploratory nature of the study, not only was the effect overall investigated but also the effect at the individual vowel level.

5.1 Data analysis decisions

The study in Jones and Blume (2022) was designed to investigate whether there was an effect in use of either prestige variety (i.e. so-called 'native speaker' varieties of English) on comparison to Global Englishes varieties (Galloway, 2013) for perceptual vowel acquisition. The study was designed in such a way that gains in perceptual ability could be measured and analysed in a straightforward way. The intuitive decision was to use paired t-tests, because this is frequently the orthodox way to compare the performances or developments of two matched groups in different conditions. The unorthodox decision to use t-tests rather than the Wilcoxon Signed Rank test or Mann-Whitney U-test was due to the difference in statistical power affordances. While the latter two tests assume non-normal distribution, and a t-test assumes normal distribution, violation of these assumptions is possible when the means behave as normal, and because the t-tests were conducted with Bayesian analysis using MCMC sampling, the simulated sampling can be assumed to be sufficiently robust. Additionally, where a Pearson p-value tends to be taken with t-tests, Wilcoxon and Mann-Whitney tests, Bayes factor analysis allows for a view of how well the data fit the model.

The models in the t-tests were actually that prestige varieties resulted in higher gains than Global Englishes ($PV > GE$) and that prestige varieties did not result in higher gains than Global Englishes ($PV \leq GE$), and these tests were undertaken for all four vowels tested overall, and for the four individual vowels tested. Attentive readers may notice that this is similar to NHST, however, the difference is the purpose and use of benchmarking values. The Bayes factor is not used in the same way that p values are used in NHST, but are indicators of the probability that the model fits the data.

As stated above, the study had a total sample of N=16, with two groups, each of n=8, therefore, while it was possible

to test this as a hypothesis, because the central limit theorem does not hold in Bayesian theory, it would be unwise due to this being exploratory work to investigate the use of PV and GE in an ecologically valid setting, i.e. outside of a laboratory setting. Additionally, warnings on Bayesian hypothesis testing are given in Schad et al. (2023). The question regarding data analysis was to what extent the evidence supports either model. The analysis showed no support for $PV > GE$, therefore we concluded that no serious support could be given for the theory that prestige varieties are better models for perceptual vowel acquisition from audiovisual media.

Beyond whether data analysis of the pretest and posttest, understanding the factors contributing to gains (or losses) in perceptual acquisition were required. The most commonly applied linear regressions to analyse such factors are analysis of variance (ANOVA) and analysis of covariance (ANCOVA). However, more detail of interactions can be taken using a Generalised Linear Mixed Model (GLMM), which use Monte-Carlo Markov Chain (MCMC) sampling of data.

GLMM affords scrutiny of the data and the interactions of the variables. In using MCMC to sample the data and using the variables in these samples to create models and compare these models' interactions, contributing factors to the behaviour of variables in the data set can be observed. Calculating BF, allows researchers to understand the probability of the model fitting to the data. As the p value is merely an expression of the probability of gaining a more extreme statistical score under a null hypothesis, the BF provides researchers greater confidence in the factor's relationship to the data as a whole rather than as a probability of comparison of two means reaching an arbitrary benchmark. This is not to say that Bayesian GLMM are a panacea for research; there is a learning curve in the use of Bayesian analyses in general, and GLMM generally require the ability to use programming languages such as R, Python or Julia for complex calculations. In spite of the difficulties, the depth of analysis with meaningful results is of greater benefit to researchers than the ease of using preset analyses without considering their suitability. Above all, the best analysis for the data at hand is required, and in order to make the decision about what is best, alternative analyses also need to be considered.

The main reason for choosing to proceed on the work for Jones and Blume (2022) was that there was a lack of empirical work on the use of GEs for L2+ acquisition and also a lack of work in a classroom setting on vowel learning. Pilot studies had also been conducted at the tail end of the COVID pandemic conditions, resulting in sample attrition and extremely noisy data, although the opportunity for fine-tuning processes was possible. However, because the participants were my own students, as part of an intact class, I understood from the outset that steps were needed to mitigate any problems that were likely to occur. This being the case, proceeding with a methodology for small samples was the plan from the outset, and qualitative data was also collected to allow for triangulation. Despite the advanced planning, the study still had room for failure, such as students becoming tired or ill and becoming unwilling or unable to complete the

work required to allow for useful findings. However, if frequentist statistics had been the only resource at my disposal, I believe that the likely outcome of the study would have been failure due to lack of statistical power and the lack of confidence in making inferences from descriptive statistics alone.

5.2 Alternative analyses

The reasons for using t-tests with paired samples is one of orthodoxy, however, I violated one of the rules in that t-tests assume data is normally distributed. Our data was very much not normal. However, given the uniform lack of normality across all of the groups in the different t-tests, it was assumed that this would not be a significant problem. Additionally, because the t-tests were being used for exploratory purposes rather than for NHST, and the use of Bayes factors to gauge how strong the evidence was for different models meant that a nuanced interpretation of the data analysis was possible, i.e. there was no binary true/false hypothesis testing, but instead a consideration of whether an existing supposition not yet tested empirically held up to scrutiny.

However, alternative non-parametric tests are available, such as the Wilcoxon ranked sign test Wilcoxon (1945), which is also available for Bayes factor analysis (Barch & Chechile, 2023) and the Mann-Whitney U test (Mann & Whitney, 1947). The downside of these non-parametric tests is that they have lower statistical power, and therefore with small sample sizes there is a greater chance that evidential strength goes undetected.

As mentioned above, ANOVA and ANCOVA are popular ways of analysing factors contributing to learning gains. Other alternatives include the ordinary Generalised Linear Model (GLM). The GLM does not provide quite as much detail as GLMM, because the GLM is an analysis of interactions between factors without random effects. Additionally, a common linear regression is also possible, but again lacks the granular detail of the GLMM.

For a frequentist alternative to MCMC, bootstrapping (Larson-Hall & Herrington, 2010) helps create robust statistical power in a similar way to Markov-chains, which resamples smaller samples to increase power. As an additional option, complex analyses could be foregone in favour of using descriptive statistics such as mean differences, standard deviation differences, and confidence intervals (Plonsky, 2015).

6. Conclusion

Jones and Blume (2022) used Bayesian analyses because as a methodology, Bayesian analysis lends itself to exploratory work more than frequentist work does, which more often than not is aimed at NHST. While it is entirely possible to interpret p values in a nuanced way, the tendency is to use them as an arbitrary value to give binary yes/no answers. Bayes factors provide an alternative to this, although Norouzian et al. (2019) caution that their guidelines for Bayes Factors should not be interpreted in similar ways to the use of p values. In conclusion, the use of any tool for overly

simplified decision making needs to be avoided and peer reviewers need to be vigilant and advise researchers on good research practices in data analysis, as they would do for other aspects of research methodology.

Notes

The original data and analysis script for R is available at OSF <https://osf.io/xb9ce/>. This means that the script may be built upon and further, multifaceted analyses of the data are possible.

References

- Barch D. H., & Chechile R. A. (2023). *DFBA: Distribution-Free Bayesian Analysis*. R package version 0.1.0, <https://CRAN.R-project.org/package=DFBA>
- Faul, F., Erdfelder, E., Buchner, A., & Lang, A. G. (2009). Statistical power analyses using G Power 3.1: Tests for correlation and regression analyses. *Behavior Research Methods*, 41(4), 1149–1160. <https://doi.org/10.3758/BRM.41.4.1149>
- Jones, M., & Blume, C. (2022). Accent difference makes no difference to phoneme acquisition. *TESL-EJ*, 26(3), 1–22. <https://doi.org/10.55593/ej.26103a3>
- Larson-Hall, J., & Herrington, R. (2010). Improving data analysis in second language acquisition by utilizing modern developments in applied statistics. *Applied Linguistics*, 31(3), 368–390. <https://doi.org/10.1093/applin/amp038>
- Mackey, B., & Ross, S. J. (2015). Bayesian informative hypothesis testing. In L. Plonsky (Ed.), *Advancing quantitative methods in second language research* (pp. 329–345). Routledge.
- McKinley, J., & Rose, H. (Eds.). (2017). *Doing research in applied linguistics: Realities, dilemmas, and solutions*. Routledge.
- Nishi, K., & Kewley-Port, D. (2007). Training Japanese listeners to perceive American English vowels: Influence of training sets. *Journal of Speech, Language, and Hearing Research*, 50(6), 1496–1509. [https://doi.org/10.1044/1092-4388\(2007/103\)](https://doi.org/10.1044/1092-4388(2007/103))
- Nishi, K., & Kewley-Port, D. (2008). Non-native speech perception training using vowel subsets: Effects of vowels in sets and order of training. *Journal of Speech, Language, and Hearing Research: JSLHR*, 51(6), 1480–1493. [https://doi.org/10.1044/1092-4388\(2008/07-0109\)](https://doi.org/10.1044/1092-4388(2008/07-0109))
- Norouzian, R., de Miranda, M., & Plonsky, L. (2019). A Bayesian approach to measuring evidence in L2 research: An empirical investigation. *The Modern Language Journal*, 103(1), 248–261. <https://doi.org/10.1111/modl.12543>
- Plonsky, L. (2015). Statistical power, p values, descriptive statistics, and effect sizes: A ‘back-to-basics’ approach to advancing quantitative methods in L2 research. In L. Plonsky (Ed.), *Advancing quantitative methods in second language research* (pp. 23–45). Routledge.
- Plonsky, L., & Oswald, F. L. (2014). How big is “big”? Interpreting effect sizes in L2 research. *Language Learning*, 64(4), 878–912. <https://doi.org/10.1111/lang.12079>
- Rubdy, R. (2015). Unequal Englishes: The native speaker, and decolonization in TESOL. In R. Tupas (Ed.), *Unequal Englishes: The politics of Englishes today* (pp. 42–58). Palgrave Macmillan.
- Schad, D. J., Nicenboim, B., & Vasissth, S. (2023). *Data aggregation can lead to biased inferences in Bayesian linear mixed models and Bayesian ANOVA: A simulation study* (arXiv:2203.02361). arXiv. <https://doi.org/10.48550/arXiv.2203.02361>
- Trafimow, D. (2024). *Methodological issues in psychology: Concept, method, and measurement*. Routledge.
- Trafimow, D., ... Marmolejo-Ramos, F. (2018). Manipulating the alpha level cannot cure significance testing. *Frontiers in*

Psychology, 9. <https://doi.org/10.3389/fpsyg.2018.00699>
Zondervan-Zwijnenburg, M., Peeters, M., Depaoli, S., & Schoot, R.
V. de. (2017). Where do priors come from? Applying guidelines
to construct informative priors in small sample research. *Research
in Human Development*, 14(4), 305–320.
<https://doi.org/10.1080/15427609.2017.1370966>

日本語の主題卓越型構造からの転移に関する一考察 —日本人初級・中級英語学習者の文産出データの比較を通して—

橋尾晋平*

*名古屋外国語大学
smp.hashio@gmail.com

要旨

外国語学習において、学習者の母語が目標言語での文産出に転移してしまうことがあるが、日本人英語学習者の文産出では、日本語の主題卓越型構造と英語の主語卓越型構造の混同がしばしば見られる。日本語の主題をもつ文は、「主題は文頭に置かれる」・「二重主語構造をとる」・「述語が代用化される」などの特徴があり、先行研究では、日本人初級英語学習者の文産出に対して、これらの特徴が負の影響を与えているとされている。本稿では、新たに日本人中級英語学習者による文産出のデータを収集し、分析を行ったところ、主題と主語が一致しない文については、英語で表現することが容易となり、転移は減少するということが判明した。一方、述語代用が生じた文では、日本人初級英語学習者と同様に、正解率が低いことが導かれたため、習熟度が上がっても、述語の代用化は、誤りを誘発する因子となりえると結論づけた。

キーワード：日本人初級・中級英語学習者、学習者言語、転移、文産出、主題、主語

1. はじめに

文部科学省は2023年6月に「第4期教育振興基本計画」を公開した。その中の「グローバル社会における人材育成」において、高等学校卒業の段階での英語力の目標に関して、ヨーロッパ言語共通参照枠（Common European Framework of Reference for Languages: CEFR）のA2レベル相当に到達する中高生の割合を5年後までに5割以上にすることを目指すと設定した。

しかしながら、「平成29年度英語力調査結果」（文部科学省, 2018）において、高等学校卒業段階の英語力を技能毎に調査したが、CEFR A2 レベル相当以上を達成した高校生の割合は、「聞くこと（33.6%）」・「読むこと（33.5%）」・「話すこと（12.9%）」・「書くこと（19.7%）」となっており、特に、「話すこと」と「書くこと」に関しては、それ以前の調査と比較しても横ばいであり、全体的に低い状況が続いている。

橋尾（2019）では、中学校・高等学校の日本語母語話者教員32名に対してアンケートを実施し、「話すこと」・「書くこと」について、中高生が具体的にどのようなことに躓いているのかを自由記述で回答してもらい、収集した回答をコード化したところ、語順・文構造が十分に身につけていない中高生が多いということを約3割の教員が感じていると判明した。

実際、橋尾（2023）などで指摘したように、日本人初級英語学習者の発話は、非文法的・非機能的なものが数多く観察される。例えば、筆者の過去の勤務校では、インターネットの利用の是非について自分の意見を述べる活動を実施した際、「インターネットは多くの情報が得られる」という内容を英語で表現しようとする

“*The Internet can get a lot of information.”と表現する学生が多く存在した。橋尾（2023）によると、学生が「インターネットは（では）」から始まる日本語文を発想したと思われる、このことは、日本語の主題卓越型構造からの転移（transfer）が関連していると推察される。

本稿では、Hashio and Yamauchi（2021）で用いた調査票と収集したデータを用いて、日本人初級英語学習者と日本人中級英語学習者の文産出について比較を行い、日本語の主題卓越型構造がどのように学習者言語に転移するか、また、学習者の習熟度に応じて、文産出で見られる転移がどのように変化するかを論じる。

2. 理論的背景・先行研究

2.1 英語学習における母語の影響

Jarvis and Pavlenko（2008）によると、転移とは、母語と目標言語が相互に正の影響あるいは負の影響を与えることである。¹ また、Odlin（2003）によると、転移は、母語・目標言語の特定の項目に関して、客観的な言語差の有無にかかわらず、学習者自身が似ていると判断された場合に生じる。学習者に影響を与える項目は、統語構造、語彙、音韻などの言語的特性から文体、言語形式・機能のような語用論的知識まで多岐に亘る（Ellis, 1985）。後述するが、日本語の主題卓越型構造は、形式的知識であるだけでなく、語用論的知識でもある。

Kellerman（1979）によると、習熟度の低い外国語の学習者は、目標言語の知識不足によって、母語を外国語とリンクさせながら学ぼうとするため、母語に依存する傾向が強く、母語の知識が外国語学習に影響を与えやすいと考えられている。

一方、習熟度が上がると、目標言語の知識が増えていくことで、母語と目標言語の違いが意識できるようになり、転移が減少していくものの、母語の知識を何でも目標言語に置き換えることができないと判断してしまうことで、かえって誤りが増えてしまう場合がある (Kellerman, 1979)。

また、Takahashi (2000) によると、語用論的知識の転移は、習熟度の高い学習者に多く見られ、その理由は、複雑な表現であっても母語から容易に置き換えられるため、過剰一般化は習熟度が上がっても消えにくいからであると考えられている。

2.2 日本語の主題をもつ文の特徴

日本語文のプロトタイプは、「X は」から始まる文であり、文頭に置かれる助詞「は」にマークされた名詞句 X は、主題 (topic) と呼ばれている。Masuoka (2017) によると、主題とは、文において解説 (comment) が与えられるときの対象 X のことであり、(1) では、「インターネット」が主題にあたり、「インターネット」の「多くの情報が得られる」という特徴や属性が解説される文となっている。

(1) インターネットは 多くの情報が得られる。

Masuoka (2017) によると、日本語は「主題＋解説」を基盤とする主題卓越型構造の言語に分類され、英語は文法上の主格の働きが強い「主語＋述語」を基盤とする主語卓越型構造の言語に分類される。Hashio and Yamauchi (2021) は、この日英語の違いが、日本人初級英語学習者の文産出に影響を与えると指摘した。

日本語の主題をもつ文の特徴として、以下の (i) ~ (v) が挙げられる (Hashio & Yamauchi, 2021; Schachter & Rutherford, 1979)。

- (i) 主題は文頭に置かれる
- (ii) 主題は提題助詞「は」にマークされる
- (iii) 空主語が認められる
- (iv) 二重主語が認められる
- (v) 述語の代用が認められる

まず、(i) については、主語以外の名詞句 X が前置されるということであり、(2a) の「その本」を主題として表したい場合、(2b) のように「その本」を前置する。(2b) では、「その本」は主題であるが、主語ではない。

- (2) a. 父が その本 を買った。
- b. その本は 父が買った。

次に、(ii) については、これまで見てきたように、主題となった名詞句には提題助詞「は」のマークがなされるということである。² 提題助詞「は」は、主語の名詞句だけでなく、さまざまな名詞句をマークするため、

「は」はさまざまな助詞の機能を代行している (三上, 1960)。実際、(2a) の「その本」は目的格を示す「を」にマークされているが、(2b) では、「その本」が主題になることで、「を」が「は」に交替するが、これは提題助詞「は」が目的格の「を」の機能を兼務している。

また、(iii) に関して、日本語の特徴としてよく知られているが、「夏休み」の話題について述べる際に、日本語の場合、(3a) と述べることもできるし、(3b) のように主語を省略して述べることもできる。

- (3) a. 私は夏休みに仙台へ行きました。
- b. 夏休みは仙台へ行きました。

(iv) に関して、二重主語構造をもつ文 (以下、「二重主語文」と呼ぶ) は、日本語における代表的な構文である。二重主語文は、「X は Y が Z」という構造をなし、X は主題、「Y が Z」の部分が述部 (解説) にあたる (野田, 1994)。³ (4) は、「うさぎ」は主題であり、「耳が長い」が述部であるが、「耳が長い」の部分においても、「耳」と「長い」が主述関係にあることにある。

- (4) うさぎは 耳が長い。

最後に、(v) についても、日本語の主題卓越型構造が強く反映されている。日本語の主題をもつ文の中では、文脈に応じて、述語が「だ」に代用されるものがある (奥津, 1978)。以下を飲食店での注文の場面であると設定した際、(5a) の「欲しい」という述語は、(5b) において「だ」に置き換わっている。

- (5) a. 私はコーヒーが欲しい。
- b. 私はコーヒーだ。

以上より、日本語の主題卓越型構造の文は、(i) ~ (v) の特徴を持っており、英語にはなかなか見られないと思われるが、日本人英語学習者の文産出にとって、しばしば支障となり得る。次節では、日英語の違いが学習者の文産出にどのように影響するかを概観していく。

2.3 日本人初級英語学習者の主題卓越型構造の転移

日本語の主題卓越型特徴が主語卓越型言語の英語の習得に影響を与える可能性について、最初に指摘を行ったのは Schachter and Rutherford (1979) であり、その後、Sasaki (1990) や梅原・富永 (2014) などは、日本人初級英語学習者が日本語の主題と英語における主語を同一視していると指摘した。すなわち、多くの日本人初級英語学習者が日本語の主題と英語の主語が似ているものであると認識したため、転移が生じたと考えられる。

(6a) を英語で表現する際、(6b) のように表現されるが、多くの日本人初級英語学習者は、主題の「インターネット」を主語と捉えて、文を作ろうとするため、(6c) のような誤った文を産出してしまうことになる。

- (6) a. インターネットは多くの情報が得られる。
b. We can get a lot of information on the Internet.
c. *The Internet can get a lot of information.

二重主語文や述語が代用された文について、小林 (2008) や Aoki (2006) などは、多くの日本人初級英語学習者が英語のコピュラ文と同一視し、転移が生じると指摘している。

例えば、二重主語文について、(7a) では、「うさぎ」が「耳」を所有しており、「耳」が「長い」という特徴を有しているので、(7b) や (7c) のように表現する。しかし、日本人初級英語学習者は、「うさぎ」・「耳」・「長い」の間の関係を理解できず、(7d) のような文を産出してしまうことが多い。なお、橋尾 (2022) は、二重主語文にはさまざまな種類の構文があり、二重主語文における X・Y・Z の間にある主述関係や修飾・被修飾の関係を把握できるかに応じて、英語への訳しやすさが変わると指摘している。

- (7) a. うさぎは耳が長い。
b. Rabbits' ears are long.
c. Rabbits have long ears.
d. *Rabbits are long ears.

また、述語代用が生じた文について、日本人初級英語学習者の多くは、文中に表示されていない述語を補完できず、(8a) を (8b) のような英文として産出するのは困難であり、(8c) のような誤った英文を産出する。

- (8) a. (カフェの注文で) 私はコーヒーだ。
b. I would like a cup of coffee.
c. *I am a coffee.

Hashio and Yamauchi (2021) は、先述の (i) ~ (v) について、どの特徴が日本人初級英語学習者の文産出に影響を与えるのかを明らかにするために、日本人初級英語学習者の大学生に対して、和文英訳テストを実施し、その回答データの分析を行った。その結果、表 1 が示すように、(i)・(iv)・(v) の特徴が日本人初級英語学習者の文産出に強く影響することが判明した。

表 1
主題卓越型構造の特徴と日本人初級英語学習者への影響

特徴	日本人初級英語学習者への影響
(i) 主語以外の名詞句の前置	非常に強い影響を与える
(ii) 「は」のマーク	あまり大きいとは言えない
(iii) 空主語	やや影響を与える
(iv) 二重主語文	強い影響を与える
(v) 述語の代用	非常に強い影響を与える

一方で、提題助詞「は」のマークや空主語は、転移の

誤りが増える要因とはなり得ないことが示された。筆者の知る限り、さまざまな場面で、日本人初級英語学習者は、「は」と「が」の識別ができていないと指摘されてきているが、実際のところ、(9a)・(9b) のどちらも (9c) ではなく、(9d) のように表現してしまうことが多いため、日本人初級英語学習者は「は」にマークされたものを主語と判断するのではなく、(i) の特徴と関連して、文頭の名詞句である主題を主語と同一視する傾向が強いと思われる (Hashio & Yamauchi, 2021)。

- (9) a. インターネットは多くの情報が得られる。
b. インターネットで多くの情報が得られる。
c. We can get a lot of information on the Internet.
d. *The Internet can get a lot of information.

以上より、日本語の主題をもつ文は、英語の文構造と大きく異なるが、多くの日本人初級英語学習者は、その違いを気づかないまま、日本語の主題卓越型構造をそのまま英語の主語卓越型構造に置き換えて表現できると考えてしまい、英語の主述関係を適切に把握することができない場合が多いと考えられている。

2.4 日本人中級英語学習者に対する母語の影響

ここまでは、日本人初級英語学習者に対する母語の影響、日本語の主題卓越型構造の転移について論じてきたが、筆者の知る限りでは、日本人中級英語学習者への転移についての研究はあまり見られない。

先述したとおり、習熟度が上がることで、母語の知識を目標言語に置き換えることができないと保守的に判断してしまう場面が増えてしまい、転移が減少していくが、誤りが増えてしまう場合がある (Kellerman, 1979)。

また、語用論的知識の転移が習熟度の高い学習者に多く見られる (Takahashi, 2000) が、日本語の主題卓越型構造は、文字通り構造という形式的知識でもある一方で、主題が提示され、新情報が含まれる解説が付け加えられる表現手段は言語機能であり、語用論的知識にあたるため、Takahashi (2000) の指摘を踏まえると、中級レベルの学習者であっても、主題をもつ文は、英語による文産出に影響が出ると推測される。

Kuribara (2004) によると、日本人英語学習者に文法性判断テストを課したが、動詞句の前に名詞が置かれている文に関しては、どのような文も容認する傾向が強く、2 種類の名詞句が動詞の前に置かれている場合も容認する学習者が一定数いると指摘し、この傾向は少なくとも中級レベルの学習者の間でもよく見られると報告している。したがって、(i) の特徴が日本人中級英語学習者の文産出にも転移する可能性がある。また、橋尾 (2021) は、日本人初級英語学習者に対して、ここまで紹介してきた構文からの転移を克服するための和文英訳の訓練を行ったが、(v) の述語の代用化に関して、「A は B だ」文の“A is B”文への過剰一般化は改善されなかったと報告している。

以上より、日本語の主題卓越型構造からの転移は、習熟度が上がっても、必ずしも克服されるわけではないと考えられ、次節では、学習者の習熟度が初級から中級に上がるにつれて、(i)～(v)の特徴が与える学習者の文産出への影響に違いがあるのかを明らかにし、日本語の転移の実態をより詳細に整理する。

3. 日本人中級英語学習者の文産出に関する調査

3.1 目的

Hashio and Yamauchi (2021) の調査を踏襲し、2.2・2.3で整理した(i)～(v)の特徴が日本人初級・中級学習者の文産出に影響を与えるかどうかを分析する。特に、初級から中級にかけて習熟度が上がることで、日本語からの転移が克服される項目とされない項目があるか、を明らかにしていく。

3.2 調査協力者

本稿の調査は、CEFR B1 レベルに該当する日本人中級英語学習者の大学生 90 名を対象とし、協力者全員が実用英語技能検定 2 級を習得している。また、比較のため、Hashio and Yamauchi (2021) で収集した日本人初級英語学習者の大学生 70 名分のデータを用いる。

3.3 調査票・統計的分析

Hashio and Yamauchi (2021) で使用したものと同一の調査票を用いて、協力者の学生には、30 問の和文英訳問題 (Q1～30) からなるテストに回答してもらう。調査票については、付録 A を参照されたい。

すべての日本語文が主題をもつ文であり、主語・主題・述語に関して、以下の (a)～(c) の特徴を含んでいる。⁴表 2 は、Q1～30 の日本語文がそれぞれどのような特徴を持っているかをまとめている。

- (a) 主語：「空主語」・「二重主語 (二重)」・「その他 (表示)」
- (b) 主題：「主語と一致している (主語)」・「主語と一致していない (非主語)」
- (c) 述語：「表示されている (表示)」・「代用化されている (代用化)」

調査実施後、回収したテストについて採点を行うが、構文的に正しい英文が産出できるかをチェックし、正誤をまとめたデータを用いて、日本人初級英語学習者 70 名分の正誤データとの比較を行っていく。まず、Hashio and Yamauchi (2021) で収集したデータを用いて、単純集計とフィッシャーの正確検定による比較を行う。なお、フィッシャーの正確検定においては有意水準を 5% と設定する。次に、30 題の問題の正解者数を目的変数として、主語・主題・述語の特徴を説明変数とした重回帰分析を行っていく。

表 2

各日本語文の主語・主題・述語の特徴

問題	主語	主題	述語	問題	主語	主題	述語
Q1	表示	主語	表示	Q16	二重	主語	表示
Q2	表示	非主語	表示	Q17	二重	非主語	表示
Q3	表示	非主語	表示	Q18	空主語	非主語	表示
Q4	表示	非主語	表示	Q19	二重	主語	表示
Q5	表示	非主語	表示	Q20	空主語	非主語	表示
Q6	表示	非主語	表示	Q21	空主語	非主語	代用化
Q7	表示	非主語	表示	Q22	空主語	非主語	表示
Q8	表示	非主語	表示	Q23	空主語	非主語	表示
Q9	表示	非主語	表示	Q24	表示	主語	表示
Q10	表示	非主語	表示	Q25	空主語	非主語	表示
Q11	表示	主語	表示	Q26	空主語	非主語	代用化
Q12	二重	非主語	表示	Q27	空主語	非主語	表示
Q13	空主語	非主語	表示	Q28	空主語	主語	表示
Q14	空主語	主語	代用化	Q29	空主語	非主語	表示
Q15	空主語	主語	表示	Q30	空主語	非主語	表示

4. 結果・考察

4.1 フィッシャーの正確検定の結果

本節では、まず、日本語の主題をもつ文の (i)～(v) の特徴に関して、日本人中級英語学習者の文産出への影響についての分析結果について述べる。

まず、Hashio and Yamauchi (2021) で収集した日本人初級英語学習者の大学生 70 名分のデータと本稿で収集した日本人中級英語学習者 90 名分のデータを踏まえて、図 1・図 2 では、全 30 問毎の正解率を習熟度別にそれぞれ示している。また、問題毎に日本人初級・中級英語学習者の間で、正解率の差が見られるかどうかについて、フィッシャーの正確検定に行い、その検定における p 値を表 3 に示し、 p 値が 0.05 未満のものについては、太字で示している。

図 1

問題毎の正解率 (Q1～15)

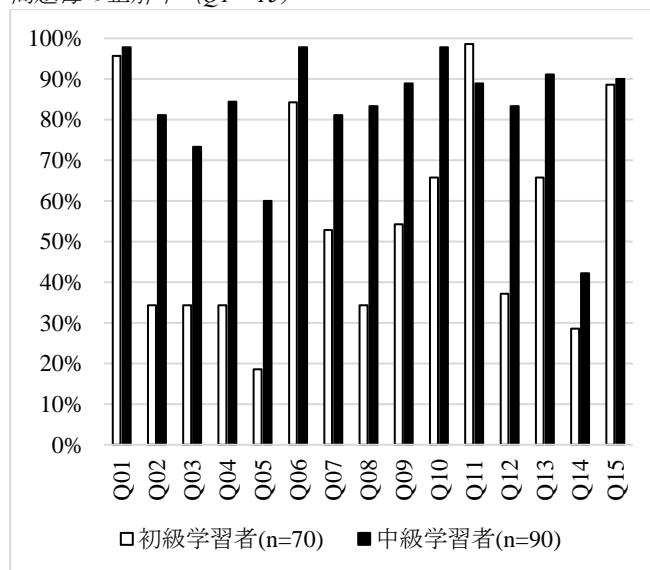


図 2
問題毎の正解率 (Q16~30)

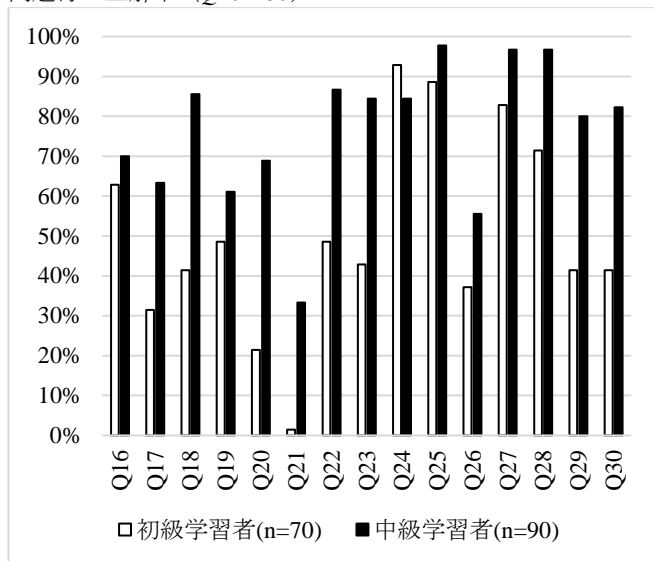


表 3
フィッシャーの正確検定における p 値

問題	p 値	問題	p 値	問題	p 値
Q1	0.65	Q11	0.02	Q21	0.00
Q2	0.00	Q12	0.00	Q22	0.00
Q3	0.00	Q13	0.00	Q23	0.00
Q4	0.00	Q14	0.10	Q24	0.14
Q5	0.00	Q15	0.80	Q25	0.02
Q6	0.00	Q16	0.40	Q26	0.03
Q7	0.00	Q17	0.00	Q27	0.00
Q8	0.00	Q18	0.00	Q28	0.00
Q9	0.00	Q19	0.15	Q29	0.00
Q10	0.00	Q20	0.00	Q30	0.00

日本人初級・中級英語学習者の間で正答者数の割合に差が見られなかった問題 (表 3 で太字になっていない問題) は、Q1・14~16・19・24 であり、個別に見ていくと、初級・中級レベル両方の学生の正解者数も 8 割以上だった問題が Q1・15・24 に該当し、どちらのレベルの学生も正解者数が 8 割未満だった問題が Q14・16・19 にあたり、Q14 の日本語文は述語の代用がなされており、Q16・19 は二重主語文を用いた問題である。

一方、日本人初級・中級英語学習者の間で正答者数の割合に差が生じた 24 問 (表 3 で太字になっている問題) の中で、Q2~10・13・18・20・22・23・25・27・29・30 の 18 問は、主語と主題が一致していない日本語文であり、(i) の特徴が反映されている。よって、習熟度が上がったことで、主語と主題の区別をつけることができるようになるかと推察される。

また、Q12・17 は二重主語文、Q21・26 は述語が代用化された文を用いているが、中級レベルの正答率の方が有意に高いことが示されているが、Q12 に関しては、中級レベルの学生の正答率が 8 割以上のため、習熟度が上がることで転移が減少できたと思われる。ただし、Q17・21・26 は、中級レベルの学生であっても正答率が

8 割未満であり、特に Q21 は 3 割程度であるため、転移を克服できたわけではないと導ける。

Q28 については、主題と主語が一致している文ではあるが、日本人初級英語学習者の正答率が例外的に低く、また、Q11 は初級レベルの学生の正答率が中級レベルの学生よりも有意に高いことが判明したが、中級レベルの学生の正答率も 8 割以上のため、こちらも本稿では例外扱いとする。

4.2 重回帰分析の結果

表 4・表 5・表 6 は、主語・主題・述語の特徴毎の要約統計量を示している。表 4 では、主語の特徴が「表示」「空主語」「二重主語」の問題数と平均正解者数・標準偏差をまとめており、表 5 は、主題が「主語」と一致している問題と主語と一致していない「非主語」の問題の数と平均正解者数・標準偏差を表している。表 6 は述語が「表示」されている問題と「代用化」されている問題の平均正解者数・標準偏差をまとめている。また、各問題の正解者数を目的変数とし、主語・主題・述語の特徴を説明変数とした重回帰分析の結果を表 54 のとおりに示す。

表 4
主語の特徴毎の要約統計量

	問題数	平均値	標準偏差
表示	11	76.45	10.32
空主語	15	70.53	17.84
二重主語	4	62.50	9.00

表 5
主題の特徴毎の要約統計量

	問題数	平均値	標準偏差
主語	9	71.22	16.45
非主語	21	71.81	14.50

表 6
述語の特徴毎の要約統計量

	問題数	平均値	標準偏差
表示	27	75.22	10.17
代用化	3	39.33	10.07

表 7
重回帰分析の結果

		係数	標準誤差	t 値	p 値
主語	(Intercept)	40.68	5.83	6.98	0.00
	二重主語	-16.34	5.36	-3.05	0.01
	表示	-1.93	3.82	-0.50	0.62
	空主語	0			
主題	非主語	-2.02	3.70	-0.55	0.59
	主語	0			
述語	表示	39.17	5.91	6.63	0.00
	代用化	0			

重回帰分析の結果、次の 3 点が明らかになった。まず

1 点目に、日本語の主語と英語の主題が一致しなかったとしても、日本人中級英語学習者の文産出において転移がほとんど生じなかったと推察される。2 点目に、主語が文中で表示されている文や空主語の文では、中級レベルの学生の誤りがほとんど見られなかったが、二重主語文を正しい英語で表現する学生数は減少したと示された。3 点目として、基本語順の文と同様に、述語が表示されている場合と代用化された場合では、正解者数に大きな差が生じていることが導かれ、述語が代用化された文を英語で表現することは、中級レベルの学生にとっても困難であると判明した。

4.3 結果のまとめ

単純集計の結果およびフィッシャーの正確検定と重回帰分析の結果を踏まえて、日本人中級英語学習者の文産出への日本語の主題卓越型構造の (i) ～ (v) による影響について、表 8 のようにまとめた。以降、学生の回答例を挙げながら、中級レベルの学生の文産出に関して、考察を深めていく。

表 8

主題卓越型構造の特徴と日本人中級英語学習者への影響

特徴	日本人中級英語学習者への影響
(i) 主語以外の名詞句の前置	あまり影響はない
(iii) 空主語	ほとんど影響はない
(iv) 二重主語文	やや影響を与える
(v) 述語の代用	非常に強い影響を与える

まず、(i) の特徴は、Hashio and Yamauchi (2021) と比較すると、中級レベルの学生となると、転移が生じなくなっていくことが示された。例えば、(10a) について、初級レベルの学生は (10b) のように英訳する傾向が強かった (Hashio & Yamauchi, 2021) が、中級レベルになると、(10c) のように英語で表現できる学生が多い。

- (10) a. そのデパートでは新鮮な魚を買うことができる。
 b. *The department store can buy fresh fish at the department store.
 c. We can buy fresh fish at the department store.

Kuribara (2004) は、文法性判断テストを通して、日本人中級英語学習者であっても、主題と主語を同一視してしまう傾向があると主張しており、本稿で行った調査においても、一部の問題で英訳の正解率が 8 割未満だったものもあるが、少なくとも中級レベルとは、主題と主語の区別がついてくるようになってくる段階であると推察できる。

次に、(iv) の特徴に関して、二重主語文の英訳の問題となっているのは、Q12・16・17・19 の 4 問であり、Q12 のみ正解率が 8 割以上となっているが、他 3 問の正解率は 6～7 割となっており、重回帰分析の結果を鑑

みても、相対的に英語で表しづらい文であると考えられる。例えば、初級・中級両方のレベルの学生の多くは、(11a) を (11b) でなく、(11c) のように表してしまう。

- (11) a. 辞書は新しいのがいい。
 b. A new dictionary is good.
 c. *Dictionaries are new good.

ただし、フィッシャーの正確検定の結果を踏まえると、Q12・17 は、初級レベルの学生よりも、中級レベルの学生の方が正確に英語で表現することができると統計的に明らかになっており、(iv) の特徴による転移は、中級レベルになると、少しずつ克服できるようになる可能性が示された。橋尾 (2022) において、二重主語文の構文のタイプに応じて、学習者の英訳のしやすさが変わると示唆されているため、日本人中級英語学習者にとって、どのような構文が英語で表現することが困難であるかについては、今後さらに検討していく必要があると思われる。

最後に、(v) の特徴に関して、述語が代用化された日本語文の英訳の問題となっているのは、Q14・21・26 の 3 問であり、中級レベルの学生の正答率は、Q14 が 4 割程度、Q21 が 3 割程度、Q26 が 5 割程度であった。フィッシャーの正確検定の結果が示すように、Q21・26 は、中級レベルの学生の方が正確に英語で表現できると示せてはいるものの、重回帰分析の結果からも明らかになったように、初級レベルの学生だけでなく、中級レベルの学生も英語で表現することが非常に難しいことが示された。(12a) は Q21 で出題された文であるが、本稿の調査結果のとおり、(12b) のように英語で書くことができる学生が 3 割程度であり、(12c) のような誤った英語を産出する学生が非常に多いことが判明した。

- (12) a. その授業は毎週火曜日です。
 b. We take the class every Tuesday.
 c. *The class is every Tuesday.

したがって、重回帰分析の結果が示したように、二重主語文以上に述語代用の生じた文の英訳は困難であり、文中に表示されていない述語を補完し、英語で表現することは極めて困難であると思われる。

以上より、習熟度が上がると、主題と主語の混同が生じなくなるが、中級レベルになっても、述語代用の文・二重主語文を英語で表現することが難しく、特に述語が日本語文中で表示されていない文の場合、それを補完して英語で表現することは難しいと示せた。2.4 において、Takahashi (2000) が指摘したように、中級レベルになり、英語の知識が増えてきたことで、複雑な表現であっても母語から容易に置き換えられるため、過剰一般化は習熟度が上がっても消えにくい。よって、習熟度が中級レベルになったとしても、「A は B だ」文の“A is B.”文への過剰一般化が消えにくいと結論づける。

5. おわりに

日本語の主題卓越型構造には、主に 5 つの特徴があるが、日本人初級英語学習者は、日本語の「主題＋解説」構造の文と英語の「主語＋述語」構造の文を無意識に同一視する傾向があることが先行研究から示されてきた。初級の段階では、目標言語である英語の知識が不足しているため、日本語の知識の借用が多く、転移が生じやすいと考えられ、日英語の違いについての「気づき」を得ていない状態にあると思われる。

中級の段階になると、英語の知識が増えてくるため、日英語の違いを意識できつつある状態になり、日本語の主題と英語の主語を区別できるようになってくる。ただし、先述のとおり、英語の知識が増えてきたことで、母語である日本語から英語への置き換えができるようになったが、「A は B だ」文の“A is B.”文への過剰一般化が消えにくく、特に述語代用が生じる文の英訳が極めて難しいことが判明した。

ただし、本稿では、二重主語文・述語が代用化された文に関する問題数が合わせて 7 題と少なめであった。このことは、日本人初級英語学習者との比較を行う目的で、Hashio and Yamauchi (2021) と同様の調査票を用いたためであるが、Hashio and Yamauchi (2021) は、主語と一致しない主題をマークする提題助詞「は」が兼務している機能にも着目していたため、二重主語文・述語が代用化された文の問題数をこれ以上増やすことができなかった。この点については、今後の課題としたい。

加えて、本稿の調査では、和文英訳テストに回答してもらい、その回答を分析データとし、学習者言語の分析を行ってきたが、Hashio and Yamauchi (2021) などでは、どのような思考プロセスで英文を産出したのかについて、数名の日本人初級英語学習者の大学生に聴き取り調査を行ってきたが、今後は日本人中級英語学習者に対しても、インフォーマントな聴き取り調査の実施も行い、文産出における学習者の思考過程をより精緻に解明していきたいと考える。

最後に、本稿で得られた知見を踏まえて、日本人初級・中級英語学習者が日本語の主題卓越型構造からの転移を克服するための方策について論じる。現行の文部科学省の『学習指導要領』などでは、「日本語と英語の違いに注意する」という趣旨の文言が盛り込まれているが、学習者たちの母語である日本語についての「気づき」を促すことが今後の英語教育にとって重要になってくる。橋尾 (2021) では、主題をもつ日本語文を英語で表現する訓練を行う際に、田地野 (2008) の開発した意味順を用いて、日英語の違いについて視覚的に気づきを与える指導法を提案・実践したが、本稿で取り扱った構文を重点的に練習する機会を設けることが大切であると主張する。

また、現在、習熟度にかかわらず、より実践的なコミュニケーション活動を取り入れた英語教育の需要が高まっており、近年、コミュニケーション活動を通して、

文法能力を養成するフォーカス・オン・フォーム (Focus on Form) の指導法が推奨されており、高島 (2011) では、日本語の言語機能の転移を克服する手段として有効であると主張している。橋尾 (2023) では、フォーカス・オン・フォームは有効であると主張したが、プレゼンテーションやディベートなどのスピーキング活動は、何らかのテーマを設定して、それについて述べる活動であり、主題を述べて、それに関する新しい情報を加えていく日本語の表現形式と通ずる部分があると考えられ、主題をもつ日本語文の英訳の訓練との親和性が高いと期待される。よって、プレゼンテーションやディベートなどを通して、日本語の影響の克服を図りつつ、コミュニケーション能力の養成することも今後目指していきたい。

注

¹ Luk and Shirai (2009) によると、正の転移は、統語構造や語彙、音韻が類似しているとより生じやすく、学習速度も速くなる傾向があるが、本稿における転移は、負の転移のみを扱うこととする。

² Masuoka (2017) によると、提題助詞「は」にマークされた主題を無標主題、「は」にマークされていない主題を有標主題と定めた。

³ 野田 (1994) を参照すると、二重主語文の中には、格成文・副詞的成分が主題の文が含まれ、(2b) などがその例に該当するが、このタイプの文は、(i) ~ (iii) の特徴が反映されている文であるため、本稿では、二重主語文として扱わないこととする。

⁴ Hashio and Yamauchi (2021) では、無標主題と有標主題をそれぞれ初級レベルの学生に英訳させるテスト調査も別途を行い、「は」のマークは日本人初級英語学習者への転移に関係しないと結論をづけているため、本稿の調査では、調査協力者への負荷などを考慮し、30 題すべての日本語文が無標主題の文とした。

謝辞

第 6 回 JAAL in JACET 学術交流学会にて有益なコメントをくださった先生方および拙稿に対して貴重なコメントをくださった匿名の査読者の先生方に、心より謝意を表する。

参考文献

- Aoki, C. (2006). Interference caused by transfer of characteristics of topic-prominent languages. 『北海学園大学学園論集』, 130, 21–35.
- Ellis, R. (1985). *Understanding second language acquisition*. Oxford University Press.
- 橋尾 晋平. (2019). 「日本人英語学習者の主題卓越構造からの転移の克服に向けた英語教育に関する一考察」. 『比較文化研究』, 137, 51–63.
- 橋尾 晋平. (2021). 「日本語の主題卓越型構造からの転移の克服に向けた意味重視の文法指導に関する一考察」. 『JAAL in JACET Proceedings』, 3, 93–100.
- 橋尾 晋平. (2022). 「日本人初級英語学習者の文産出に見られる日本語の主題卓越型構造からの転移に関する一考察—日本語の二重主語文に焦点をあてて—」. 『JAAL in JACET Proceedings』, 4, 86–93.

- 橋尾晋平. (2023). 「初年次英語教育における「シンプル・ディベート」の実践報告—日本人初級英語学習者のコミュニケーション能力養成の観点から」. 『JAAL in JACET Proceedings』, 5, 67–74.
- Hashio, S., & Yamuchi, Y. (2021). A note on the influence of topic prominence in Japanese on Japanese beginner-level EFL learners' interlanguage: An empirical study. In J. R. Jaskow & A. Wolanin (Eds.), *Facing diversity in child foreign language education* (pp. 291–311). Springer.
- Jarvis, S., & Pavlenko, A. (2008). *Crosslinguistic influences in language and cognition*. Routledge.
- Kellerman, E. (1979). The problem with difficulty. *Interlanguage Studies Bulletin*, 4, 27–48.
- 小林亜希子. (2008). 「国文法を利用した英文法教育の試み」. 『島大言語文化—島根大学法文学部紀要言語文化学科編』, 25, 41–75.
- Kuribara, C. (2004). The identity of pre-verbal noun phrases in Japanese learners' English. 『外国語教育研究』, 7, 1–17.
- Luk, Z. P., & Shirai, Y. (2009). Is the acquisition order of grammatical morphemes impervious to L1 knowledge?: Evidence from the acquisition of plural *-s*, articles, and possessive *'s*. *Language Learning*, 59, 721–754. <https://doi.org/10.1111/j.1467-9922.2009.00524.x>
- Masuoka, T. (2017). Topic and subject. In M. Shibatani, S. Miyagawa & H. Noda (Eds.), *The Handbook of Japanese Syntax* (pp. 97–122). Mouton de Gruyter.
- 三上章. (1960). 『象は鼻が長い』. くろしお出版.
- 文部科学省. (2018). 「平成 29 年度英語力調査結果（高校 3 年生）の概要」.
- 文部科学省. (2023). 「第 4 期教育振興基本計画」.
- 野田尚史. (1994). 「日本語とスペイン語の主題化」. 『言語研究』, 105, 32–53.
- Odlin, T. (2003). Cross-linguistic influence. In C. J. Doughty & M. H. Long (Eds.), *Handbook of Second Language Acquisition* (pp. 436–486). Blackwell.
- 奥津敬一郎. (1978). 『「ボクハウナギダ」の文法—ダとノ』. くろしお出版.
- Sasaki, M. (1990). Topic prominence in Japanese EFL students' existential constructions. *Language Learning*, 40, 369–385. <https://doi.org/10.1111/j.1467-1770.1990.tb00667.x>
- Schachter, J., & Rutherford, W. (1979). Discourse function and language transfer. *Working papers on bilingualism*, 19, 3–12.
- 田地野彰. (2008). 「新しい学校文法の構築に向けて—英文作成における『意味順』指導の効果検証」. 小山俊輔・西堀わか子・田地野彰. (編). 『平成 20 年度英語の授業実践研究』 (pp. 8–21). 奈良女子大学国際交流センター.
- 高島英幸. (編). (2011). 『英文法導入のための「フォーカス・オン・フォーム」アプローチ』. 大修館書店.
- Takahashi, S. (2000). Reassessing L1 transfer in second language acquisition research: New perspectives. *Studies in Languages and Cultures*, 10, 51–75. <https://doi.org/10.15017/5467>
- 梅原大輔・富永英夫. (2014). 「日本人英語学習者は主語をどう捉えているか—量的・質的研究」. *JACET Kansai*, 16, 103–122.
- 「私のおばあちゃんは先月 9 調査票で用いられた日本語文 0 歳になった *grandmother 「おばあちゃん」
 - 「そのコンサートは Takahiro が主催した。」 *concert 「コンサート」 *host 「主催する」
 - 「清水寺は毎年多くの観光客が訪れる。」 *Kiyomizu-dera temple 「清水寺」 *tourist 「観光客」
 - 「そのプログラムは多くの学生が参加した。」 *program 「プログラム」 *attend 「参加する」
 - 「彼の雑貨屋さんはたくさんのお客さんが来る。」 *general store 「雑貨屋さん」 *customer 「お客さん」
 - 「今年の 2 月には私の姉が結婚する。」 *get married 「結婚する」 *February 「2 月」
 - 「映画館は隣の町にあります。」 *movie theater 「映画館」 *neighboring town 「隣の町」
 - 「その公民館は子供たちが英語を勉強することができます。」 *public hall 「公民館」
 - 「その地図ではその運転手がホテルにたどりつけなかった。」 *map 「地図」 *driver 「運転手」 *hotel 「ホテル」 *get to 「たどりつく」
 - 「昨年は 3 人の中国人が私たちの大学に入学した。」 *Chinese 「中国人」 *enter 「入学する」
 - 「私の母はその公民館でアルバイトをしている。」 *public hall 「公民館」 *work part-time 「アルバイトする」
 - 「その公民館は屋根が赤い。」 *roof 「屋根」
 - 「Ayumi には 9 年前大学で会った。」 *university 「大学」
 - 【Ayumi の仕事について】「彼女は旅行代理店だ。」 *travel agency 「旅行代理店」
 - 「彼女は出張の時、何でも予約してくれた。」 *on my business trip 「出張の時」 *everything 「何でも」 *book 「予約する」
 - 「うどんはその店が美味しい。」 *udon noodle 「うどん」 *shop 「店」
 - 「そのうどんはスープが美味しい。」 *soup 「スープ」
 - 【お店の HP によると、】「その店には車で行くことができない。」
 - 「辞書は新しいのがよい。」 *dictionary 「辞書」
 - 「辞書は英語の授業で必要です。」 *class 「授業」
 - (20)に関連して「その授業は毎週火曜日です。」 *Tuesday 「火曜日」
 - (現在辞書を書店で注文しており、)「来週にはそれが手に入る予定です。」
 - 「そのデパートでは新鮮な魚を買うことができます。」 *department store 「デパート」 *fresh 「新鮮な」
 - 「そのデパートでは毎年物産展を開いている。」 *food fair 「物産展」 *hold 「開く」
 - 「去年はその物産展に行きました。」
 - (予定について)「明日は大阪で同窓会だ。」 *reunion 「同窓会」
 - 「そのため、家は 4 時に出なければならない。」 *leave 「出る」
 - 「その店にはたくさんの雑貨がある。」 *grocery 「雑貨」
 - 「カードでは商品を買うことができません。」 *merchandise 「商品」 *credit card 「カード」
 - 「このお皿はそのお店で買った。」 *plate 「皿」

付録 A. 本稿で使用した調査票

次の 1～30 の「 」内の日本語文を英語に訳してください。なお、各日本語文に与えてある語彙は用いてもよいです。原則 1 問目から順番に解答してするようにしてください。また、辞書や文法書を用いずに、すべての問いに対して解答を記入するように心がけてください。

日本語母語英語学習者の時制誤用の変化 —使用（用法）基盤モデルと事態把握の観点から—

長 加奈子*

*福岡大学
kcho@fukuoka-u.ac.jp

要旨

英語と日本語は認知類型を異にしており、その認知類型の違いに基づく出来事のとらえ方（事態把握）の違いが、日本語を母語とする英語学習者が英語を使用する際に大きな影響を及ぼす。本研究は出来事のとらえ方の違いが表出している文法項目の1つである時制を取り上げ、認知言語学の知見に基づいた文法指導に加え、多読学習を通じた意味ある言語使用事例との接触、そして日本語から英語への出来事のとらえ方の転換を促す仕組み作りを通して、日本語を母語とする英語学習者の産出する誤用がどのように変化するかについて分析する。学習者のライティング課題の分析の結果、英語の出来事のとらえ方を示す時間軸を明示することで、学習者の時制の誤用が減少し、日本語から英語の出来事のとらえ方に移行する可能性があることが示唆された。本研究では、これまであまり注目されていない英語の出来事のとらえ方を習得させる教材の必要性を主張する。

キーワード：認知言語学、使用基盤モデル、時制、事態把握（出来事のとらえ方）、多読学習

1. はじめに

学習指導要領の改訂によって、実際のコミュニケーション活動の場面で活用することを通して文法などの知識を学習、そして評価することが求められている（文部科学省, 2018）。これまでの正しい文法知識をルールとして身に付けさせる教育から、言語使用の場面を通して知識を習得し、そしてコミュニケーション活動において、文法知識を正しく活用することができる教育へと変化している。実際のコミュニケーション活動に表れる文法事項は、短文で提示されるものではなく、大きな文脈の中に位置づけられているものである。さらに日本語を「英語に訳す」ものでもない。その点を考えると、このような変化はごく自然な流れだと考えられる。また、大学入試も大きく変わっている。大学入試センター試験では、文法知識にフォーカスした空欄補充問題や整序問題が単独で出題されていた。令和2年度大学入試センター試験の外国語（英語）第2問では、空欄補充と整序問題に47点の配点がなされている（大学入試センター, 2020）。一方、令和3年度の大学入学共通テストでは、すべての設問が長文に関するものとなり、空欄補充問題や整序問題のような文法知識を単独で問う問題の出題がなくなっている（大学入試センター, 2021）。このように文法教育は大きな転換を求められている。

実際の言語使用の場面から文法の知識を学習するというのは、本来の学習の方法であり、ある意味で生きた文法知識を身に付けることができるであろう。しかし、日本のように英語の使用が極めて限られる環境下ではなかなか容易なことではない。さらに、日本語と英語は

認知類型を異にしており、出来事のとらえ方（事態把握）の違いが文法的な特徴の隔たりを生み出している。日本語と英語は認知類型とその文法への表出という観点から見ると、ある意味、対極的な言語であるとも言える（中村, 2004, 2009, 2019）。我々は母語を習得する際に、母語の出来事のとらえ方を習得している。長（2021, 2023）は、学習者の母語である日本語の出来事のとらえ方が影響し、英語の発話において時制や相の誤用が多く出現していることを明らかにしている。長（2021）では、比較的英語力が高い学習者の発話を分析し、学習者は動作が完了しているのか未完了なのかに着目し、未完了であると認知される出来事を現在時制で表現しており、一方完了していると認知される出来事を過去時制で表現することを明らかにしている。また長（2023）では、JEFL コーパスを分析し、現在完了、過去完了を単純過去の代わりに使用している誤用が多いことを報告している。そこで本研究は、認知言語学の枠組みに則り文法項目を当該言語の出来事のとらえ方の表れとしてとらえ、日本語を母語とする学習者が英語の出来事のとらえ方を身に付けるためにはどのようなアプローチが必要になるのかについて論じる。

2. 理論的背景

2.1 認知言語学と事態把握

1980年代以降に深化および拡大した認知言語学では、言語には認知主体である我々人間の認知能力が反映されていると考え、生物としての人間の身体感覚に起因する能力に基盤を置き、言語を体系的に記述している（e.g., Langacker, 1987, 2008）。Langacker（2008）は、“…

linguistic structure is seen as drawing on other, more basic systems and abilities (e.g., perception, memory, categorization) from which it cannot be segregated.” (p. 8)と述べ、言語特有の能力を仮定せず、人間が持つ認知能力や認知システムに基づいて説明を行っている。また、山梨 (2020) は「言語能力は、生物の延長としての人間の身体性を反映する一般的な認知能力によって動機づけられ、この認知能力からの発現の一形態として位置づけられる」(p. 13) としている。この考えに基づく、言語は言語使用から切り離された抽象的な原理や原則によって生成される文の集合体ではなく、言語使用者 (母語話者) と認知的要因が密接に関係して創発されたものと理解される。また、認知言語学は実際の言語使用を重視する。Langacker (1987)は、それを使用基盤モデルと呼び、次のような特徴を持つと説明する。

Usage-based approach Substantial importance is given to the actual use of the linguistic system and a speaker's knowledge of this use; the grammar is held responsible for a speaker's knowledge of the full range of linguistic conventions, regardless of whether these conventions can be subsumed under more general statements. A nonreductive approach to linguistic structure that employs fully articulated schematic networks and emphasizes the importance of low-level schemas. (p. 494)

Langacker (1999) は、文法は我々言語使用者が言語活動で聞いたり読んだりした言語表現の集大成であり、具体的な言語表現から抽出された認知パターンがボトムアップ的に立ち上がるとする。つまり文法は具体的な言語使用に内在しているものと考えられる。

さらに我々が言語を用いて目の前の出来事を表現する際、出来事に含まれるすべてのものを言語化するわけではない。例えば、バスとタクシーの衝突事故が発生したという出来事を描写する際、周りに歩行者がいたのか、天気はどうだったのか、事故現場にどのような建物が建っているのか等は、話し手と聞き手に関係がある (価値のある情報である) と話し手が考えない限り言語化されない。つまり、話し手は、出来事を認知し、それを言語化する過程において、自分に関わりの深いものを取り上げ言語化するということを行っている。我々が耳にする言語表現は、出来事と言語形式が客観的に結び付いたものではなく、話し手の事態把握というフィルターを通して見た結果であり、第一言語習得では言語活動を通して、出来事をどのように言語化するかを学習し習得している (e.g., Tomasello, 2003)。そして、この事態把握は言語ごとに特徴が見られる。特に英語と日本語では、池上 (1981, 1991, 2006a, 2006b) の「する」言語と「なる」言語、中村 (2004, 2009, 2019) の D-mode 言語と I-mode 言語、濱田 (2016) の場面外視点言語と場面内視点言語に代表されるように、出来事のとりえ方が大きくことなることが明らかとなっている。

濱田 (2016) は英語の出来事のとりえ方を次のようにまとめている。

英語話者は母語習得の過程で知覚された出来事をメタ認知処理することが慣習化している。そのために、「場面の外から見る感覚」で出来事を認識し、その出来事の中で一番目立って認識されたモノに視点を置いて「A は B だ (A is B)」あるいは「A が B を～する (A do B)」と表現する。(p. 32)

一方、日本語は次のような出来事のとりえ方をしている。

日本語話者は知覚と認識が融合した認知の仕方出来事を把握することが言語習得の過程で慣習化している。そのため場面の中に自分をおいて直接インタラクションすることで出来事をとらえる感覚で、出来事の中の何かをトピックにして、その出来事をそれが展開する順にたどるように言語化する。(pp. 26–27)

濱田 (2016) も述べているように、我々は母語の習得の過程で、これらの出来事のとりえ方を言語とともに習得している。そのため、同じ出来事を日本語と英語で言語化した際に違いが生じる。池上・守屋 (2009) では、出来事のとりえ方の言語による違いとして、以下の例をあげている。

〔戦死者に言及して〕
彼は戦争で死んだ。
He was killed in the war. [直訳：彼ハ戦争デ殺サレタ。]

〔道に迷って尋ねるとき〕
ここはどこですか。
Where am I? [直訳：私ハドコニイマスカ。]
(池上・守屋, 2009, p. 18)

池上・守屋はこれらの表現はそれぞれの場面において日本語母語話者、英語母語話者が最も一般的に使用する表現であると述べている。例えば、「ここはどこですか。」という日本語を “Where is here?” と英語にそのまま訳した場合、その英語は非文法的な文というわけではない。しかし、母語話者は使用しない。これまでは、このような「普通は言わない」というものに対して、明確な説明ができていなかったが、認知言語学の深化・発展とともに、言語特有の出来事のとりえ方という観点から、これらの違いについても説明可能となっている。

2.2 第二言語習得と使用基盤モデル

認知言語学が立脚する使用基盤モデルでは、言語は言語使用を通じ、個々の言語使用からスキーマを抽出

し、それが慣習化され定着することで言語構造が構築されるとする (Langacker, 1987, 2008; Tomasello, 2003)。さらに、学習者が接するインプットが重要であることも指摘されている (Tomasello, 2003; Tyler, 2008)。しかし、既に母語を習得した学習者における第二言語習得は、第一言語習得とは異なることが容易に想像がつくであろう。

Ellis and Cadierno (2009) は第一言語習得とは異なる点として、学習対象言語の母語話者が出来事や状況を解釈し言語化する一連の流れにおける意味と形式の対応関係を第二言語学習者は新たに構築する必要があることをあげている。我々は母語の習得過程において、母語の出来事のとりえ方とそれに対応する言語形式の対応関係を習得している。第二言語を習得するということは、場合によっては母語とは全く異なる出来事のとりえ方と言語形式の対応関係を習得する必要があることを意味している。このような使用基盤モデルの立場からの第二言語習得研究の代表的なものが Nick C. Ellis を中心とする一連の研究である。Ellis and Ferreira-Junior (2009) は、動詞の項構造に着目し英国在住の移民労働者と英語母語話者との対話データを分析した。その結果、英語母語話者の発話においてトークン頻度が高く、かつプロトタイプ性も高い動詞が、移民労働者の発話でも同様に多く出現し、そこからより頻度やプロトタイプ性が低いものに移行することを明らかにしている。また、Ellis et al. (2014) はドイツ語、スペイン語、チェコ語を母語とする英語力が高い学習者を対象に動詞の項構造に関する心理実験を行っている。その結果、頻度や偶発性、プロトタイプ性が大規模コーパスデータと同様のパターンを示すことを明らかにしている。これまで英文法の指導は、概してルールを教えるもの为中心となってきたが、彼らの一連の研究からも、第一言語習得同様、言語使用に基づいた学習が重要だと言える。

しかし日本のように授業外でのインプットが極めて貧弱な外国語学習 (EFL) 環境下では、どのように対象言語の意味ある使用を学習者に与えることができるのか、また、十分なインプットをどのように確保するかという点で課題がある。EFL 環境下において、実際の豊かな言語使用事例に触れさせる学習方法の 1 つとして、多読学習があげられる。多読学習が学習者にどのような影響を与えているかについて様々な研究がなされているが、その中で英語力が向上したとの報告もある (林他, 2020 ; 西澤他, 2006 ; 大槻・高瀬, 2012)。しかし、これらの研究では、文法テストや検定試験等を用い「文法知識」の有無を測定しており、英語の出来事のとりえ方の習得という観点での検証は十分になされていない。つまり、学習者が英語を用いる際、眼前の出来事を英語の事態把握を用いてとらえ、そして言語化できるようになっているかについては明らかではない。

2.3 日本語と英語の時制の違い

池上 (1981, 1991, 2006a, 2006b)、中村 (2004, 2009, 2019)、濱田 (2016) が述べるように、出来事のとりえ方の違いは、文法項目として表れている。例えば中村 (2004) は、英語の非人称構文や主語優先の特徴は、D-mode 言語である英語の出来事のとりえ方に起因するものだとしている。また、濱田 (2016) が詳述している視点の取り方の違いは、日本語と英語の時制の違いとしても表れている。以下の表 1 は英語の時制と相の関係を表したものである。

表 1
英語の時制と相の関係

相	時制		
	過去	現在	(未来)
単純	過去	現在	(未来)
進行	過去進行	現在進行	(未来進行)
完了	過去完了	現在完了	(未来完了)

長 (2023), p. 108 表 1

英語では、「時制」と「相」という 2 つの要素が独立して存在する。場面外視点を取るため客観的な時間軸が存在し、その時間軸に出来事をポジショニングし言語化している。一方、日本語では、英語のような独立した関係はない。表 2 は日本語の時制と相の関係を表したものである。

表 2
日本語の時制と相の関係

	過去・完了	現在	未来
動態述語	タ形	テイル形	ル形
静態述語	タ形	ル形	ル形

菅井 (2012), p. 57 表 17

場面内視点を取る日本語は客観的な時間軸が存在しない。以下の例文から分かるように、過去の出来事が必ず「タ形」で表されるわけではない。

彼らは米国へ行くときは結婚していたが、帰ってきたときは離婚していた。
(樋口, 2001, p. 68)

樋口 (2001) の例文では、「米国へ行く」、「結婚する」、「(米国から) 帰ってくる」、「離婚する」という 4 つの出来事が描写されている。いずれも過去の出来事である。しかし、日本語では過去の出来事に対しル形も許容される。同じ「とき」に接続する出来事にもかかわらず、「米国へ行く」はル形が、「(米国から) 帰ってくる」はタ形で出現している。樋口では、「行く」が許容されるのは、日本語では「行く」という行為がこれから起きるという視点を取っているためだと主張している。樋口は「日本語の時制は、そもそも発話時や主節時といった特定の基準時を持つのではなく、事態との相対的な位

置に視点を持ってくる」(p. 57)と主張している。また、池上・守屋 (2009) は、このような日本語の出来事のとりえ方と時制の関係について次のように説明している。

ル・タはそれぞれ非過去、過去などを表す形式だと言われることがよくあります。これは物理的・客観的な時間軸を前提にした考え方です。日本語の場合、…説明がつかないこともしばしば見られます。それは日本語では事態の出来の事実を、物理的・客観的な時間軸にあてはめてとらえるというより、むしろ事態の出来を発話者のイマ・ココの認知との関わりにおいて捉える傾向があるためです。(p. 101)

さらに児玉 (2012) は、「日本語の時間が話し手の<イマ>しかない」(p. 229)のに対して、「英語は発話時(<イマ>)を起点に時間差が時制によって示される」(p. 229)と述べ、日本語と英語とで出来事をとらえる時間の概念が異なることを述べている。

日本語母語話者は母語である日本語を習得する際に、客観的な時間軸を用いない出来事のとりえ方を身に付けている。一方、英語は客観的な時間軸を用いて出来事をとらえる言語であるため、英語を習得する際に、出来事のとりえ方を習得するということが必要不可欠である。

以上の点を踏まえると、これまでの「ルールとしての英文法教育」ではカバーしきれなかった出来事のとりえ方の違いを、どのように学習者に習得させるかを考える必要がある。そこで本研究では、英語の時制に焦点を当て、英語と日本語の事態把握の違いに基づく時制の指導と多読によるインプットが学習者の出来事のとりえ方の習得にどのような影響を与えているかについて考察を行う。

3. 方法

3.1 参加者および多読学習方法

本研究には日本語を母語とする英語学習者 45 名が参加した。彼らの専門は英語であり、英語のレベルは CEFR で A2 から B1 である。参加者達は、英語のリスニング、リーディング、スピーキングの授業に加えて英語の文法に関する授業を受講している。授業は認知言語学の知見に基づいて英文法を学習するもので、前期に主に五文型、関係詞、時制、態を学習し、後期に助動詞、不定詞・動名詞、可算・不可算名詞、冠詞、some/any を学習している。この授業は英語学の授業ではないため、認知言語学の専門用語は一切用いずに、ワークシート等を活用して指導を行っている。

参加者達は、授業の課題として多読学習を行っている。これは英文法を意味ある文脈の中で理解してもらうためである。オンライン上の多読学習システムを教材として指定し、学習開始時にプレイスメントテストを行った後、各自のレベルに合ったものを読んでいる。本を 1 冊読むごとに簡単な内容確認テストが出され、

60%以上のスコアを取ると本の語数が積算されるという仕組みとなっている。つまり内容が理解できなかった場合は、語数がカウントされない仕組みになっている。学習者は授業を通して英語の出来事のとりえ方が文法項目にどのように表出しているかについて知識を得るだけでなく、自分の英語力に応じた多読学習を通して個々の文法の実際の使用を大きな文脈の中で触れている。

本研究では、文脈を持つ言語使用事例のインプットが与える影響を見るため、夏休み中に多読を自主的に行ったグループ 8 名と自主的な多読を行わなかったグループ 8 名を抽出して分析を行う。両グループは 6 月に英語力テスト (VELC テスト) を受験しているが、2 グループ間に統計的に有意な差は見られなかった ($t = .39, n.s.$)。今回の分析対象となる学習者 16 名のうち、9 月および 11 月のライティング課題を行わなかった学習者 3 名を除いたため、最終的な分析対象者は多読を自主的に行ったグループ 7 名、自主的に行わなかったグループ 6 名となった。

3.2 時制に関する指導

7 月に 3 回に渡って英語の時制と相に関する指導を行った。学習者の母語である日本語は場面内視点であるため客観的な時間軸が存在しないが、英語は場面外視点を取るため、客観的な時間軸が存在する。指導にあたっては、まず、日本語と英語の例を示し、英語には客観的な時間軸があるが、日本語は完了・未完了の対比で出来事をとらえていることを示し、時間軸を用いて、英語の時制と出来事のとりえ方を結びつける練習を行った。それぞれの時制と相が、出来事をどのように時間軸に位置づけているか、また話し手がどのように出来事を眺めているかという点を強調した。特に英語の時制・相は、空間概念からのメタファーとしてとらえられる。

(濱田, 2016; 川瀬, 1999; 中野, 2023) そのため、空間領域における物体の見方を時間領域へマッピングさせるという形で出来事のとりえ方を教えた。例えば、現在完了は過去の出来事について描写しているが、時制は現在である。この出来事のとりえ方を空間領域にマッピングすると次のように説明可能である。我々は高い建物の上に「現在」おり、そこから空間的に離れたところにあるタワーを見る場合、我々がいる建物からタワーまでの街並みも視野に入ってくる。この見方を時間軸にマッピングしたものが現在完了である。「現在」という視点に立って、時間的に離れた過去の出来事を眺めているため、過去から現在までの時間軸上の出来事が視野に入る。学習者たちはこれまでに高い建物の上から街並みを眺めるという経験をしたことがある。このため、比較的容易に英語の現在完了の表す出来事のとりえ方を理解した。

3.3 ライティング課題

夏休み明けの 9 月および、その 2 ヶ月後の 11 月にラ

ライティング課題を実施した。JEFLC コーパスの叙述文のトピックにならない、9月の課題では「浦島太郎のその後の話」を、11月の課題では「高校の文化祭（大学の学園祭）」について書いてもらった（投野，2007）。なお、コロナ禍により学校行事が制限された世代であることを考慮し、11月の課題では、高校だけでなく大学の学園祭もトピックに含めることにした。また、9月の課題の際には、浦島太郎の物語の動画を見せ、内容を確認した後に、続きを書いてもらうようにした。参加者は与えられたトピックについて、辞書を使わずライティングを行うが、英語の単語が分からない場合は、ローマ字で表記することを許可した。制限時間は20分とし、制限時間終了後に回収した。

また、9月と11月のライティング課題において、異なる課題用紙を使用した。9月は図1に示す通常の罫線が入った用紙を使用した。一方、11月は、図2に示す用紙を使用した。

図1
9月の課題用紙

Title: _____

学籍番号: _____ 日付: _____ 氏名: _____

図2
11月の課題用紙

Title: _____

学籍番号: _____ 日付: _____ 氏名: _____

過去 現在 未来

図2は、文法の授業において時制を指導する際に授業で参加者に提示した時間軸を9月の課題用紙に追加したものである。2.3で述べたように、日本語母語話者は場面内視点を取る為、出来事をとらえる際に客観的な時間軸を必要としない。一方、英語は場面外視点を取る為、出来事をとらえる際に、客観的な時間軸が存在し、そこに出来事を位置づけて言語化を行う。英語における客観的な時間軸の重要性については、文法指導の際

に強調したが、出来事を描写する際に学習者に客観的な時間軸を意識づけるために、課題用紙に時間軸を提示した。なお、課題に使用する用紙が異なるだけで、11月の課題実施にあたって時制に関する指導や言及は一切行っていない。

3.4 分析方法

学習者のライティング課題において、時制を伴う動詞すべてについて、時制が正しく使えているかという観点から分類を行った。時制以外の要素が正しくなくても、時制のみの正誤で判断している。例えば、以下の文の場合、動詞 look の後には前置詞が必要であるが動詞の時制は正しいため、「正」と分類した。下線部が対象となった動詞である（以下、誤用例の表記は原文のまま）。

He looked his house and his parents after he became uncle.

また、時制を伴う動詞すべてを分析する為、主節だけでなく従属節の時制も対象とし、正誤の判断を行った。

Urashimataro who has came to his hometown opened the Tamatebox.

出現したすべての動詞の正誤を判断した後、9月のライティング課題と11月のライティング課題において、有意な違いがあるかどうかについて、カイ二乗検定を行い、検討した。

4. 結果と考察

4.1 学習者の多読量と誤用の出現頻度

学習者の多読量を表3に示す。夏休み前までの多読は、授業の課題として授業時間外に実施したものである。夏休み中の多読については、教員側の指示のもと学習者が自発的に行っているものである。9月以降の多読は、夏休み前と同様に授業の課題として行っているものである。

表3
学習者の多読量の平均語数

自主多読	夏休み前 まで	夏休み中	9～11月	計
あり (7名)	46,763.1	24,450.6	49,289.0	120,502.7
なし (6名)	3,9449.3	1,312.3	37,675.2	78,436.8

9月および11月に行ったライティング課題の平均語数を表4に示す。自主的な多読を行っていないグループは、9月の課題ではあまり記述出来ていなかったが、11月の課題では総語数を大きく伸ばしている。

表 4

ライティング課題の平均語数

自主多読	9 月	11 月
あり	121.7	155.4
なし	66.2	125.0

次に、動詞の出現頻度と完了相の出現頻度を表 5 に示す。今回のライティング課題では、完了相の使用が少なく、長 (2023) で観察されたような誤用は一般化が可能なほど出現しなかった。

表 5

動詞の出現頻度

	自主多読	時制		完了相
		現在	過去	
9 月	あり	41	90	4
	なし	30	38	3
11 月	あり	25	114	4
	なし	23	72	4

次に、時制の使用について、正誤の頻度を表 6 に、その分布を図 3 に示す。

表 6

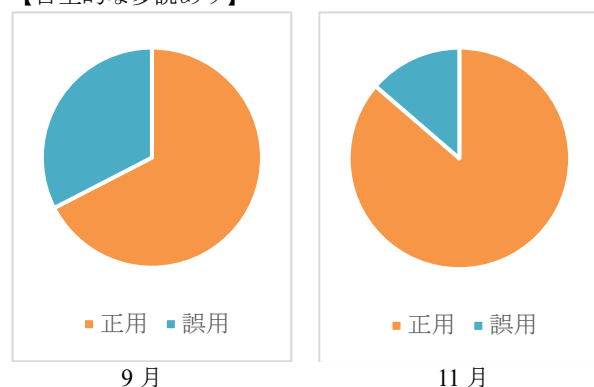
時制使用の正誤の頻度

	自主多読	正誤	
		正	誤
9 月	あり	89	42
	なし	35	33
11 月	あり	120	19
	なし	67	28

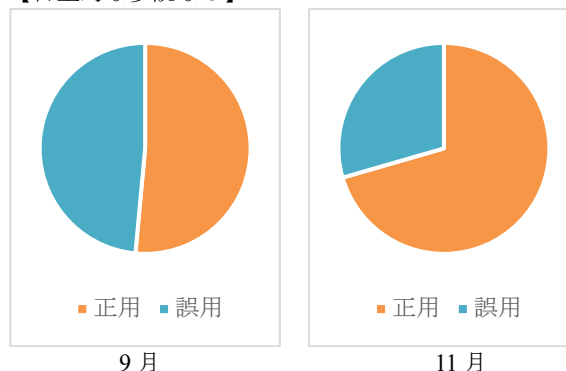
図 3

時制使用の正誤の分布

【自主的な多読あり】



【自主的な多読なし】



9 月から 11 月にかけていずれのグループも時制の誤用が減っていることがわかる。そこで、カイ二乗検定を用いて検定を行ったところ、有意な結果となった ($\chi^2(3) = 29.43, p < .01, \text{Cramer's } V = .26$)。

さらに、自主的な多読学習の有無が時制の使用の正誤に影響を及ぼしたかどうかを分析するため、9 月の課題について 2 グループを比較したところ、有意ではあるものの効果量は小さかった ($\chi^2(1) = 4.49, p < .05, \text{Cramer's } V = .15$)。また、英語の場面外視点による出来事のとらえ方を促す時間軸を明示した課題シートが影響を及ぼしたかどうか、各グループについて、9 月のライティング課題と 11 月のライティング課題を比較分析した。自主的な多読を行わなかったグループは、有意な結果となったものの効果量は大きくなかった ($\chi^2(1) = 6.15, p < .05, \text{Cramer's } V = .19$)。自主的な多読を行ったグループでも有意な結果となり、自主的な多読を行わなかったグループよりやや効果量が高かった ($\chi^2(1) = 13.05, p < .01, \text{Cramer's } V = .22$)。

4.2 学習者の誤用の特徴

次に、学習者の具体的な誤用について詳細に見ていく。次の 2 例は自主的な多読を行わなかったグループの 9 月の課題からのものである。

A man called Urashima opened Tamatebako. Then his hair turn white. He become very old.

最初の文において学習者は過去時制を使用しているが、2 文目以降は現在時制となっている。特に、2 文目は冒頭で “Then” という過去の 1 時点を表す副詞が用いられているにもかかわらず、現在時制を使用している。同様の過去時制から現在時制への移行は、他の学習者にも観察された。

He noticed that he became old person. He is very angry to Otohime and Kame.

この学習者は 1 文目では、主節も従属節も正しく過去時制を用いることが出来ているにも関わらず、2 文目では現在時制に移行してしまっている。

この、過去時制から現在時制への移行は、自主的な多読を行ったグループにおいても観察された。

He decided to write book about this life and sell it. His book become so famous around village. He become rich man.

また主節では正しい時制が使用されているにもかかわらず、従属節において時制が誤っている例も観察された。

He thought this is so fantastic experience.

両方のグループに共通して言えることは、時制の知識が欠如しているわけではないという点である。どの学習者のライティングにも過去時制を用いた動詞が出現していた。部分的には正しく用いることができていたということは、「過去の出来事は過去時制で表現する」という文法知識があるということである。それにもかかわらず、現在時制への移行が見られるということは、過去時制の使用を阻害する要因があると考えられ、そこには母語である日本語の場面内視点が影響しているものと考えられる。2.3 で見たように、日本語の場合、樋口 (2001)、池上・守屋 (2009)、児玉 (2012) が主張するように、出来事をとらえる視点が移動するという特徴を持つ。英語の過去時制のように過去の出来事と一対一対応している文法項目が存在しない。その結果、最初は過去の出来事であると認識し過去時制を用いているが、書き進めるうちに、母語である日本語の場面内視点に移行したためだと考えられる。

11 月のライティング課題では、英語特有の場面外視点を意識させるために、時間軸を課題用紙に明示した。いまだ誤用が観察される一方、誤用の出現頻度はどちらのグループにおいても減少していた。以下の例は、自主的な多読を行っていないグループのものである。

We had to practice very hard. Because, we want to won.

また、従属節の時制の誤用も 9 月の課題と同様に、観察された。

I thought I want to strive and succeed it.

その一方、時制を正しく使用できている例も増えている。

Last week, I joined Gakusai for the first time. There were a lot of people came for Gakusai.

接続詞で結ばれた従属節の中や、関係代名詞の中でも時制が正しく使用できるようになっている。

The corona virus was range around the world when I was enter the high school.

But my friends who was same high school went to Wataame shop to see me.

一方、自主的な多読を行ったグループも、誤用が大きく減少し、以下の例からもわかるように過去の出来事を過去時制を用いて表現することができている。

I went to this festival at last day. Because 3 and 4 had a game of my club in Yamaguchi.

また、従属節の中でも正しく使える例が増えている。

We had a mistake that we returned to high school on different time.

Because of corona virus, school festival wasn't held when I was a first-year and second-year student.

時制の誤用が減少しているということは、学習者が場面外視点を用いて出来事の描写できている可能性を示唆している。特に従属節の中で正しく使うことができていたということは、複数の出来事を複合的に描写する際に取る視点に母語の影響が少なくなり、学習対象言語である英語の視点を取れるようになっているのではないだろうか。

以上の結果から、英語の時制について、時制に関する文法知識の指導と十分な言語使用事例に触れることに加え、出来事のとりえ方の転換を促す時間軸を提示することに一定の効果がある可能性があることが分かった。

5. まとめ

日本語と英語は認知類型を異にしており、その違いが文法という形で表れている。時制はその代表的なものである。場面外視点を取る英語と場面内視点を取る日本語では、出来事のとりえ方が異なっており、その結果、日本語を母語とする学習者の時制の誤用につながっている。出来事のとりえ方について、認知言語学の枠組みの中で様々なことが明らかになっている。本研究では、その知見に基づき英語の時制の指導を行い、あわせて多読学習によって意味ある言語使用事例のインプットを増やした結果、自主的な多読を行ったグループの方が時制の誤用が有意に減少しているものの効果が大きいとは言えなかった。さらに、英語の出来事のとりえ方である「場面外視点」に存在する時間軸を課題用紙に明示したところ、どちらのグループも時制の誤用が有意に減少しており、一定の効果が観察された。時制に関する文法指導と文脈のある言語使用事例との十分な接触に加え、出来事のとりえ方の転換を促す仕組みを明示的に提示することで、学習者の出来事のとりえ方の習得に一定の効果がある可能性が示唆された。しかし今回の調査では、大きな効果が観察されておらず、今後、指導法や出来事のとりえ方を促す仕掛けについて、

さらなる改善が求められる。また、本研究では多読によるインプットの効果について、十分に検証ができていない。さらに、分析の対象となった学習者の数も限られている。学習者の数をさらに増やすとともに、外国語学習者にとって意味ある言語使用との接触とはどのようなものかについても、さらなる分析が必要であろう。

我々は母語の習得過程において、母語の出来事のとらえ方を習得している。そのため、外国語を学ぶ際に、新しい出来事のとらえ方を習得する必要がある、単に語彙や文法の知識を学ぶだけでは十分ではない。Hosokawa (2023) は、日本語の「コト」に着目し、英語から日本語、日本語から英語という使用言語の切替は Code-switching ではなく Mode-switching であると主張している。出来事のとらえ方もまさしく Mode の違いの 1 つであろう。本研究において、時間軸を提示することで「場面外視点」へと Mode の転換が促進された可能性があると考えられる。

日本語を母語とする学習者に英語を教えるということは、英語の文法を通して、母語とは異なる出来事のとらえ方を教えるということである。近年、文法を意味ある文脈の中で教えることが主流となってきたが、出来事のとらえ方の観点からアプローチすることも重要だと考えられる。しかし、認知言語学の知見をそのまま教育の現場に持ち込むことは、現実的ではなく、また不可能である。認知言語学の知識がなくとも、英語の出来事のとらえ方を指導できる教材の開発が求められる。

謝辞

本研究は JSPS 科研費の JP20H01296 および JP22K00701 の助成を受けたものである。

参考文献

- 長 加奈子. (2021). 「英語学習者の母語の事態把握が与える影響—英語の時制に着目して」. 『福岡大学人文論叢』, 53(1), 35–56.
- 長 加奈子. (2023). 「完了形の使用と日本語母語英語学習者の母語の事態把握の影響—認知言語学の観点から」. 『比較文化研究』, 151, 105–117.
- 大学入試センター. (2020). 『大学入試センター試験 外国語〔英語（筆記）〕』.
- 大学入試センター. (2021). 『大学入試共通テスト 外国語〔英語（リーディング）〕』.
- Ellis, N., & Cadierno, T. (2009). Constructing a second language: Introduction to the special section. *Annual Review of Cognitive Linguistics*, 7, 111–139. <https://doi.org/10.1075/arcl.7.05ell>
- Ellis, N., & Ferreira-Junior, F. (2009). Constructions and their acquisition: Islands and the distinctiveness of their occupancy. *Annual Review of Cognitive Linguistics*, 7, 187–220. <https://doi.org/10.1075/arcl.7.08ell>
- Ellis, N., O'Donnell, M. B., & Römer, U. (2014). Second language verb-argument constructions are sensitive to form, function, frequency, contingency, and prototypicality. *Linguistic Approaches to Bilingualism*, 4, 405–431. <https://doi.org/10.1075/lab.4.4.01ell>
- 濱田英人. (2016). 『認知と言語—日本語の世界・英語の世界』. 開拓社.

- 林幸代・丸尾加奈子・川瀬義清・長加奈子. (2020). 「多読学習が英語読解に与える影響」. 『LET Kyushu-Okinawa BULLETIN』, 20, 27–38.
- 樋口万里子. (2001). 「日本語の時制表現と事態認知視点」. 『九州工業大学情報工学部紀要 人間科学篇』, 14, 53–81.
- Hosokawa, H. (2023, December 2). *Not code but mode switching from Japanese to English: Enhancing English interactional proficiency among Japanese learners of English* [paper presentation]. 第 51 回九州英語教育学会熊本研究大会, 熊本.
- 池上嘉彦. (1981). 『「する」と「なる」の言語学—言語と文化のタイポロジーへの試論』. 大修館書店.
- 池上嘉彦. (1991). 『<英文法>を考える—<文法>と<コミュニケーション>の間』. 筑摩書房.
- 池上嘉彦. (2006a). 「<主観的把握>とは何か—日本語話者における<好まれる言い回し>」. 『言語』, 35, 20–27.
- 池上嘉彦. (2006b). 『英語の感覚・日本語の感覚<ことばの意味>のしくみ』. 日本放送出版協会.
- 池上嘉彦・守屋三千代. (編著). (2009). 『自然な日本語を教えるために—認知言語学をふまえて』. ひつじ書房.
- 川瀬義清. (1999). 「英語の完了形—認知的観点から」. 『西南学院大学英語英文学論集』, 40, 115–136.
- 児玉徳美. (2012). 「日本語と言語類型」. 『立命館文学』, 627, 244–221.
- Langacker, R. W. (1987). *Foundations of cognitive grammar, vol. 1, Theoretical prerequisites*. Stanford University Press.
- Langacker, R. W. (1999). *Grammar and conceptualization*. Mouton de Gruyter.
- Langacker, R. W. (2008). *Cognitive grammar: A basic introduction*. Oxford University Press.
- 文部科学省. (2018). 『高等学校学習指導要領（平成 30 年告示）解説 外国語編 英語編』.
- 中村芳久. (2004). 「主観性の言語学—主観性と文法構造・構文」. 中村芳久. (編). 『認知文法論 II』 (pp. 3–51). 大修館書店.
- 中村芳久. (2009). 「認知モードの射程」坪本篤朗・早瀬尚子・和田尚明. (編著). 『「内」と「外」の言語学』, 開拓社.
- 中村芳久. (2019). 『認知文法研究—主観性の言語学』. くろしお出版.
- 中野研一郎. (2023). 『シン・認知文法論入門：イノベティブ思考が明かす英語と日本語』. 開拓社.
- 西澤一・吉岡貴芳・伊藤和晃. (2006). 「英文多読による工学系学生の英語運用能力改善」. 『電気学会論文誌 A（基礎・材料・共通部門誌）』, 126(7), 556–562.
- 大槻きょう子・高瀬敦子. (2012). 「多読用図書教材が英語習得に及ぼす影響—L1 児童用英語絵本と中学校英語教科書との違い」. 『英語教育研究』, 35, 63–78.
- 菅井三実. (2012). 『英語を通して学ぶ日本語のツボ』. 開拓社.
- Tomasello, M. (2003). *Constructing a language: A usage-based theory of language acquisition*. Harvard University Press.
- 投野由紀夫. (編著). (2007). 『日本人中高生一万人の英語コーパス—中高生が書く英文の実態とその分析』. 小学館.
- Tyler, A. (2008). Cognitive linguistics and second language instruction. In P. Robinson & N. C. Ellis (Eds.), *Handbook of cognitive linguistics and second language acquisition* (pp. 456–488). Routledge.
- 山梨正明. (2020). 「認知科学と認知言語学」池上嘉彦・山梨正明. (編). 『認知言語学 I』 (pp. 1–16). ひつじ書房.

日本人大学生の会話に対する前提認識が英語会話に与える影響 —英語会話データとフォローアップインタビューからの考察—

大谷麻美*

*京都女子大学
ootanm@kyoto-wu.ac.jp

要旨

本研究は、日本人大学生が会話の前提として持っている規範や認識と、それらが会話に与える影響を、実験会話とフォローアップインタビューの分析から明らかにした。データは、年代も英語の能力も類似する日本人大学生とアジアからの留学生との間の初対面の英語会話5本（計2時間30分）である。参加者はともに英語の非母語話者であったにもかかわらず、ほとんどの会話は留学生によってリードされ、日本人学生は発話が少なく、聞き手になり、受け身になりがちであった。会話後のインタビューの分析では、日本人大学生の英語会話に対する認識には、母語の日本語の規範からの転移の影響が見られた。この前提認識が、英語会話への円滑な参加を難しくしている可能性を指摘した。さらに、この結果を受けて、英語の会話の規範とそれに沿ったインタラクション方法の明示的な指導が必要なことも指摘した。

キーワード：英語会話、規範、認識、インタラクション、明示的指導

1. はじめに

日本人の英語の話し方に関して、大人しい、控えめだという意見をしばしば耳にする。実際に日本人のビジネスパーソンの英語を調査した寺内他（2008）は、「積極的な外国人に先を越される日本人のメンタリティ、穏やかさ、性格的弱さ」を課題として指摘している。また、Iwata（2010）は、談話分析の観点から日本人と英語母語話者の間の英語会話を分析して、日本人は自分のことを話す自己開示が英語母語話者と比較して非常に小さく、その結果、会話を英語母語話者にリードされてしまっていると指摘している。

また、日本人の大人しさは、単に話す量が少ないという問題だけではなく、相手との人間関係や信頼関係を損なう結果にもつながりかねない。例えば、Otani（2007）では、日本人と話したアメリカ人が、積極的に話さない日本人について、「日本人の話し方は失礼だ、もう話したくはない」と述べる事例が挙げられている。つまり、日本人の大人しさは、友好的な人間関係を損ないかねない問題でもある。

その原因としては、日本人の気質や、英語の言語能力の不足、指導方法の問題など様々なことが複雑に絡み合っていることが推測できる。本論では、その中でも特に、日本人と話し相手との間にある、インタラクションの方法とその背後の文化規範の違いに焦点を当てて考察を行う。

Celce-Murcia（2007）は、コミュニケーション能力には、文法等の言語知識に関する *Linguistic competence*、決まり文句などの運用が出来る *Formulaic competence*、その言語が使用される社会の文化を理解する *Socio-*

cultural competence 等、いくつかの下位能力があることを挙げている。その中で、会話のやりとりを適切に行える *Interactional competence* の重要性も指摘している。

インタラクションは、事前準備が出来ない、その場での即興的なものであり、会話参加者双方が歩み寄りながら成立させるものである。従って、一人でスピーチやプレゼンテーションを行うための能力とは異なる能力を必要とする。そのため、ヨーロッパ言語共通参照枠（CEFR）においても、インタラクション能力はスピーキング能力とは分けてその重要性が指摘されている。

また、インタラクションの方法は、その言語社会の文化、価値観、規範と密接に結びついている。そのため、英語のインタラクション能力を身につけるには、その社会文化的能力を身につけることも不可欠である。

本研究では、日本人英語学習者が持つ会話に関する規範や認識を調査し、それらが彼らの英語会話のインタラクションにどのような影響を与えているのかを分析する。

2. コミュニケーションと規範

コミュニケーションの行い方は、その話者の文化規範（*cultural norm*）の影響を大きく受ける。Ting-Toomly and Chung（2005）は *cultural norm* を “the collective expectations of what constitute proper or improper behavior in a given interaction scene” と定義する。また、Spencer-Oatey and Kádár（2021）は、文化規範やそこから生じる期待に背くと人は悪く評価されるが、一方で、その規範や期待は文化によって異なる、と指摘する。つまり、学習言語のインタラクションの規範を理解し、それに添

って話すことは、円滑な会話を進める上で不可欠であるばかりか、相手との円滑な人間関係を築く上でも重要なのである。しかし、日本の英語教育では、英語のインタラクションの規範にまではほとんど踏み込めていないのが現状である。

3. 目的

そこで、このような日本人の英語インタラクションの現状を把握し、その指導の糸口を探るために、本研究では以下の3つの点を明らかにしたい。

- 1) 談話分析の手法を用いて、日本人の英語会話を分析し、彼らのインタラクションの問題点を明らかにする。
- 2) 会話のフォローアップインタビューの分析を通して、その問題点の背後にある、彼らの会話に対する認識や規範を明らかにする。
- 3) この問題点を克服するためにどのような指導方法が可能かを考察する。

4. 調査方法

日本人学生とアジアからの留学生（中国人、韓国人）の間で行われた5本の実験英語会話と、そのフォローアップインタビューをデータとして用いた。会話相手に、英語母語話者を避けて、アジアの留学生を選んだのには二つの理由がある。

一つは、母語話者と非母語話者との力関係を排除するためである。Fan (1994) は言語接触場面では、母語話者はホスト、非母語話者はゲストとして振る舞うことを指摘している。先述した Iwata (2010) の研究では、英語母語話者との会話で日本人の発話量が少ないことが指摘されている。しかし、それは、日本人のインタラクション方法に要因があったからのか、あるいは、日本人がゲストとして振る舞ったからなのかがはっきりしない。そのため、本調査では、あえてホスト・ゲストの関係になりにくい非母語話者同士の会話を調査することで、日本人のインタラクションの特徴をより明らかにしたい。

もう一つは、英語が様々な人々の間で共通語として使用されている現実を考えると、日本人の英語を英語母語話者と比較して分析するのではなく、ELF (English as a Lingua Franca) 環境でのありようとして見ることに意味があると考えた。また、この結果が ELF 環境を目指した英語教育のあり方を考える一助になり得ると考えたからである。

会話参加者は、19歳から34歳までの日本の大学に通う女性の大学生、大学院生であった。日本人学生は、全員が英語を専攻とする学生で、留学生は、英語の専攻ではないが、外国語として英語を学び話せる者であった。性別や年齢をそろえたのは、性や年齢差による力関係を極力排除するためである。

参加者の英語能力は、自己申告でおよそ CEFR の

A2-B1 レベルであった。ただし、実際に会話を観察すると、B2-C1 レベルと思われる者も、わずかではあるが混ざっており、全員が全く同等のレベルとは言えない状況であった。しかし、現実の英語の会話場面では、会話参加者が全く同じ英語のレベルになることは考えにくい。ため、今回はこの差は黙認することとした。日本人も留学生も、英語を主に大学の入学や卒業の試験対策として学び、実際の生活で使用する機会はほとんど無い者であった。

いずれも日本人学生1人と留学生2人の3人一組の会話であった(表1)。日本人と留学生は初対面であった。留学生同士には、面識のある者も混ざっていた。また、参加者探しに困難があり、留学生には相手を変えて数回実験に参加してもらった。このことが結果に幾分か影響を与えた可能性は否定できない。

会話は、先生宅のホームパーティーに招かれて、日本人と留学生が初めて出会った場面という設定で行った。先生宅の居間での気楽な会話という設定で、会話のテーマなどは決めずに自由に話してもらった。各会話時間は30分で、5本で計2時間30分のデータである。

会話の後に、日本人と留学生とに対して別々にフォローアップインタビューを行い、会話の印象、困難点、気になる点等を自由に話してもらった。紙面調査ではなくフォローアップインタビューを選択したのは、インタビューのほうが、紙面調査だけではとらえきれない率直で詳細な意見や本音がうかがえると考えたからである。

会話とインタビューは録音、録画し、文字おこしをしてデータとして用いた。

表1
会話参加者 (J:日本人, C:中国人, K:韓国人)

グループ	No.1	No.2	No.3	No.4	No.5
参加者	J1	J2	J3	J4	J5
	C1	C1	C3	C1	C3
	C2	C2	K1	C2	K1

5. 分析

5.1 話題を膨らませない日本人

まず、各会話のフォローアップインタビューで、留学生から出た日本人学生への印象を見てみよう。参加者は、留学生も日本人学生もともに、会話を「楽しかった」と述べていた。しかし詳しくたずねると、いくつかの興味深いコメントが出てきた。留学生の何人かは、日本人学生について「大人しい」「話題が少ない」と指摘した。また中国人留学生の C3 は、「自分が話すぎたかも」と反省しており、「相手が話さないから自分が盛り上げようとした」とも述べていた。また、中国人留学生 C3 と韓国人留学生 K1 は、「会話が途切れることがあるので、仕切らないと、沈黙を埋めないと、と思った」と述べていた。日本人学生があまり話さないことに対して

困惑をし、会話を盛り上げたり繋いだりしなくてはならないと会話をリードしようとしていた様子うかがえた。

また、それ以外にも、C1 と C2 は、日本人学生の話し方を評して、「母国でああいう話し方をされると、私に興味が無いんじゃないかなと思う。友達になりたくないんだと思う。多分それ以上は話しかけないと思う。今回は会話のアルバイトだから一生懸命話しましたが。」とも述べていた。つまり、日本人学生のあまり積極的に話さない態度に対して不安や不満を抱いていたことがわかった。

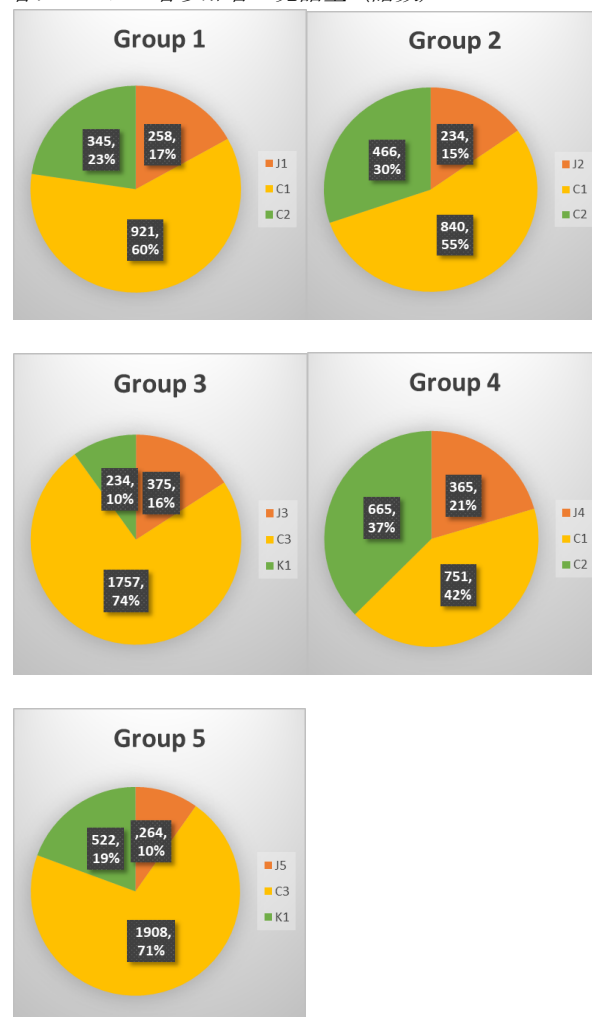
次に、実際に日本人学生が、どれほど会話に参加していたのかを分析した。以下のグラフは、各グループの会話の参加者ごとの発話量の分布である。会話の中には、大きく分けて意味交渉を行う発話（意味や内容を確認する発話など）と、情報交換を行う発話（相手に情報を求めたり提供したりする発話）がある。各会話から、参加者が特に自分に関しての情報提供を行っていた発話を抽出し、その単語数を数えた結果が図 1 である¹。なお、自分に関しての情報提供を行う発話とは、以下のように定義をした。1) 自分に関する情報（名前、趣味、嗜好、体験など）を提供する発話、2) 自分の属する集団（自分の大学、家族、出身地、国など）に関する情報を提供する発話、3) 自分の意見、考え、感想などを述べる発話²。このような発話は、自己を開示して互いを理解し合うためにとりわけ重要な発話だからである。

その結果、5つのグループのうち4つのグループ（グループ 1, 2, 4, 5）で、日本人の発話が3人の内で最も少なかった。その割合は、一番少ないグループ5で全体の10%、最も多く話していたグループ4でも21%であった。グループ3では、日本人が中国人に次いで2番目に話してはいるものの、その割合は、全体の16%で、決して多くはなかった。従って、フォローアップインタビューで留学生が述べていた、日本人はあまり話さず話題が少ない、留学生が沈黙を埋めるべく頑張って話した、という意見は、実際の発話の量でも裏付けられた。

また、日本人学生のインタビューからも、話題を見つけたり膨らませたり出来ずに苦労したという意見が出た。J1は「一番苦労したのは話題、話題ですかね。」、J3は「なんか、話題を振ったほうがいいんだろうと思ったんですけど、なんかもうそれまでに、自分の頭が本当パンクしそうになっちゃって。で、沈黙になって、えー、とっ、なんかしゃべらないと、とは思ったんですけど、（中略）これについて教えてとか、もっとつながられるはずの質問だったりとかがあると思ったんですけど、もうそれがあまり出てこなくて」、J5は「何か話さなきゃなっていう、思ったんですけど、何から話そう？というので迷いとか。この、いったん会話が途切れた後の自然な会話って何だろうとか、ていうのを考えてたら、なんか、どうしたらいいのかわちょっと分かんなくなったりとか。」と述べていた。日本人学生も話題を探して話

そうとはしていたようだが、それがうまく出来ずに苦労していた様子うかがえる。

図 1
各グループの各参加者の発話量（語数）



では、実際の会話はどのようなものであったのだろうか。以下の会話例を見てみよう³。

会話例 1 (Group 2)

- 01 J2: When did you come here?
 02 When did you...
 03 C1: I came here September, last month, last...
 04 J2: September?
 05 C2: Last...
 06 C1: Last, last year.
 07 J2: Wow! One year.
 08 C1: Almost 1 year, about.
 09 C2: Mm-Hmm, yeah. I came here
 10 two years ago...
 11 J2: Two years?
 12 C2: About - eh - almost three years.
 13 J2: Three years, very long. Yeah.
 (6-second silence)

14 C1: And what's your favorite fruit?

01-02 行目で、J2 が留学生に来日の時期をたずねている。それに対して、C1 と C2 は、03-12 行目で来日時期を答えている。しかし J2 は、自分からこの質問をしたのにもかかわらず、それに対して、“Wow! One year.” (07 行目)、“Three years, very long. Yeah.” (13 行目) と最小限の反応を示すだけで、それ以上にこの話題を膨らませたり、追加の質問をしたりすることはしなかった。その結果、13 行目の後に 6 秒の沈黙が生じてしまう。そして 6 秒後に C1 が新しい話題を持ち出すことで、ようやくその沈黙は解消される。

このように、日本人学生は相手の発話に関連づけて話したり、質問したりすることはほとんどなかった。その結果、話題が膨らまず、沈黙が生じてしまう事例が多く見られた。そしてその沈黙を破るのはほとんどが留学生であった。

5.2 質問をためらう日本人

上記の会話例 1 では、日本人学生は、01-02 行目で留学生に質問をしているものの、その後は、話題を掘り下げるような追加の質問をすることもなく、話題は発展しないままであった。このように日本人学生が相手に積極的に質問をしない背景には、質問をすることに対して、否定的な認識を持っているようである。上記の J2 同様に相手にほとんど質問をしなかった J3 は、フォローアップインタビューで、筆者からなぜもっと質問をしなかったのかとたずねられて、次のように答えている。J3「急にその話が終わって、質問っていうのもあんまり良くない、みたいな。そんな質問し過ぎないほうがいい、みたいな」。質問をすることに対してあまり良い印象を持っておらず、質問をすることをためらっていたことがうかがえる。

Shigemitsu (2021) は、日本語の会話の分析から、日本語では相手に質問することはあまり好まれないと指摘する。そしてその理由として、質問とは、相手に答えることを強いる可能性、相手の領域をおびやかす可能性、相手と対立を生む可能性、相手の話を遮る可能性、相手の能力のなさを明らかにする可能性があるからだとして説明する。

会話例 1 が示すように、J2 は最初は自分から質問をしたものの、それ以上のより詳しい情報をたずねる質問はしなかった。J3 同様に、J2 も質問をし続けることを好ましいことではないと判断していた可能性が大きい。

しかし、このような日本人のインタラクション方法は、先述した通り、C1 と C2 からは、話し相手に関心を示さない失礼な態度と解釈されていた。

5.3 沈黙を厭わない日本人

会話例 1 には 6 秒の沈黙が生じていたが、同じように長い沈黙が生じた会話は、本データにはいくつも見

られた。会話例 2 を見てみよう。この会話では、J4 が自分のペットの犬のことを話している (01-07 行)。そして、それについて C1 と C2 が犬と猫を比較するコメントを、面白おかしく述べている (08-16 行目)。その後、6 秒の沈黙が生じている。ここは、明らかに、C1 と C2 のコメントに対して J4 が何かを言うべき場所で、実際 C1 と C2 もそれを待っている様子であった。しかし、J4 は、自分の犬がボール遊びをすると飼い主の言うことを聞かなくなるという自分の話題 (01-07 行) について、“I'm so sad” (19 行目) と短い追加のコメントを述べるだけで、それ以上は何も言わず 6 秒の沈黙が生じてしまう。そして結局、その沈黙破ったのは C1 であった。J4 にターンがふられているのに、彼女はそれを埋めようとはしなかった。本データでは、このような比較的長い沈黙は、ほぼすべてが留学生によって埋められていた。

会話例 2

- 01 J4: Like and I called Coo, Coo. She –
02 he is very brat, but I call – I couldn't
03 call, she's little angry,
04 Coo, Coo, Coo. @@ And she – he play
05 ball. He likes play – playing ball.
06 So, I throw away the ball and he will
07 – he likes this. @@
08 C2: Sport games.
09 C1: You know, cat just can't @@
10 just can't have ball games. But you play
11 with the dogs.
12 C2: Maybe, cats brain is too small to have this
13 difficult mission to complete. They
14 only knows, yes, I eat, I sleep. I... @@
15 C1: And play with human and not what you
16 like and when I'm angry I bite them.
17 J4: Hmm.
18 All: @@
19 J4: I'm so sad.
20 All: @@
(6-second silence)
21 C1: So -um- if you start to live alone, do you
22 want to keep pets?

フォローアップインタビューで、このような沈黙についてどう思うかをたずねたところ、ある日本人学生からは次のような回答があった。J2「やっぱり、国も違って、初対面ですし。まあ、考える時間とか、そういう話すことを考えることは、まあ、あるのかな、お互いに、と思ったんで、そこは そんなに気にはならなかったです。」日本人学生には、比較的長い沈黙もさほど気にならなかったようである。

大谷 (2022) は、英語と日本語の会話を比較し、沈黙の長さや機能の違いを分析している。会話の中の沈黙の長さや頻度は、日本語の方が英語以上に圧倒的に長く多いことを明らかにしている。その理由として、日本

語の沈黙は、話題の転換を参加者相互が確認し合うための重要な働きをしているからだ指摘する。話題の終結部では、参加者は相互に沈黙を繰り返すことで話題の転換の可否を確認し合い、その後に話題転換が行われると述べている。日本人学生が沈黙をすぐに埋めようとしなのは、このような日本語の沈黙の機能が、かれらの英語会話にも転移されている可能性がある。

一方、大谷 (2022) によると、英語母語話者は沈黙を否定的に捕らえ、沈黙が生じるとなるべく即座に埋めなくてはならないと感じていると指摘する。このように、沈黙に対する規範は文化差があり、本データの留学生達も、沈黙には日本人ほどは肯定的ではなく、慌てて埋めようとする様子が幾度も見られた。

5.4 表面的な話題を好む日本人

本データの会話は初対面会話で、時間も 30 分と限られていたため、あまり深い話題を話し合う場面は少なかった。しかし、それを考慮しても、日本人が導入した話題は非常に表面的で浅い話題ばかりであった点は注目に値する。会話例 3 を見てみよう。J2 が 01 行目で新しい話題を導入し、留学生に寿司が好きかどうかをたずねている。それに対して C1 と C2 は、02-08 行目で各自の好みを伝えている。しかし、J2 は 09-10 行目で彼らの返事に理解を示したものの、それ以上、その話題を膨らませることはしない。そして 11 行目で次の話題（留学生がいつ来日したのか）に移ってしまっている。そのため寿司の話題は、留学生が寿司を食べられるのか否かを確認するだけの会話にとどまってしまっている。

会話例 3

01 J2: Ohhh, do you like sushi? Sushi?
02 C1: I do.
03 J2: Mm-hmm.
04 C2: Raw fish is little bit hard - difficult to
05 - I also - ah - eat - I always eat
06 - Mm...
07 C1: Cooked food?
08 C2: Cooked, cooked food.
09 J2: Cooked food. Cooked, food
10 hmmm.
11 When did you come here?
12 When did you...

このような、非常に表面的で深まらない話題は、会話例 3 のように日本人が導入した話題だけではなく、留学生が導入した話題に答える日本人の会話にも見られた。会話例 4 は、外国語学習について話している場面である。01-02 行目で、C2 が J1 に第二外国語の学習についてたずねている。03-04 行目で J1 はフランス語を学習していることを伝えている。しかし、最小限の情報提供をするだけで、それ以上の詳細は話さない。そのため 05-06 行目で留学生は、J1 がもう少し話すことを期待し

て待っている。C1 は “how . . .” とたずねて J1 から何かの情報を引き出そうとして、また、C2 は 06 行目で、“French” と J1 の発話を受け止めながら、J1 からの続きの発話を待っている。それを受けて、07-08 行目で J1 は追加情報（学習期間とフランス語が難しいこと）を提供している。しかし、やはり非常に少ない情報しか提供できていないため、C1 と C2 が次々に質問を重ねること（09、13、16-17、19、20-21、22、24 行目）話題を深めようとしている。しかし、J1 の返答は、“Yeah, yes, yes.” や “I - I’m not sure, but maybe related - maybe.” などと最小限の返答にとどまっている。英語を専攻している J1 にとっては、留学生以上に知識も情報も持っているような話題であるにもかかわらず、この話題を深めようとする様子はない。一方、留学生は、29 行目以降で、フランス語と英語を比較したり、同じヨーロッパの言語のドイツ語とフランス語の文法などを比較したり、自分の YouTube 鑑賞の体験談を話したり、自分が知っているフランス語を披露したりと話題を広げたり深めたりする糸口を提示している。

会話例 4

01 C2: Do you guys have any second - second
02 language to study in college?
03 J1: Ahhh, I started to - ah - I started to
04 study French.
05 C1: Ah-hah. How...
06 C2: French.
07 J1: One year, about 1 years ago. @@
08 But it’s difficult.
09 C2: Is that hard?
10 J1: Yes.
11 C2: Difficult.
12 J1: It’s difficult.
13 C1: French and English, which is difficult?
14 J1: French. @@
15 C2: I think it’s French. @@
16 C1: I heard that English came from French,
17 is that true?
18 J1: Yeah, yes, yes.
19 C1: That’s true?
20 C2: It’s not from - came from - ah -
21 German - Germany?
22 C1: Is that from German or French?
23 J1: Hahhhhhhh?
24 C2: It’s - it’s or it’s not?
25 C1: I don’t know. I - I heard that was from....
26 J1: I - I’m not sure, but maybe related - maybe.
27 C1: Maybe...
28 J1: French, German, English. I don’t know.
29 C1: Well, European language, they’re also the same.
30 C2: European.
31 J1: @@ Yeah, yeah.
32 C1: They’re also the same. @@
33 Germany, it’s difficult. I think it’s more
34 difficult than French. Is it? I - I don’t know.

35 It's just – you – sometimes you – you watch
 36 those people talking YouTube, like French
 37 people and –and German – and German, like
 38 you can still understand some of the words
 39 in French.
 40 J1: Mmm.
 41 C1: But nothing in German.
 42 C2: Seems like – seems like English, but En...
 43 C1: But different.
 44 C2: Ah – different.
 45 C1: But different, anyway. All I know is
 46 Bonjour, Bon Appetit. That's all, that's all.
 47 C2: But the European languages grammar are all
 48 – is also same – the same way it is wrote –
 49 it is true or not? European language,
 50 European languages grammar...

これ以外にも、留学生には自分の情報や体験談などを多くの言葉を用いて詳細に語る場面がしばしば見られた。時には、自分の国の政策や、日本の文化への自分なりの解釈などについても、文法などの誤りは多いものの、まとまった量を話し、それに対する自分の考えや意見を披露する場面が何度も見られた。

一方、日本が導入した話題は、会話例3同様に、好きな果物は何か、コーヒーと紅茶はどちらが好きか、コーヒーにはミルクや砂糖を入れるか否か、犬と猫のどちらが好きか等の話題が多かった。しかも、相手からの返答を聞くと、それに納得し、それ以上話題を膨らませない事例が多かった。そのため、各話題は深まらず、例3のように非常に短く、その結果、次々と新しい話題を探して導入する必要性が生じていた。

これについて、日本人学生にフォローアップインタビューでたずねてみたのが以下のやりとりである。筆者が、日本人学生に、会話が面白かったかどうかをたずねている場面である。

筆者 「なんか、最初のほうさ、コーヒーが好きとか、砂糖入れるとか・・・」
 J2 「はい、面白かったです、そこも。」
 筆者 「イチゴが好きとか、マンゴーが好きとか。ああいう会話っていうのは面白い？」
 J2 「面白いです。なんか、最初やっぱ緊張するんですけど、あの、簡単とかじゃないですけど、何て言うんですか、話しやすい、フルーツとか、飲み物とかって、あの、そんなに難しいことじゃないと思うので、自分の好きなこと、なんで。」

このことから、話題を深めるよりも、表面的な話題を日本人学生がむしろ面白いと感じていることがわかる。そして、初対面の会話では、このような難しくなく、話しやすい無難な話題をあえて選んでいた様子が見える。

6. 考察

このように、日本人学生のインタラクションの行い方には、留学生を困惑させたり、人間関係を形成しにくくする事例も見られた。その解決の糸口がどこにあるのかを考察する。

6.1 日本人のインタラクションに関する前提認識、規範の問題

以上の会話とインタビューの分析から、日本人学生は初対面会話で以下のような認識を持って話していたことがわかる。

- ・相手に質問をするのはあまり良くない。
- ・浅く無難な話題が好ましい。
- ・沈黙はあまり問題がない。

質問などを避けて相手や相手の話題に深く踏み込もうとはせず、沈黙などを用いて様子を見ながら、表面的で無難な話題に終始することで、相手と適度な距離を取りながら会話を進めている様子が見えた。また、会話例からもわかる通り、自分の情報を積極的に提供することとも控え、最小限の情報提供にとどめようとしていたこともわかる。

しかし、このようなインタラクション方法のために、話題を深めたり継続したりすることは難しくなり、話題はあまり発展せず、途切れがちになっていた。その結果、すぐに次の新しい話題を見つける必要が生じていた。しかし、相手に踏みこまずに新しい話題を見つけることは難しく、インタビューで日本人学生が述べていたように、話題を見つけられないという悩みが生じたようである。

このような会話の前提となる認識や、相手との距離の取り方の規範は、彼らの母語である日本語の規範からの転移の影響が大きいと考えられる。津田他(2015)は、日本語と英語の会話を比較しながら、各言語のインタラクションの特徴を明らかにしている。本データの日本人学生のインタラクション方法は、そこで明らかになっている、相手と距離を保とうとする日本語のインタラクション方法とほぼ一致している。

日本人学生は、英語でどのようなインタラクションが期待されているのかという英語会話の前提認識や規範を理解できていなかったようである。日本語のインタラクション方法のまま英語を話したことで、上記のような問題が生じたと考えられる。英語では、もっと自分の情報や考えを述べ、相手にも関心を示して情報を求め、相互に詳細な情報交換を行うことで、相手との心的距離を縮めることが期待されていたのである。

このような文化によって異なる会話の認識や規範を理解する能力は、先述した Celce-Murcia (2007) の Communicative competence の Socio-cultural competence に当てはまるだろう。また、その規範に沿ったインタラクションを実践できる能力が Interactional competence だと

いえるだろう。今回の日本人学生には、これらの能力が十分に習得されていなかった可能性が大きい。

6.2 指導方法の問題

一方、フォローアップインタビューで日本人学生 J3 は、「いつもなら話せるのに、(今回の会話では) ことばが出てこなかった」と述べていた。日本学生は全員が英語を専攻する学生で、普段は、英語母語話者教員が担当する英語会話の授業を多く受講している。この「いつもなら」とは、そういった授業の中の会話を指していた。J3 によると、その授業形態は、英語母語話者の教員が解説や指示を行うものの、会話の練習は同じ授業を受講する日本人の学生同士でペアやグループで行うというものであった。つまり、同じ日本語の規範を共有する者同士で、英語会話の練習を行っているということである。

このことからわかることは、たとえ英語の会話練習でも、彼らは日本語の規範のままで英語を話していた可能性が高いということである。そのため、話題に困ったり、話題を継続したりする際に問題が生じることも少なく、「話せる」と感じていたのであろう。また、英語の会話の規範に気づく機会もほとんどなく、その習得も出来ていなかったと考えられる。本データの日本人学生には、日本人以外の相手と 30 分にわたって英語で会話を行うのは今回が初めての経験であった。自分とは異なる会話規範を持つ相手と話して、はじめて会話がうまくいかないことに気づいたのである。

つまり、同じ規範を持つ者同士で英語会話の練習をしても、英語の会話規範に気づく機会は少ないことがわかる。そして、異なる規範をもつ相手と会話練習を行うことがいかに重要であるかも明らかとなった。

しかし、実際には日本の教室内の授業で、日本人以外と会話練習を行う機会を頻繁に設けることは難しい場合がほとんどであろう。その際は、英語の会話規範に関しての明示的な指導が効果的であろう。本データの日本人学生も、英語母語話者教員の英語に触れる機会は多かった。しかし、それだけでは、会話の規範にまでは気づくことは難しかったようである。そこで、明示的に、英語では相手との距離を縮める話し方が期待されていることを解説し、そのために必要なインタラクション方法を練習するとよいだろう。例えばその一案として、本会話データで日本人学生がうまく出来ていなかったインタラクション項目を参考にして、以下のような点を解説、訓練をすることも有効であろう。これらの訓練によって、本会話データに見られたような問題を減らすことが出来るのではないかと考えられる。

<英語の話し方の規範>

相手との距離を縮める話し方が求められている。そのために相手にもっと関心を示し、自分のことをもっと開示することが期待されている。

<そのための具体的インタラクション方法>

- ・自分について語る練習 (自分の情報、意見、考え等を詳細に話す練習)
- ・質問をする練習 (相手になるべく多くの質問をして、相手や話題に関心があることを示す練習)
- ・質問への返答の練習 (相手からの質問に対して、詳細に返答して自分の情報を詳細に提供する練習)
- ・沈黙を避ける練習 (沈黙が生じそうなら、相手への質問や自分の話題を提供することで沈黙を埋め、会話を継続させる練習)。

これらの訓練を行うことで、話題は深まり、会話は継続しやすくなるだろう。その結果、「話題がない、話すことが見つからない」という悩みもかなり回避できるのではないかと考えられる。

7. まとめ

会話のやりとりとフォローアップインタビューを詳細に分析することで、日本人が英語会話で大人しい、話さないと言われる問題の原因が、英語の文法や表現の能力の不足だけではなく、インタラクションの能力の不足も関わっている可能性が高いことを指摘した。そして、日本人の話し方は、単に発信する情報量が少ないという問題だけではなく、相手から話したいと思われたい、友人になりたいと思ってもらえない危険を孕んでいた。

このようなインタラクション方法の要因として、日本人学生がもつ会話の規範や認識が大きな影響を与えていると考えられる。そしてこれらの認識は、日本語の母語からの転移によるところが大きい。従って、文法や発音の指導と並行して、このような会話の規範や認識に対して、明示的な指導が必要だと考えられる。その具体的な指導項目の一案は、上記のとおりであるが、それ以外にも、まだ明らかに出来ていないインタラクション指導項目が残っている。今後も、このような談話の分析を通して、それらを明らかにしていく必要があるだろう。

英語は国際語であり、その多様性を認めるという観点から考えると、日本人の英語の話し方を矯正する必要があるのか、日本人は日本人らしい英語の話し方のままで良いのではないかという疑問もあろう。しかし、多様であるからこそ、その話し方も双方からの歩み寄りが必要だろう。「英語でのコミュニケーションを通じて、多様な人々と友好的な人間関係を形成する」ことが英語学習の目標の一つであるなら、少なくとも誤解を招いたり、人間関係を悪化させたりするような話し方は避け、相手の話し方に歩み寄る必要があるだろう。

(もちろん、これは相手の留学生にも言えることである。) 従って、英語と日本語のインタラクションとの違いを理解させることは極めて重要だと考える。その上で、学習者が場面や相手ごとに話し方を調整したり、選択したり出来るようにしておくことが肝要だろう。

本調査は、5 組のデータを対象とした小規模調査で、

会話参加者の性別も女性に限られ、会話の場面設定も初対面の会話だけであった。そのため、この結果を一般化することは難しい。そのため今後は、被験者の性別、英語のレベル、場面設定などを拡大し、より広範囲な調査を行う必要があると考えている。また、インタラクションの具体的な指導方法や効果を検証していくことも重要な課題である。

注

1. 発話の中での、言いよどみ、言い直しなどの発話は語数から省いた。
2. 意見や感想を述べる発話の中で、相手の発話の間に述べる“great” “good”などの短い単独発話は、語彙的あいづちと見なして感想の発話からは除外した。
3. 会話の文字化の記号については以下の通り。
@ 笑い () 筆者注

謝辞

本研究は JSPS 科研費 20K00849 の助成を受けたものである。

参考文献

- Celce-Murcia, M. (2007). Rethinking the role of communicative competence in language teaching. In E. A. Soler & M. P. S. Jordà (Eds.), *Intercultural language use and language learning* (pp. 41–57). Springer.
- Fan, S. K. (1994). Contact situations and language management. *Multilingua*, 13(3), 237–252.
<https://doi.org/10.1515/mult.1994.13.3.237>
- Iwata, Y. (2010). Pragmatic failure in topic choice, topic development, and self-disclosure by Japanese EFL speakers. *Intercultural Communication Studies*, 19(2), 145–158.
- Otani, M. (2007). Topic shift by Japanese and Americans: A cause of mis-interpretation in intercultural communication. *Memoirs of the Nara University*, 35, 69–84.
- 大谷麻美. (2022). 「日本語と英語における話題終結の相互行為—ポーズの使用に焦点を当てて」. 『京都女子大学人文論叢』, 70, 1–27.
- Shigemitsu, Y. (2021). Question-answer sequences in Japanese first encounters: Wishing to get to know new persons vs. dispreferred behavior of asking questions. In C. Ilie (Ed.), *Questioning and answering practices across contexts and cultures* (pp. 285–312). John Benjamins.
- Spencer-Oatey, H., & Kádár, D. Z. (2021). *Intercultural politeness: Managing relations across cultures*. Cambridge University Press.
- 寺内一・小池生夫・高田智子. (2008). 「企業が求める英語力調査」小池生夫研究代表『第二言語習得研究を基盤とする小、中、高、大の連携をはかる英語教育の先導的基礎研究—平成 16 年度～平成 19 年度科学研究費補助金(基盤研究(A)) 研究成果報告書』, 447–476.
- Ting-Toomy, S., & Chung, L. C. (2005). *Understanding intercultural communication*. Oxford University Press.
- 津田早苗・村田泰美・大谷麻美・岩田祐子・重光由加・大塚容子. (2015). 『日・英語談話スタイルの対照研究—英語コミュニケーション教育への応用』. ひつじ書房.

混合型研究法における妥当性検証の枠組み —教室におけるパフォーマンステストの形成的評価での事例研究—

松村香奈*

*鶴見大学

matsumura-k@tsurumi-u.ac.jp

要旨

混合型研究（以下 MMR）では、量的妥当性と質的信用性に加えて混合型研究の妥当性・正当性、方法論的整合性が求められ、Creswell and Plano Clark（2018）は、MMR デザインの手続きが、特定の理論や哲学の枠組みの中に組み込まれることの重要性を唱える。本研究は、ライティングとスピーキングのパフォーマンス評価 2 事例の実証研究を例に、MMR デザインを Assessment Use Argument（Bachman & Palmer, 2010）妥当性検証の枠組みで方法論的内容の検討という視点で論じた。量的・質的研究の結果を統合・解釈するためのジョイントディスプレイでのメタ推論から、統計的有意差が確認できない項目での群間での差異の要因推定等、MMR が、単一の方法では得ることが難しい課題の解明に貢献し、研究デザインと結果の解釈の質の確保のため、全体を包括する妥当性検証の枠組みの適用の有効性が示唆された。

キーワード：混合型研究法、方法論的整合性、形成的評価、Assessment Use Argument、メタ推論

1. はじめに

近年、言語教育研究分野では、量的ならびに質的研究アプローチでのデータの収集、分析、結果の統合を行う混合型研究法に高い関心が寄せられる。APA 第 7 版には、新たに混合型研究（以下 MMR）デザインの執筆基準が加わり、量的妥当性と信頼性に加えて混合型研究の妥当性・正当性、方法論的整合性が示される。Creswell and Plano Clark（2018）は、MMR デザインの手続きが、特定の理論や哲学の枠組みの中に組み込まれることの重要性を唱える。本研究は、MMR デザインを Assessment Use Argument（Bachman & Palmer, 2010）妥当性検証の枠組みで実施した 2 つの事例研究を例に、方法論的内容の検討という視点で論じる。

両事例は、教室内で特定の教育手法を用いた効果を測る実証研究である。事例 1 はペア型スピーキング（ $n = 69$ ）、事例 2 は、パラグラフライティング（介入群 $n = 23$ 、統制群 $n = 22$ ）で、それぞれ、「順次的収斂デザイン」、「介入を伴う複合的な収斂デザイン」を用いた。量的分析には多変量一般化可能性理論、多相ラッシュモデルでの当該評価尺度使用の妥当性を、また、クラスター分析で群間比較、2 要因混合型多変量分散分析で介入効果を測った。質的分析は、質問紙の自由記述に対してテーマ分析を行い、事例 2 では、談話注釈タグ付けツールにより学生の作文を樹状図化したものも分析対象とした。本研究¹では、MMR における研究デザインと結果の解釈の質の確保のため、全体を包括する言語評価妥当性検証の枠組みを適用することの有効性を論じる。

2. 理論的背景・先行研究

2.1 混合型研究法（MMR）の考え方

混合型研究法は、特定の学術分野にとどまらず、健康科学および社会科学において近年広く採用されている研究アプローチであり、量的研究により現象の全体的傾向を把握し、質的研究により個々人の多様な経験を把握できる。量と質の両面でのアプローチにより、研究領域での現象が起こる頻度、影響の程度や蓋然性の高い状況や様々な要因の関係を確証でき、活動プログラムの評価を可能にする手法とされている（Gibbs, 2007, 砂川他訳, 2017, p. 13）。この概念は、 $1 + 1 = 3$ （全体は部分の総和に勝る）（Fetters & Freshwater, 2015; ジョンソン・フェターズ, 2016）の言葉に表現される。留意すべき点として、MMR は従来の量的・質的データや分析結果の併用（combination）ではなく、意図的な統合（integration）を旨とし、両者の差異化のポイントは「哲学的視座の有無」と「（二つの研究アプローチの）統合によるシナジーの徹底的追及姿勢の有無」である（抱井, 2016, p. 3）。MMR の今一つの特徴に、ジョイントディスプレイという結果の提示方法がある。これは、質的研究・量的研究の結果を統合し、その解釈を示すために 1 つの図や表に両結果を配置したもの（Creswell, 2015）で、このディスプレイの作成によりメタ推論を行う（亀井, 2019 他）。

2.2 MMR での研究結果の質の評価

実証研究での結果の質の確保に必要な概念として、量的研究における妥当性（validity）、質的研究における信用性（trustworthiness）がある。中でも質的研究の評価

における中心的概念として充実性 (fidelity) と有用性 (utility) から成る方法論的整合性² (methodological integrity) では、論文における方法論のプロセスの議論の展開の重要性が示される (Levitt et al., 2018, p. 32)。これに加え、混合型研究法においては、正当性 (legitimization) の確立が必要とされ、Onwuegbuzie and Johnson (2006, pp. 56–60) は正当性に関する 9 つの下位レベル視点として以下を挙げている。(1) サンプル統合の妥当性 (Sample integration legitimation) ; (2) 内部者・外部者の視点に関する妥当性 (Inside-outside legitimation) ; (3) 短所最小化の妥当性 (Weakness minimization legitimation) ; (4) 順序の妥当性 (Sequential legitimation) ; (5) 変換の妥当性 (Conversion legitimation) ; (6) パラダイムの・哲学的妥当性 (Paradigmatic mixing legitimation) ; (7) 共約可能性への接近の妥当性 (Commensurability legitimation) ; (8) 統合の妥当性 (Multiple validities legitimation) ; (9) 社会政治的妥当性 (Political legitimation)。なお、用語の日本語訳は、ジョンソン・フェターズ (2016)³ の抱井訳に倣う。個別の混合型研究における妥当性の検証について、実際にどのように上記の項目を適用すべきかについて、ジョンソン・フェターズ (2016) は次のように述べている。

研究を評価する上では、上記に挙げた正当性の確立項目の一部のみが関わってくる。しかしながら、複数の項目を用いて正当性の確立を示すことが重要であり、そのために特定の研究に関連の深い適切な項目を選ぶことが必要になってくる。(p. 11)

そこで、教室内での形成的評価における言語テストの事例研究を扱う本稿で重視したいのは、(4)、(6)、および (8) である。個別の事例研究において、透明性 (transparency) をもって「研究の読者に具体的な言葉で研究がどのように進んだかわかる」(Flick, 2007, 上淵訳, 2017, p. 189) 程度を高めることは、研究の手続きおよび結果の「質」を高める上で重要であるのは言うまでもないが、異なる手法や分析を組み合わせる MMR では、研究デザイン⁴に関する事柄に加えて、質的・量的の一連の方法論的アプローチにおいて「縦断的視点」でつなぐことが特徴である (Kluge, 2001)。従って、上記の (4) 順序の妥当性を示すことは特に必要であろう。

(6) のパラダイムの・哲学的妥当性は、量的アプローチと質的アプローチの根底にある、研究者のパラダイムが、どの程度うまく組み合わせられているか、あるいは利用可能なパッケージとして融合されているかを検証するもので、本稿での言語テストという特定の文脈での分析、解釈においてどのような枠組みで研究を捉えるかを示すことは極めて重要であると考えられる。さらに、前節で挙げたジョイントディスプレイは、(8) の統合の妥当性に関わるもので、質的・量的研究結果の積み上げによって導き出されるメタ推論が理にかなっているかどうかの評価のポイントとなる。

本稿では以上 3 項目の妥当性検証の観点から論じることが、次節では、特に (6) に関して、教室の形成的評価を MMR のアプローチで言語テストを使用するという視点で、Bachman and Palmer (2010) の Assessment Use Argument (AUA) を説明するとともに、同枠組みを MMR で用いる意義を記す。

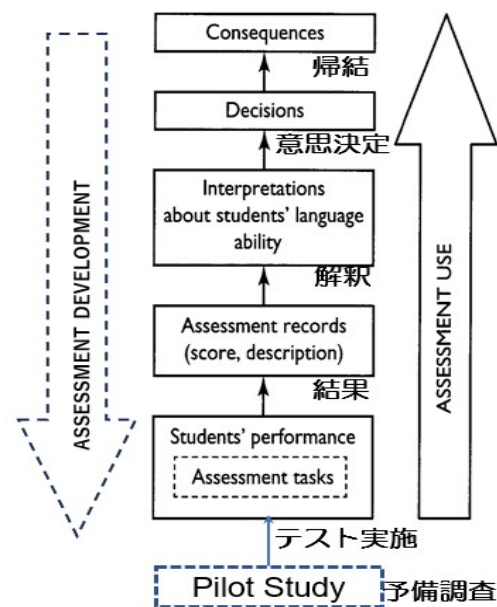
2.3 AUA の枠組み

Assessment Use Argument (AUA) は、Bachman and Palmer (2010) が提唱した言語評価の作成・開発および使用に関する指針、枠組みで、次のように定義される。

AUA は、評価における受験者の「パフォーマンス」、評価から得られる得点または質的記述である「評価の記録」、評価したい能力に関する「解釈」、下されるべき「(意思) 決定」、および評価の使用と下された決定の「帰結」の間の概念的つながりを特定する一連の主張で構成される。(p. 30, カギ括弧および日本語訳は本稿筆者による)

図 1 は、AUA の帰結 (Consequences) から評価パフォーマンス (Assessment Performance) への推論の繋がりを図式化したもので、現存する評価の使用 (Assessment Use) と評価の作成・開発 (Assessment Development) では流れが反対方向となる。

図 1
帰結からテスト実施への一連の推論の流れ



注. Bachman and Palmer (2010, p. 91) から引用。日本語注釈は本稿筆者により追加。

表 1 は、一連の推論の各主張 (Claims) の内容の記述で、研究の目的に適う主張を検証することで、当該の評価テストの妥当性を確認することができる。さらに、

AUA を用いた実証研究では、この Claim に沿った研究課題が設定される。

表 1

AUA における 4 つの主張 (Claim)

主張	Claim#	検証内容
帰結	Claim 1	このテストは有益である (か) beneficial
意思決定	Claim 2	このテストの意思決定は平等 (か)、価値観を反映する (か) value sensitive, equitable
解釈	Claim 3	このテストの解釈は意味がある (か)、適切である (か) meaningful, impartial, generalizable, relevant, sufficient
結果	Claim 4	このテストの結果 (test records) は一貫性、信頼性がある (か) consistent

注 Bachman and Palmer (2010, p. 103) を基に筆者が作成。

2.4 AUA と混合型研究アプローチ

本稿で注目したいのは、Bachman and Palmer (2010) が AUA の「評価の記録」として、量的な結果である「得点」だけでなく、「質的記述」を明記している点である。

... an assessment record, which is the score or qualitative description we obtain from the assessment... (p. 30)

ここでの「質的記述」というのは、一般的には Excellent, Good, Poor といった質の程度の評価結果が想定される。しかし、正誤で採点できる客観テストと異なり、ライティングやスピーキング能力を測るテストにおける評価は、通常、評価尺度における記述子 (descriptor) での定義に照らして決定されるため、得点化される場合であってもその定義に従って数値が付与される、つまり、質的記述が数値に変換されていることになる⁵。これらの産出能力のテストでは「量」と「質」の評価は地続きで、両者の統合が前提となる評価であり、AUA ではそういった概念が明記された枠組みであることから、混合型研究法との親和性は高いと考える。

加えて、AUA と混合型研究法 (MMR) の親和性について挙げるべき点は、研究の限界に対する議論である。フェッターズ (2021) は、混合型研究法を評価する 10 の基準の一つとして「混合研究法デザインの妥当性を脅かす要素 (調査の限界)」の議論を挙げている。一方で AUA は、その論理的基盤をトゥールミン論証モデル (Toulmin, 2003) に置くが、同モデルの特徴の一つとして、構成要素に「反駁 (Rebuttal) : 主張が成立しえない条件を予め提示する」があることが挙げられる。混合型研究法はパラダイム論争⁶以来、常にその方法論が批判にさらされることが少なくなっただけに、妥当性検証の視点として、批判的視点を内包することが求められてきたが、その点においてトゥールミン論証モデルを基盤に置く AUA と価値を共有すると考えられる。

MMR を支える哲学として一般的に用いられるのが

プラグマティズムで、これは「知識の有用性を重視するため、これを哲学パートナーとすることで質的・量的研究を統合することに矛盾を生じさせないという考えが背景にある」(抱井, 2015; 2021, p. 72)。一方、AUA が紹介される Bachman and Palmer (2010) の著書 *Language Assessment in Practice* は、タイトルに表わされる通り、一貫して「実用」が主眼に置かれ、更に、表 1 にある AUA の帰結 (Consequences) の主張 (Claim) は、「評価を用いた結果や下された決定は、利害関係者にとって有益なもの (beneficial) である」(Bachman & Palmer, 2010, p. 103) というもので、MMR の哲学とその価値を共有するものであると考える。

次節では、AUA の枠組みで行った MMR アプローチでのパフォーマンス評価の 2 事例を紹介する。

3. 事例研究 1 : ペア型スピーキングテスト

3.1 はじめに (事例研究 1)

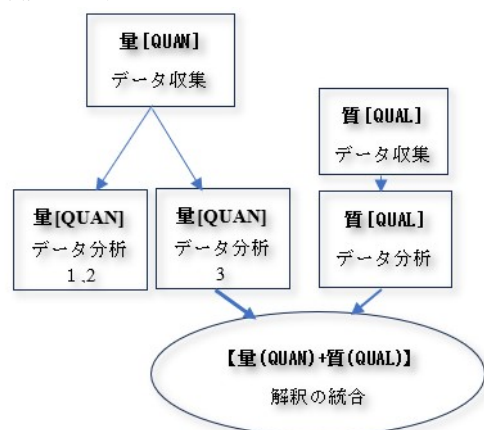
ペア型スピーキングテスト (Paired oral test) は、自分と同等程度の英語力である英語学習者同士が対話を行うテスト形式で、面接官と 1 対 1 で行うスピーキングテストとは異なり、相手とのより自然な会話を通し、お互いに協力して主体的な対話維持の努力をする機会が得られるという利点が指摘される (Galaczi, 2014)。本研究は、ペア型スピーキングテストを、教室でのアクティビティーの一環で、学習および診断的フィードバックの効果を測る形成的評価として実施し、教室内でのテストの実行可能性、評価尺度の実用性、さらには習熟度別の結果の特徴ならびに学生によるテストに対する認識についてプロファイリングすることで、言語学習支援における教育的示唆を提供することを目的とした (松村・守屋, 2019a)。

3.2 事例研究 1 の方法

a. 事例研究 1 のデザイン

本研究では、順次的収斂デザイン (Creswell, 2015) を採用した (図 2 参照)。このデザインにより、スピーキングの 4 観点 (やり取り、内容、正確さ、流暢さ)、各 5 段階評価での得点 (量的データ) で、評価尺度の信頼性 (量的分析 1) ならびに実行可能性 (量的分析 2) の 2 種類の分析による妥当性検証が行われ、妥当性、信頼性が確認された。次に、量的データを基にクラスター分析 (量的分析 3) で分けられた 4 つの習熟度別グループの分野別の特徴と併せて、テスト実施後のインタビュー、振り返りコメントシートによる学習者のテストに対する受け止めについてのコメント言語データ (質的データ) をテーマ分析によりコーディングし、質的・量的分析の解釈を統合することで、より包括的かつ詳細なプロファイリングが可能となった。

図 2
順次的収斂デザインのダイアグラム



b. 事例研究 1 の研究対象

都内の私立大学に通う日本語を母語とする英語学習者計 69 名⁷に研究の協力を得た。彼らの専攻は経営学部と社会学部で、2018 年 4 月から 7 月に本稿筆者が指導者として担当する英語 4 技能の育成を目的にした必修科目 3 クラスの受講生（2 年生～4 年生）である。全員が英語を専門とせず、英語習熟度は CEFR の A2 レベルが最多で、B1 レベルが若干名いる。海外在住経験 3 カ月以上が 2 名である（大学 1 年次に豪州に 1 年語学留学生 1 名、大学入学前に英国に通算 3 年滞在者 1 名）。

c. 事例研究 1 の手続き

本研究は次の手順により進められた。まず (1) 予備的研究として、調査評価表の作成、トピック選択⁸、テスト実施方法検討、評価者（筆者ならびに共同研究者）トレーニングが行われた。(2) テスト実施に当たって、トピックは授業で学習した内容に則し 2 つが選定され、事前に受験者に告知された。なお授業内では、それまでにテストと同形式のやり取りの練習が何度か様々なペアで行われた。(3) 次に本稿筆者である教員がファシリテータとなりペア型スピーキングテストが実施された。ペアはランダムで決められ、テスト当日発表された。(4) テスト内容は参加者の承諾を得て、後日評価のため録音された【量的データ収集】。(5) テスト実施後、アンケート・インタビュー調査実施（録音）【質的データ収集】。(6) 評価者が評価表に基づき得点化し、アンケート・インタビュー回答集計、(7) 得点から得られる学生の英語スピーキング能力を分析・解釈（ $n=69$ ）【量的分析 1・2】。(8) 探索的にアンケート回答内容分析・回答テーマ抽出（ $n=34$ ）【質的分析】、(9) 学生への feedback 内容決定・作成・返却、(10) 得点に基づきグループ化（ $n=38$ ）【量的分析 3】、最後に (11) ジョイントディスプレイの作成により、質的・量的分析の統合を行った。

d. 事例研究 1 の研究課題および分析方法

研究課題および分析方法は表 2 のとおりである。研

究課題は AUA の枠組みに則って作成され、図 2 にある順次的収斂デザインで実施された。

表 2

AUA の枠組みでの研究課題および各課題の分析方法

RQs	AUA	内容	分析方法
RQ 1	Claim 4	本研究で作成した評価尺度は適切に機能しているか	MFRM【量】 mG-theory【量】
RQ 2	Claim 3, 4	教室での実行可能性の観点から、十分な信頼性を確保した上で必要な評価者とタスクの数はいくつか	mG-theory【量】
RQ 3	Claim 2	得点から学生をいくつかのグループに分類できるか	クラスター分析【量】
RQ 4	Claim 1	ペア型スピーキングテストのグループごとの受け止めはどうか。同テストの実施は利害関係者に有益か。	テーマ分析【質】 ジョイントディスプレイによる分析【質・量】

注 松村・守屋（2019b）の内容を表にまとめたもの。

MFRM = 多相ラッシュモデル mG-theory = 多変量一般化可能性理論

3.3 事例研究 1 の結果

AUA に則り設定した各研究課題の結果は次のとおりである。【研究課題 1：評価尺度の信頼性】Many-facet Rasch Measurement（MFRM）のモデル適合度は 89.9%（58/69）でアンダーフィット（ミスフィット）が 4.4%（3/69）であった。スピーキングテスト等のパフォーマンステストでは比較的小ミスフィットの値が大きく出る傾向がある（小泉, 2016）ことを考慮しても、概ねモデルに適合していることが確認された。また、変数マップによる受験者の分布が概ね正規分布を示し、受験者能力分布平均ロジットが 1.63 であることから、学生にとって取り組み易く難易度が適切であったことが確認できた。4 観点難易度、評価者の厳しさ、ならびに 2 つのトピックに難易度についてもいずれも平均ロジットが 0.0 で問題がなかった。また multivariate Generalizability theory（mG-theory）の推定分散成分の分析結果から、4 観点の得点の分散変動要因が平均で 64%と他の変動要因より最も高いことから、この評価尺度で受験者の能力を推定できることを示唆した。以上から、使用評価尺度の信頼性は確認された。

【研究課題 2：評価尺度使用の実用性】mG-theory の D-study（Decision study: 決定研究）でテストデザイン（タスクと評価者の組み合わせの数）の分析結果から、4 観点での信頼度指数（ Φ ）が 0.8 以上となる受験者の能力差で説明できる組み合わせの中で評価者を 2 までに限定した箇所を確認したところ、評価者 1-タスク 3；評価者 2-タスク 2；あるいは評価者 2-タスク 3 という結果が得られ、いずれの場合も教室実施の実効性とローステイクなテストの性質から、これらの組み合わせが十分妥当で実用性があることが確認できた。

【研究課題3：学生の習熟度による分類】対象学生（ $n=38$ ）をクラスター分析で分類するための変数は、4観点での各得点および合計点を用いた。同分析で作成されたデンドログラム（樹形図）の結果、高習熟度群 Group 1（以下 G1）（ $n=7$ ）；Group 2（以下 G2）（ $n=17$ ）；Group 3（以下 G3）（ $n=8$ ）；低習熟度群 Group 4（以下 G4）（ $n=6$ ）に分類できた。

【研究課題4：質的研究結果および質・量の統合結果】

質的データのテーマ分析の結果、ペア型スピーキングテストのコメントは、＜感想＞と＜効果＞の2つに大別でき、コメント回答数はそれぞれ 65、110 であった。＜感想＞については、「上手くできた」などの自己肯定的内容（ $n=14$ ）、「適切に応えられなかった」などの否定的（反省的）（ $n=30$ ）、「語彙力を増やしたい」など今後の課題・改善点（ $n=21$ ）であった。＜効果＞は、「発音の大切さがわかった」などの言語・非言語に関するもの（ $n=65$ ）、「留学に行きたくなった」などの自律学習に関するもの（ $n=41$ ）、「新しい視点を発見できた」などのその他に分類されるものの各テーマが抽出できた。習熟度群別の特徴についても G1 から G4 の特徴を抽出した結果、図 3 の上部吹き出しにあるようにまとめることができた。G1 では、「この経験を英語学習に具体的に結び付けられる」といった言語面での自律的学習面での効果があり、G2 では、「テストというより対話の機会として捉え、やり取りの難しさ」を指摘する声があり、G3 では「自信がなく緊張度も高かったが、英語でやり取りできたことへの満足感」を示し、G4 では、「英語力不足を痛感し、発言にも苦労したものの、聞き取りができ相手の発言を理解できたことに満足感」を示すことがわかった。

以上の習熟度群別の質的分析結果と図 3 の 4 つの観点での得点グラフを統合したジョイントディスプレイによるメタ推論から分かることは、全ての観点で安定した能力を示す G1、全体に占める人数割合が最多の G2、得意・不得意分野で能力にばらつきの見られる G3、事前学習効果のある「内容」観点以外は英語能力の全体に低調な G4 と、量的分析・解釈では大きく特長を異にする学生が混在するが全ての学生において、ペア型スピーキングテストについて共通に感じていることとして、「自分の発音への注意、相手の発話への傾聴、記憶、そしてこれらに注意を払いながら適切に反応し、協力的にやり取りを行うことの重要性への認識」があることが確認できた。

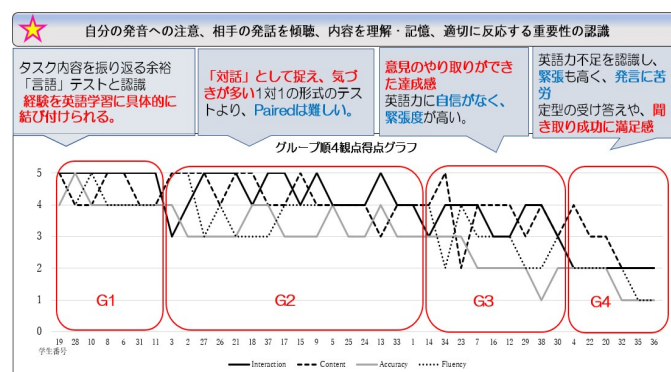
3.4 事例研究1の考察、限界と今後の展望

半構造化された評価基準が明確な議論型タスクという条件下ではあるが、教室での Paired oral テストが、一定程度信頼性の確保されたテストとして、実行可能性が示唆された。限界点としては、トピックを事前に告知するなど、即興性に関する能力が必ずしも測れていない点と、4 観点の評価はフィードバックとしては丁寧で

はあるが、その場での採点は難しく、録音での評価が評価者にとっては負担感がある点である。今後は 4 観点での「やり取り」の評価記述子の精査を含め、より単純化した評価尺度での有効性も探る必要がある。

図 3

質的・量的分析の統合の結果（ジョイントディスプレイ）



注 松村・守屋（2019a）掲載の図を本稿用に一部改変。

3.5 事例研究1の結論

今後の検討課題はあるが、質的・量的分析の統合の結果の解釈として、学生にとって、習熟度レベルに差はあっても、どの群の学生もテストからスピーキングの対話の重要性を認識できることから、今後の英語学習にとって有益であると推定できた。また指導者（教員）にとっては、学生の習熟度層の分布比率やスピーキング 4 観点の達成度合いの把握・理解の向上ができることから、今後の指導にとって有益であると判断できた。

4. 事例研究2：議論パラグラフライティングテスト

4.1 はじめに（事例研究2）

日本の英語学習者は、大学の英語学術的ライティング授業において、パラグラフライティングの主題文、支持文、結語文の基本的な論理の展開の基礎を系統的に学習する。英語らしい論理的構成では、主語と述語の選び方、言語情報の提示する順序や構成、接続詞の選択や文のつながり方など、まとめたり一貫性のある文を書く難しさがある。それに加え、ライティングにおける構成評価は、様々な評価観点から成るため、その評価および評価結果を効果的に提示することの難しさが指摘されている（Knoch, 2007 他）。教室での形成的評価に関する研究である事例研究2では、「構成」の要素である「一貫性」を、他の文の解釈に基づく談話の意味的特性（van Dijk, 1977）とした上で、文章が適切な順序で配置され、切れ目なく繋がり、文の間で淀みなく次々とアイデアが提示されることと定義し、学生への英語ライティング指導における形成的評価としての図示化された形成フィードバックが学生の英作文能力および作文修正行動や一貫性の観点への理解に及ぼす影響を、介入群と対照群を設けて探究した（Matsumura, 2023）。

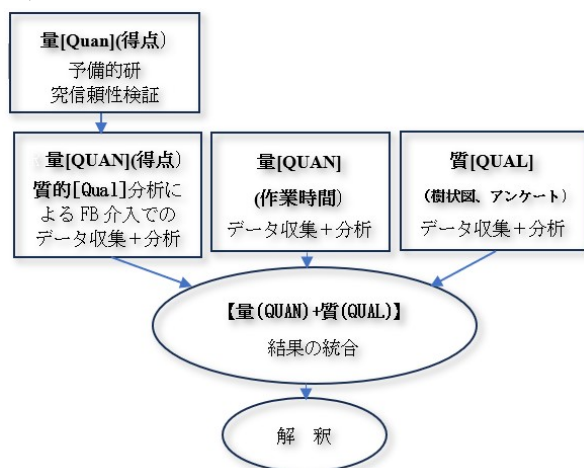
4.2 事例研究2の方法

a. 事例研究2のデザイン

本研究では、樹状図での教員フィードバック (FB) の介入効果を測ることに主眼を置きながら、質的・量的研究法の特長を活かして分析結果を統合するために、介入を伴う複合的な収斂的混合研究法 (Creswell, 2015; 抱井, 2015) を採用した。本研究の MMR ダイアグラムは、複数のデータ収集と分析が組み合わされて複雑だが、大きくは、Study 1 (質的分析介入を伴う量的研究)、2 (量的研究)、および 3 (質的研究) の 3 つから成り、それを最終的に統合するものである。本稿では紙面の都合上、概略版デザインを図 4 に示す。

図 4

介入を伴う複合的収斂的混合研究法デザインの概略版ダイアグラム



注 Matsumura (2023) 掲載のダイアグラムを本稿用に改変。

b. 事例研究2の研究対象

本研究は、東京の私立大学で教育学を専攻する大学2年生2クラス計50名(うち研究対象者45名)を研究対象とした。介入効果を測るため、一方を介入群($n=23$)、他方を対照群($n=22$)とした。両クラスの英語能力は同程度(CEFR B1~B2レベル)、どちらのクラスも全員が日本語母語話者で英語圏に長期滞在経験者は不在であった。どちらの群を介入群とするかは、無作為で決めた。評価者としては、大学で英語ライティング教授経験のある6名の日本人英語教員が調査に参加した。

c. 事例研究2の研究マテリアル

本研究の手続きは、次節 d の研究課題および分析方法に詳細を記載するが、本節では研究に使用されたマテリアル各種を紹介したい。(1) ライティング課題は、難易度の異なる英検2級(課題A)と英検準1級(課題Bおよび転移課題C)から学生の興味関心に配慮したトピック(大学生が社会人になった際に考える内容)を選んだ。(2) 評価尺度は Jacobs et al. (1981) の ESL Composition Profile 評価尺度に修正を加えたものを採用し、(3) 注釈タグ付けツールは、操作性、機能性の観点

から、分析と同時にパラグラフの文と文の繋がりを樹状図で自動的に表示できる Tool for Interactive Argument Annotation (TIARA) (Putra et al., 2020) を採用した。

d. 事例研究2の研究課題および分析方法

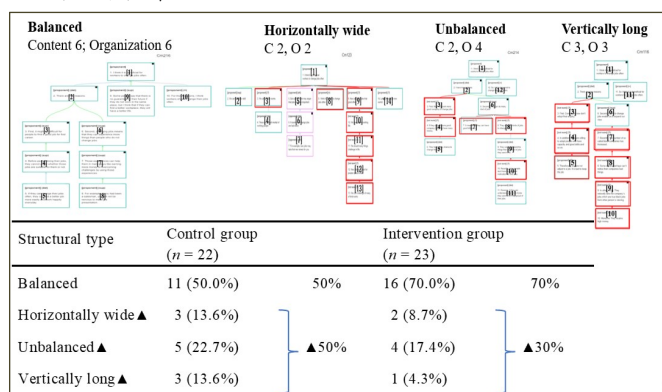
量的研究では、Study 1-1 (Claim 4) で多相ラッシュモデルを用い、本研究のために既存の尺度に修正を施した評価尺度について、評価者間の採点の信頼性および学生の能力やタスク難易度の観点での尺度使用の適切さを分析し、評価尺度がうまく機能しているかを検証した。Study 1-2 (Claim 3) では、総合的な得点の特徴の分析および解釈のため記述統計に基づく平行座標プロットで分析した。Study 1-3 (Claim 1) では、図示化フィードバックの効果の分析のため、介入群と対照群(2水準)での従属変数4観点(内容、構成、言語使用、語彙)における採点結果比較を初稿、初稿修正、転移タスクの反復測定(3水準)で行い、分析には2要因混合型多変量分散分析の結果を用いた。さらに、Study 2 初稿修正タスクでの作業所要時間の比較も行った。Study 3 の質的研究においては、2種類の質的データを以下の方法で分析した。まず、TIARA を用いて学習者のテキストを注釈タグ付けした上で文同士の関係を樹状図に図示化し、タグ付けから一貫性の破調 (coherence anomaly) を特定し、構成の形状の特徴を分析した。また、両群の修正行動とフィードバックの受容の違いを確認するため、自由記述での質問紙から得られた回答をテーマ分析の手法で探索的に分析、検証した。分析の途中および最終段階で、質的・量的研究分析結果を統合して課題の検証を行った (Matsumura, 2023)。

4.3 事例研究2の結果(一部)

本稿では、紙面の都合上、Study 3 で樹状図による質的分析と得点での量的分析の統合の結果を表わすジョイントディスプレイの結果(図5)のみ報告する。樹状図は、均衡型 (Balanced)、横広型 (Horizontally wide)、

図 5

4 タイプの樹状図と得点の関係 (介入群 vs. 対照群) を示すジョイントディスプレイ



注 Matsumura (2023) 掲載の図を本稿用に編集。

不均衡型 (Unbalanced) および縦長型 (Vertically long) の 4 タイプに概ね分類され、その中で均衡型の Organization と Content の観点での得点が高い傾向 (図 5 の例では 6-6) にあり、他の 3 つは得点が低い傾向 (図 5 の例では 2-2; 2-4; 3-3) にあるが、介入群のパラグラフの方が対照群に比べ均衡型の樹状図の割合が高いことが量と質の結果の統合から確認された。

4.4 事例研究 2 の考察・研究の限界と今後の展望

研究事例の限界点については、本稿筆者が指導者として担当する授業に指導者兼研究者 (teacher-researcher) として行った実証研究であることから、学習者の理解という点では利点であるが、バイアスがかからない工夫をすることの難しさが挙げられる。今一つの限界点として、質的研究分析が、学生を全体として捉えた分析に留まっている点であろう。今後は、より個別で深いレベルでの量と質のデータの統合と解釈ができれば、混合型研究法を活用する価値が高まると考える。

4.5 事例研究 2 の結論

混合型多変量分散分析 (Mixed-between-within MANOVA) では、図示化の形成フィードバックの効果は、介入群と対照群との間で統計的に有意な差は確認されなかった。しかし、質と量を統合する MMR を用いたことで、図示化による形式的フィードバックが初稿修正文や転移タスクのライティングに与えるさまざまな肯定的な影響が明らかになった。介入群において、書き直し作業における時間対効果が上がることで、得点には表れにくい一貫性の破調の出現頻度が低く、テキスト単位での評価が高いこと、図示化されたテキストの構造の形状のバランスがとれている点、更には、学生の英作文における「構成」や「一貫性」への理解や気づきといった情意面においての深まりがあることなどから、一定程度の介入効果が期待できるという結論に至った。

5. 本稿の考察

本稿は、混合型研究法 (MMR) における妥当性・正当性の検証項目と言語テスト評価の妥当性検証の枠組みの一つである Bachman and Palmer (2010) の Assessment Use Argument (AUA) の構成概念を、方法論的整合性という観点で照合した上で、言語テストの文脈で AUA を、混合型研究法を用いた実証研究に適用することの有効性について、順序の妥当性、パラダイムの・哲学的妥当性、ならびに統合の妥当性 (Onwuegbuzie & Johnson, 2006) の 3 つの視点から論じた。本稿筆者が過去に実施したスピーキングおよびライティングのパフォーマンステストの実証研究 2 事例では、AUA の枠組みで設定した研究課題を、適切な MMR デザインに沿って実施することで、質的および量的データを効果的に統合し、ジョイントディスプレイによるメタ推論により、解釈がより充実したものとなり、2 つの対象事例の評価をより包括かつ複眼的に行うことができた。従って、AUA の

枠組みの MMR での適用が、先に示した 3 つの視点の妥当性を確保しつつ、研究の実施、結果の分析・解釈に有効であることが推認できた。

6. 本稿の限界と今後の展望

本稿では、混合型研究法において AUA の枠組みを適用する有効性について論じてきたが、特定の枠組みで妥当性の検証が常に首尾よく進むわけではないことは言うまでもない。この点に関して澤木 (2011) が、「一つの推論に関して最初に予定した研究が完了したとしても (中略) 一度渡った「橋」についても繰り返し再確認しながら妥当性研究を進める」(p. 58) 重要性を説いているように、妥当性検証に関しては、常に柔軟かつ自己批判的に適宜修正を加える姿勢を持つ必要がある。

7. 本稿の結論

本稿で紹介した 2 事例において、AUA の枠組みに沿って設定した研究課題を、量的および質的データ収集や分析、統合、解釈に至る行程を適切な混合研究デザインで構築することにより、妥当性・正当性の質を確保しつつ研究を実施することができた。Moeller (2016) は、「教室での本格的なタスクベースのコミュニケーションの評価には多数の変数があり複雑であるため、1 つの研究手法では言語スキルの複雑さを完全に捉えることはできない」(p. 8, 本稿筆者訳) と指摘している。ここに、教室における言語評価の難しさと共に、混合型研究アプローチの価値がある。指導とテストはしばしば分離不可能で、基本的にコミュニケーション型言語能力評価は混合型研究アプローチと親和性が高い。

質的・量的分析での段階的で複雑な過程を伴う混合型研究の結果の質の確保には、方法論的妥当性 (順序の妥当性・統合の妥当性) の視点が重要で、その場合の包括的研究デザインの構築に Assessment Use Argument の適用は有効な手段の一つだと結論づけられる。

注

¹ 本研究の一部には、謝辞にある研究助成での著作物ならびに筆者の未刊行の博士論文を本研究の目的である方法論的視点に焦点を絞って再構成したものが含まれる。

² 方法論的整合性は、具体的には、十分な文献レビュー、適切な研究アプローチの選択、適切な方法や手続きの使用、方法や結果に基づく結論によって検証される。

³ ジョンソン・フェターズ (2016, pp. 10–11) では、引用元となる論文の著者の一人であるジョンソンは、正当性 (legitimation) は全て妥当性 (validity) に置き換え、また、いくつかの用語については変更・追加がされているが、本稿の英語用語表記は原文となる Onwuegbuzie and Johnson (2006) に従った。

⁴ 研究デザインとは、①研究へのアプローチ、②データ収集戦略、③データ分析戦略の 3 つを組み合わせたものである (Levitt et al., 2018, p. 32)。

⁵ この変換に関しては、2.2 での Onwuegbuzie and Johnson (2006) の 9 つの妥当性の中の (5) 変換の妥当性が問題となる。

⁶ 1970 年代から 1990 年代にポスト実証主義を背景とする量的

研究者と構成主義を背景とする質的研究者との間で展開された人間研究の方法論の優劣をめぐる論争。

⁷ 69名全員がペア型スピーキングテストに参加したため量的データ分析1および2の対象としたが、クラスター分析の対象は、質的データの紐づけのため、インタビューを実施できた1クラス38名とした。さらに質的データ（インタビューおよびアンケート）分析対象者は、38名の内、全ての質問項目を完了できた34名とした。

⁸ トピックは①オンラインショッピングと店舗ショッピングのどちらが良いか、②高校生のアルバイトの是非。いずれも授業内で学習したトピックである。

謝辞

第6回 JAAL in JACET 学術交流学会にて有益なコメントをくださった先生方および査読者の先生方に、心より謝意を表す。本研究の一部は、JSPS 科研費 JP20K00750 の助成および日本英語検定協会の第31回研究助成を受けたものである。また、本研究の一部は、筆者が早稲田大学大学院教育学研究科に提出した未刊行の博士論文を、本研究の目的である方法論的視点に焦点を絞って再構成したものが含まれる。

参考文献

- Bachman, L. F., & Palmer, A. (2010). *Language assessment in practice*. Oxford University Press.
- Creswell, J. W. (2015). *A concise introduction to mixed methods research*. Sage.
- Creswell, J. W., & Plano Clark, V. L. (2018). *Designing and conducting mixed methods research* (3rd ed.). Sage.
- フェッターズ, マイク. (2021). 「トレジャーハント (宝探し) で学ぶ混合研究法」. フェッターズ, マイク・抱井尚子. (編).『混合研究法の手引き—トレジャーハントで学ぶ研究デザインから論文の書き方まで』(pp.9-14). 遠見書房.
- Fetters, M. D., & Freshwater, D. (2015). The 1 + 1 = 3 integration challenge. *Journal of Mixed Methods Research*, 9(2), 115–117. <https://doi.org/10.1177/1558689815581222>
- Flick, U. (2007). *Managing quality in qualitative research (The Sage qualitative research kit 8)*. Sage. (上淵寿. (訳). (2017).『質的研究の「質」管理 (SAGE 質的研究キット 8)』. 新曜社.)
- Galaczi, E. (2014). Interactional competence across proficiency levels: How do learners manage interaction in paired speaking test? *Applied Linguistics*, 35, 553–574. <https://doi.org/10.1093/applin/amt017>
- Gibbs, G. R. (2007). *Analyzing qualitative data (The Sage qualitative research kit 6)*. Sage. (砂川史子・一柳智紀・一柳梢. (訳). (2017).『質的データの分析 (SAGE 質的研究キット 6)』. 新曜社.)
- Jacobs, H., Zinkgraf, S., Wormuth, E., Hartfiel, V., & Hughey, J. (1981). *Testing ESL composition: A practical approach*. Newbury House.
- ジョンソン, R. B.・フェッターズ, M. D. (2016).「混合研究法入門」. 抱井尚子・成田慶一. (編).『混合研究法への誘い』(pp. 5–13). 遠見書房.
- 抱井尚子. (2015).『混合研究法入門—質と量による統合のアーチ』. 医学書院.
- 抱井尚子. (2016).「混合研究法『古くて新しい』研究アプローチ」. 抱井尚子・成田慶一. (編).『混合研究法への誘い』(pp. 2–4). 遠見書房.
- 抱井尚子. (2021).「介入デザイン論文のトレジャーハント」. フ

ェッターズ, マイク・抱井尚子. (編).『混合研究法の手引き—トレジャーハントで学ぶ研究デザインから論文の書き方まで』(pp. 67–79). 遠見書房.

亀井智子. (2019, 1月7日).「混合型研究の評価—データ統合の質に焦点をあてて. Part 2. データの統合方法とその評価」混合研究法 Webinar シリーズ.

<https://www.youtube.com/watch?v=78TzkyNW7Q0>

Kluge, S. (2001). Strategien zur Integration qualitativer und quantitativer Erhebungs- und Auswertungsverfahren. Ein methodischer und methodologischer Bericht aus dem Sonderforschungsbereich 186 ‘Statuspassagen und Risikolagen im Lebensverlauf’. In S. Kluge & U. Kelle (Eds.), *Methodeninnovation in der Lebenslaufforschung. Integration qualitativer und quantitativer Verfahren in der Lebenslauf- und Biographieforschung* (pp. 37–88). Juventa.

Knoch, U. (2007). *Diagnostic writing assessment: The development and validation of rating scale*. [Doctoral dissertation, University of Auckland]. Researchgate.

<https://www.researchgate.net/publication/37986429>

小泉利恵. (2016年12月23日).「Facetsを使った多相ラッシュ分析—パフォーマンステストの妥当性検証に向けて」言語教育とデータ分析に関する連続ワークショップ, 早稲田大学.

Levitt, H. M., Bamberg, M., Creswell, J. W., Frost, D. M., Josselson, R., & Suárez-Orozco, C. (2018). Journal article reporting standards for qualitative primary, qualitative meta-analytic, and mixed methods research in psychology: The APA publications and communications board task force report. *American Psychologist*, 73(1), 26–46.

<https://doi.org/10.1037/amp0000151>

Matsumura, K. (2023). *A mixed methods study on formative assessment of EFL writing of college students in Japan: Focusing on the effectiveness of schematized teacher feedback on coherence*. [Unpublished doctoral dissertation]. Waseda University.

松村香奈・守屋亮. (2019a).「教室における Paired oral test の診断的評価および学習者の授業に関する調査—混合研究法を用いて」.『第31回「英検」研究助成報告書」, 212–234. https://www.eiken.or.jp/center_for_research/pdf/bulletin/vol31/vol_31_p212-p234.pdf

松村香奈・守屋亮. (2019b).「英検 Web ショートプレゼンテーション (研究助成/委託研究)」

https://www.eiken.or.jp/center_for_research/movies/

Moeller, A. J. (2016). The confluence of language assessment and mixed methods. In A. Moeller, J. W. Creswell, & N. Saville (Eds.), *Second language assessment and mixed methods research* (pp. 3–16). Cambridge University Press.

Onwuegbuzie, A. J., & Johnson, R. B. (2006). The validity issue in mixed research. *Research in the Schools*, 13(1), 48–63. <https://ocd.lcwu.edu.pk/cfiles/Gender%20&%20Development%20Studies/GDS-502/Thevalidityissueinmixedresearch.pdf#page=55>

Putra, J. W. G., Teufel, S., Matsumura, K., & Tokunaga, T. (2020). TIARA: A tool for annotating discourse relations and sentence reordering. *Proceedings of the 12th International Conference on Language Resources and Evaluation (LREC)*, 6912–6920. <https://aclanthology.org/2020.lrec-1.854>

澤木泰代. (2011).「大規模言語テストの妥当性・有用性検討に関する近年の動向」.『言語教育研究』, 2, 54–63.

Toulmin, S. E. (2003). *The use of argument*. Cambridge University Press.

van Dijk, T. A. (1977). *Text and context: Explorations in the semantics and pragmatics of discourse*. Addison-Wesley Longman.

必修英語授業における ChatGPT 活用の試み

大賀まゆみ* 豊島知穂** 山中司***

*立命館大学
moga@fc.ritsumeai.ac.jp
**立命館大学
ctr22145@fc.ritsumeai.ac.jp
***立命館大学
yaman@fc.ritsumeai.ac.jp

要旨

生成系 AI が飛躍的に発展したことにより、教育現場における ChatGPT や機械翻訳の使用については、賛否両論含めて様々な議論がなされており、大学によって対応も様々である。本研究においては、大学の正課英語授業に ChatGPT を取り入れ、教員が ChatGPT 利用を許可した場合、学生はどのように課題に取り組むのか調査を行った。プロジェクト発信型英語プログラムにおいて、ChatGPT を使ってプロジェクトにタイトルをつけるというアクティビティを実施したところ、多くの学生が最終的には ChatGPT の力を借りずに自力でタイトルをつけるという結果となった。このことから、学生が自身の課題に対する愛着や誇りがあれば容易に AI に頼ることはないということが 1 つの事例として明らかとなった。今後の課題として、様々な場面に応じた利用方法を比較・分析し、より有効な使い方を模索していく必要がある。

キーワード：ChatGPT、生成系 AI、英語教育、プロジェクト発信型英語プログラム

1. はじめに

生成系 AI が飛躍的に発展したことにより、教育現場における ChatGPT や機械翻訳 (Machine Translation: 以下, MT) の使用については、賛否両論含めて様々な議論がなされており、大学によってその対応も様々である。立命館大学の生命科学部・薬学部プロジェクト発信型英語プログラム (Project-based English Program: PEP) では、MT をはじめとした生成系 AI を積極的に活用し、英語教育の質の向上に取り組んできた。安易に精度の高い MT が使えるようになったことから、多くの語学学習者が MT を利用していることは明らかであり (Jolley & Maimone, 2015; 小田, 2019)、学習者の生成系 AI 使用は避けて通れないように思われる。国内外においても、MT を授業で利用することを前提としたカリキュラムポリシーや倫理的な指導を導入する動きがある (Bowker, 2020)。学習者個別に合った問題の解決法やヒントを AI が瞬時に与えてくれることから、機械が個別教師のような役割を果たし、うまく使いこなせば効果的な学習に繋げることが期待される。加えて、国内においても複数の研究において、MT の使用は語学学習の動機付けを奪うどころか、適切に使えば学習意欲を高めることができると報告されている (坂本, 2021; 下山, 2023)。MT を正課英語授業に導入した取り組みについては、近藤他 (2023) によって報告されている。学生が MT を利用して英語の課題に取り組んだ場合、MT で作成した英文と自作英文の比較や、ポストエディットの

作業を行うなど、学生が能動的に学習した場合においてのみ、その作成した課題が学生にとって愛着のあるものになるという。ICT を活用した教育が今後も増えていくことが予測されるが、教育現場に導入する上での研究は現段階で十分とはいえず、学生が生成系 AI の利用を許可された状況でどのような行動をとるのか、実態や意識を把握しておくことは重要である。そこで、本研究においては必修英語授業において教員が ChatGPT の利用を許可した場合、学生はどのように活用し、実際に使用した結果どのように感じるのか調査を行った。

2. 調査方法

本調査は、理系分野を専攻している大学 3 年生春学期の必修英語科目、プロジェクト発信型英語プログラム (PEP) の JP1 (Junior Project 1) 授業内で実施した。この授業では、学生が教員に与えられた課題ではなく、自分で興味・関心のあるテーマを追い求め、共通の興味を持つ学生がグループで集まり、プロジェクトの立案から最終発表まで全て学生が決定する。専門分野の教員のアドバイスを受けながらグループで話し合い、数週間から数ヶ月に渡り、紆余曲折を経て最終的に自分たちのプロジェクトを完成させ、最終発表は英語でポスター発表を行う。今回の調査ではその最終発表のタイトルを考える際に、教員が ChatGPT を許可した場合、学生がどのような反応をするのか調査した。

参加者は 69 名で、15 グループ (各グループ 4 名から

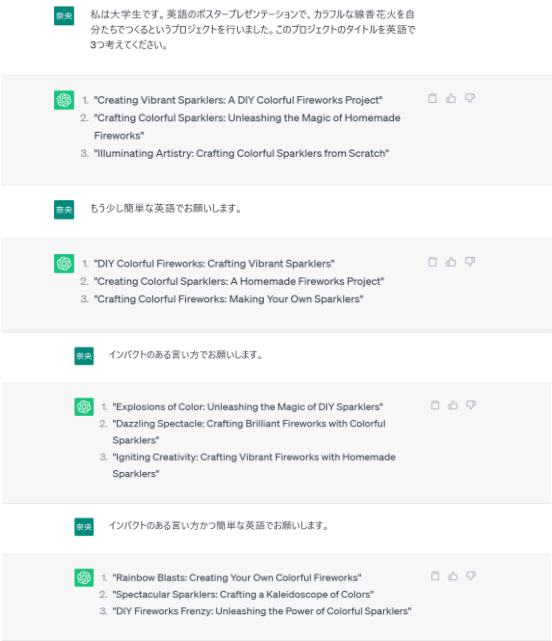
5 名)に分かれてそれぞれ課題に取り組んだ。本プログラム受講生の学年全体の TOEIC 平均点は 550 点程度である。はじめに、ChatGPT を使わずに各グループが自力でプロジェクトのタイトルをつけた上で、ChatGPT を利用してタイトルを考えた。ChatGPT に人間が与える指示をプロンプト (prompt) と呼ぶが、学生には以下のような 3 つのプロンプトを出すように事前に指示している。実際に学生が ChatGPT に出した指示例は図 1 の通りである。

ChatGPT を使う際の学生への指導

ChatGPT のタイトルをそのまま使うと退屈なタイトルになるので、以下のようなプロンプトを出すこと。

- (1) リサーチの内容 (目的・方法・結果) を説明し、タイトルをつけて欲しい
- (2) リクエストをする (大学生がわかるように・アカデミックに・Catchy なもの 3 つ考えてほしい・サブタイトルも etc.)
- (3) さらにリクエストをする (もっとオリジナリティがあるもの、注目を惹くもの etc.)

図 1 学生が実際に ChatGPT に出したプロンプト例



以上のような手順を通して ChatGPT より得たタイトルと、自力で作成したタイトルを各自比較しながらグループでの話し合いを行った。ChatGPT を使いながらまずは暫定的なタイトルを決め、その後話し合いを重ねて、最終発表に向けたポスター発表のためのタイトルを決定した。アクティビティ後、参加者のうち 6 名に ChatGPT とアクティビティに関するインタビューも併せて実施した。

3. 結果

3.1 ChatGPT を利用したグループ

表 1 において各グループが「(1) (ChatGPT を使わず) 自力でつけたタイトル」、「(2) ChatGPT 授業後のタイトル」、「(3) 最終的に採用したタイトル」を示す。それぞれ太字で記載されている部分が ChatGPT の提案を採用した箇所である。

表 1 学生が作成したタイトルの比較 (緑：一部採用, 黄：採用)

グループ	(1) ChatGPT を使わず 自力でつけたタイトル	(2) ChatGPT 授業後のタイトル	(3) 最終的に採用したタイトル
1	What is earthquake-resistant structures? -To prepare for future great earthquakes -	Beat with a hammer to beat a disaster!! -Design the dream building preparing for the earthquake-	Beat with a hammer to beat a disaster!! -Design the dream building preparing for the earthquake-
2	It is Time to Eat Garbage!! -Proposing solution by making edible containers-	It is Time to Eat Garbage!! -Suggesting for edible containers to reduce garbage-	It is Time to Eat Garbage!! -Suggesting for edible containers to reduce garbage-
3	Oh, Come on!(OK/COME) -How to eat rice deliciously-	Oh, Come on!(OK/COME) -Exploring the Hidden Powers of Grains Through Amazing Taste Experiments-	Oh, Come on!(OK/COME) -Exploring the Hidden Powers of Grains Through Amazing Taste Experiments-
4	Big Bag Makes You Feel Bad!! -Consider the method of packing your safety and comfort-	Rainbow Sparkles -Simulate the Influenced School Bags on Young Students and Consider How to Solve This Problem (this may be changed) -	School Bag Is Too Heavy -Simulation of dice and considering of centre of gravity-
5	Applications of Vitamin C and Vitamin D Derivatives -The Secret to Keeping Young Skin-	THE SECRET TO KEEPING YOUR SKIN YOUNG -Applications of Vitamin C and Vitamin D Derivatives-	THE SECRET TO KEEPING YOUR SKIN YOUNG -Applications of Vitamin C and Vitamin D Derivatives-
6	To US!! No, To BKC!! -To improve congestion on the bus queue to BKC-	To US!! No, To BKC!! -To improve congestion on the bus line to BKC-	To US!! No, To BKC!! -To Improve Congestion on the Waiting Line for Bus to BKC-
7	Let's make coffee with vegetables! -Preparing for a future where global warming will make coffee undrinkable-	Let's make coffee with vegetables! -Preparing for a coffeeless world-	Let's make coffee with vegetables! -Preparing for a coffeeless world-
8	Protect apples from oxidation!! -Using the best method-	The Apple Odyssey -How to protect apples from oxidation-	The Apple Odyssey -How to protect apples from oxidation-
9	The Way of Speaking and Listening	Impressive Communication -Unveiling the Secrets Through Videos Comparison and Survey-	Let's Take Better Communication !!
10	Rainbow Sparkles -Making sparklers by ourselves using chemical reactions-	Rainbow Sparklers -Making sparklers using chemical reactions-	Rainbow Sparklers -Making sparklers using chemical reactions-
11	I have a sweet tooth!	The Best Food That can Replace Sweets -Adverse Effects of Sugar Intake-	The Best Food That can Replace Sweets -Adverse Effects of Sugar Intake-
12	Curb Alcohol Absorption -Foods that reduce alcohol absorption-	FOODS VS. ALCOHOL -How Certain Foods Can Help Delay the Effects of Alcohol-	FOODS VS. ALCOHOL -How Certain Foods Can Help Delay the Effects of Alcohol-
13	Beyond the Preservation Border -The natural way to keep vegetables fresh for a long time -	Beyond the Preservation Border -The natural way to keep vegetables fresh for a long time-	Beyond the Preservation Border -The natural way to keep vegetables fresh for a long time-
14	A Day to Grow Intestinal Bacteria	Get Good Gut	Get Good Gut -How to improve gut environments-

結果として、「(2) ChatGPT 授業後のタイトル」において、ChatGPT が回答したタイトルの一部を利用したグループは 14 グループのうちのうち 7 グループ、また「(3) 最終的に採用したタイトル」で ChatGPT の回答を利用したグループは 5 グループのみであった。

3.2 ChatGPT 導入後のインタビュー調査

参加者のうち 6 名にインタビューをした結果、多くの学生が ChatGPT を最適な答えを導き出すツールとしては認識しておらず、ポスター発表を見るオーディエンスの立場に立って、ChatGPT のつけたタイトルを客観的に評価し、取捨選択した上でタイトルを決定していた。

一方、インタビュー回答者全員が ChatGPT の利用やその回答を採用することについて、「結果がより良いものになれば全く違和感はない」と回答し、ChatGPT の利用そのものについては抵抗がないことが分かった。それに関わらず、ChatGPT が提案したタイトルを使わなかった理由についてインタビューの詳細を確認すると、今回はプレゼンのタイトルとしてオーディエンスの興味をいかに引き、わかりやすく伝えるかということを重視したために、採用をやめたグループが多かった。採用を決めたグループについても、ポスター発表のタイトルとして、ChatGPT の提案を確認した上で、状況に応じて使えるものは利用しながら、自分たちが考えたタイトルと組み合わせつつ利用している様子が窺える。学生のインタビュー結果詳細は表 2 の通りである。

表 2

学生が作成したタイトルの比較（緑：一部採用，黄：採用）

回答者	自力で作成したタイトル	ChatGPT授業後 (暫定的なタイトルへの) ChatGPT利用	最終発表における ChatGPT利用	ChatGPT利用後の感想	最終発表のタイトル	ChatGPT感想	ChatGPT使用経験	ChatGPTのタイトルを採用する違和感
A	重要なポイントを簡単にまとめたというタイトルの。	不採用	不採用	みんなあんまりしっくりこなかった。ChatGPTに聞いた内容だけを見ると、説明だけ、説明調、説明ばっかりで、長すぎたり単語が難しい。そもそも伝える側がわからない。なんか無機質な感じ。	伝わりやすさ的にちよっとわからないからそこはちよっと、ですがまとまってるかな。	ゼロから1を生み出そうとして使うのは弱い感じがする。自分がわからないことをChatGPTに書かせよう、みたいな感じだとあんまり使い物にならないと思うが、自分がわかっていることをまとめようみたいな感じだと使えると思う。	なし。信じられないから使わない。自分で考えないと怖い。	内容としてばっちりあってるんだったらもうそれに越したことはない。
B	みんなキャッチーなタイトルだと思って気に入っていた。	一部採用	一部採用	サブタイトルだけ変えようって話していた。何か科学的で出来れば実験 (experiment) という単語が入っているものがよかった。一番このサブタイトルが面白くてScientificでよかったと思った。	選んだサブタイトルは自分たちもオーディエンスも知ってる単語も多いしわかりやすいと思った。	信頼しすぎるのは良くない。あくまでも考える上で何かしらのアイデアをもらうという感覚で使うなら良い。	聞いてみて、その後文獻で確認したりして、ああ、あつてるんだとかいう使い方はしている。	使うことに抵抗はない。
C	メインのタイトルは気に入っていた。サブタイトルは長いしポスターのレイアウト考えて変えようとは思っていた。	一部採用	一部採用	(サブタイトルを) 先ず短くして内容もちよっと変えようと思った。(Cofeelessという単語が) 聞いたことない単語やけど意味わからな。みたいな。そう言う言葉があるんだ、何これ、みたいな。出てきた中でも注目を集めた。タイトルも聞いたが学術的で、しっかりした真面目なやつだった。	その後も結構考えて見たりもしたがピンとくる、これやなって言うのがなかった。	返ってくる速さがすごい。自分が打ち込んだやつがどこで管理されているんだろうというのが怖い。ひねりがなく、普通の回答が返ってきて単語も難しい場合がある。オリジナリティが必要なのはやっぱり人間。どんな頭が悪くなっているかわからない。便利やとは思いうけど使い方が考えないとなあ。	これがきっかけでアカウント作った。	良ければ (タイトルも) 変えていた。
D	満足はしてしたが、少し物足りない、シンプル過ぎ。	採用	不採用	Week 10 (自力作成) のタイトルだと話し方と聞き方の方法だけだったから伝わらないかな、と考える具体的な内容 (動画作成) に触れたかった。	少し長すぎるかなと言う話になり、シンプルにわかりやすくしたい。わかりやすく伝わりやすいポジティブな感じの方がオーディエンスに響くかなあ。オーディエンス目線で考えたら良いタイトル。	ChatGPTはアドバイスとか的確に教えてくれるが、目的によっては自分で判断しなければいけない。用途によって使い分けないと、そのまま使ってしまうと単語が難しすぎたりとか、論点がずれてしまったりとかもあるんで、場面に応じて使い分けないといけない。	課題で一回使用	良ければいい。
E	教員に何か指摘されたのは覚えていて、良いのがあれば変えたら良いかなと思っていました。	採用	採用	Food VS. Alcohol というのが短くて他にタイトルにしてるところがないやそうなので、ってインパクトに惹かれてつけた。	みんな気に入った。	便利なので自分で考えることが少なくなってしまうそうだが、効率を図るなら使用してもいいかな。良いものになるんだったら採用せざるを得ない。	今回初めての使用	全くない。
F	これでいいこうってなっていたが、命令文なのがよくないみたい。			教員がChatGPTの使い方をデモで行った際にでてきた中の一。みんなでこれいねって話した。サブタイトルは教員のWeek 10のポスターのコメントを採用 (ChatGPT) した。自分たちで授業中にやったものは単語が難しすぎた。これが一番簡単でわかりやすかった。	タイトルは結構気に入っててみんな一瞬でこれがいいってなったのでそのまま。Week 10より良くなったと思う。	数学系、計算とかは弱い。間違っている。科学的な事もあり得意じゃない。コピペはせず、自分の言葉で書き直す。文献と照らし合わせるなどして、疑う姿勢、怪しいと思う姿勢があれば考えなくなると思うことないと思う。	GoogleのBard (科学系に強い) と組み合わせでめっちゃ使っている。英語は未来翻訳と照らし合わせて使っている。1時間かかっていたところが5分くらいでできてとても便利。効率的。	より良くなればそれでいい。

注: 回答者 E と F については、同じグループの学生である。
1 人目の記憶が曖昧であったため、2 人目にもインタビューを実施した。

4. 考察

ChatGPT や MT を語学教育に導入する上で懸念される一般的な理由として、「思考力の低下」や「語学力の低下」などが挙げられ、学生が全ての意思決定を機械に任せてしまうのではないかと危惧されているが、今回のアクティビティに関して、学生は自分の意思で ChatGPT の回答を選択しない事例が多く見られた。インタビューの回答を確認すると、以下のような意見が得られた。

- ・ (ChatGPT を) 信頼しすぎるのは良くない。あくまでも考える上で、何かしらのアイデアをもらうという感覚で使うなら良い。
- ・ ChatGPT はアドバイス等を的確に教えてくれるが、目的によっては自分で判断しなければいけない。用途によって使い分けないと、そのまま使ってしまうと単語が難しすぎたりとか、論点がずれてしまったりすることもあるので、場面に応じて使い分けないといけない。
- ・ 数学系、計算とかは弱い。間違っている。科学的な事もあり得意じゃない。コピペはせず、自分の言葉で書き直す、文献と照らし合わせるなどして

疑う姿勢、怪しいと思う姿勢があれば考えなくなると言うことないと思う。

以上の回答からも分かるとおり、学生自身も全面的に ChatGPT を信頼してはいけないということは十分に理解しており、アイディアを得るためのツールとして利用したことが窺われる。今回の課題においては、多くの学生が自力で作成したタイトルを採用したが ChatGPT に対して否定的な意見ばかりではなかった。アンケートの回答者全員が、機会があれば便利なツールとして取り入れていきたいと回答し、利用に対し前向きな意見も散見された。

- ・ (ChatGPT の回答が) 内容としてばっちりあっているのであれば、もうそれに越したことはない。
- ・ (ChatGPT の回答について) みんな気に入った。
- ・ (ChatGPT は) 便利なので自分で考えることが少なくなってしまうそうだが、効率を図るなら使用してもいいかな。良いものになるんだったら採用せざるを得ない。
- ・ Google の Bard (科学系に強い) と組み合わせでよく使っている。英語は未来翻訳と照らし合わせて

使っている。1 時間かかっていたところが 5 分くらいでできてとても便利。効率的。

今回の課題においては、春学期を通じて自分たちが作り上げてきたプロジェクトのタイトルとして、AI の提案をそのまま採用するのではなく、自分たちで吟味して調べ、考えた結果あえて ChatGPT の力を借りずに課題を完成させるグループが多く見られた。確かに ChatGPT で作成した英語の方が、英語としては素晴らしいものになるかもしれないが、自分たちが作成したプロジェクトに対する思い入れや、愛着、誇りがあれば、容易に生成 AI が作り出した回答に流れにくい傾向にあると言えるのではないだろうか。

5. おわりに

本論文では、大学の正課英語授業で ChatGPT の利用を許可した場合の学生の反応について観察した。大学生の MT や ChatGPT の利用については様々な議論がなされており、適切な利用について模索することが喫緊の課題であるといえる。今回の調査を通して、授業として教育がきちんと機能しているのであれば、学生は自らの意思で選択し、全てを便利なツールに投げることはしないという 1 つの事例として示すことができた。だとするのであれば、このようなツールの進歩は過度に恐れる必要はなく、むしろ新たな勉強のツールとして活用することが期待できる。学生は安易に ChatGPT や MT に流れることはなく、現在の英語教育の助けとして利用していくことができるのではないだろうか。今回の調査ではインタビューの人数が限られていたため、今後の課題として、より広範な範囲で調査を実施したい。加えて、取り組むアクティビティに関しても、創造性やオリジナリティが必要とされるものではなく、情報や知識を得るための調査型のアクティビティについて、学生がどのように ChatGPT の答えを取捨選択していくのか、比較・分析する必要がある。幅広い英語教育における導入を調査することによって、より有効な使い方を模索していきたい。

参考文献

- Bowker, L. (2020). Machine translation literacy instruction for international business students and business English instructors. *Journal of Business & Finance Librarianship*, 25(1-2), 25-43.
<https://doi.org/10.1080/08963568.2020.1794739>
- Jolley, J., & Maimone, L. (2015). Free online machine translation: Use and perceptions by Spanish students and instructors. In A. J. Moeller (Ed.), *Learn language, explore cultures, transform lives* (pp. 181-200). *The 2015 Central States Conference on the Teaching of Foreign Languages*.
- 近藤雪絵・木村修平・坂場大道・豊島知穂・中南美穂・山下美朋・山中司. (2023). 「機械翻訳の英語正課授業への大規模導入とその課題—英語発信力向上のための機械翻訳活用に向けて」. 『CIEC 春季カンファレンス論文集』, 14, 41-44.

小田登志子. (2019). 「機械翻訳と共存する外国語学習活動とは」. 『人文自然科学論集』, 145, 3-27.

坂本輝世. (2021 年 12 月 18 日). 「英語コミュニケーションへの意欲と機械翻訳」 日本英語表現学会第 50 回全国大会, オンライン.

下山幸成. (2023). 「発信型英語授業における英会話文作成課題と機械翻訳利用」. 『東洋学園大学教職課程年報』, 5, 49-63.

語用論に着目した TOEIC Part 2 応用活動 —ChatGPT を利用して—

藏屋 伸子*

*東洋大学
kuraya171@toyo.jp

要旨

本研究の目的は、大学での TOEIC を扱う 1 年次必修の英語科目において、TOEIC Part 2 応答問題を通して、学生の語用論的意識を高めることである。TOEIC の Part 2 応答問題に取り組む際、あり得る会話の一部であることを忘れ気味である。そこで、事前活動で、初見時に、正答を想定した文脈情報をイラストと説明で提示して回答させた。そして応用活動では、同じ質問/発言に対して、様々な状況・場面を想像させ、その回答/応答を考えさせた。数人の案をクラス全体でシェアした後、GPT-3.5 から得た、英語でのあり得る回答/応答とそれぞれの文脈を紹介した。その結果、英語力や経験値の不足によって、聞いた内容を現実にはあり得る会話として認識するのが妨げられている様子も見られたが、学生がやり取りを想像して楽しんでいる様子が見られ、本活動は語用論理解に向けた第一歩として有効であると確認できた。よって、本活動継続によって語用論的運用への意識が高まり得ると言える。

キーワード：語用論、文脈、TOEIC、ChatGPT

1. はじめに

語用論を論じられる際によく用いられる Crystal (2003, p. 364) の定義では、語用論とは使用者の視点からの言語の研究であり、特に “the choices they make, the constraints they encounter in using language in social interaction and the effects their use of language has on the other participants in the act of communication” と、言語の使用者の選択、直面する制約、相手への効果が強調されている。

一方、語用論能力については、Taguchi (2019) によると、“L2 learners need a range of linguistic resources, as well as the ability to evaluate contextual information, select appropriate resources, and use them efficiently in a real-time interaction. (p. 2)”と述べ、L2 学習者が語用論能力を身に付けるためには、幅広い言語知識を持ち、文脈を評価して適切な言語を選択し、それをリアルタイムに効率的に使用できる必要があるとしている。

2. 先行研究と背景

語用論研究の典型的な分野として、Roever (2022, p. 8) は、Speech acts, Implicature, Routine formulae, Politeness, Interaction を紹介している。これらはいずれも文字通りの内容を意味していない発言についての本来の意味や機能、効果等の調査が中心であると言える。文字通りの意味ではないということは、L1 話者でも理解が難しい場合もあり、L2 話者にとってはさらに困難を極めることは想像に難くない。また、理解には L1 転移の問題も出てくる。Roever (2022) は、L2 語用論の指導方法を

具体的に紹介した後、語用論指導の重要性が叫ばれているのにもかかわらず体系的な指導が行われていないことを指摘している (p. 170)。さらに、大規模語学試験でも少し工夫すれば語用論に関する設問を問うことが可能であると主張している (p. 176)。大野 (2011) は、語用論的意識について、授業内の 1 回 5 分程度の指導でも効果が見られたと報告している。

一方、TOEIC の Part 2 は、応答問題と呼ばれており、1 人目の質問/発言を聞いて、一瞬で状況を把握し、引き続いて聞こえる 2 人目の回答/応答の 3 つの選択肢の中から、最も適切に、あるいは自然にコミュニケーションが成り立つものを選ばなければならない。TOEIC Part 2 の典型的な対策としてよく紹介されているのが、1 人目の質問の文頭の 5W1H に注目することだろう。実際には、これによって直接的に回答されることもあれば、そうでないこともある。さらに、I don't know. のような曖昧な発言が正答であることが多いと言われることもあり、質問に質問で返すこともある。これらはすべて実際にあり得るコミュニケーションの中から隣接ペア、あるいは少なくとも連続する 2 つのターンを切り出したものであると言える。日常のコミュニケーションは単純なものばかりではなく、隣接ペアに隣接ペアが挿入されたり、論理が飛躍したりということも起こる。従って、2 人目の発言は、1 人目の発言に対する文字通りの応答であるとは限らず、語用論を経験的あるいは理論的に理解していなければ、正答を見つけることが難しいことも多くなっている。このことから、TOEIC Part 2 の応答問題が隣接ペアあるいは連続する 2 つのターン

であることを思い出させることが語用論理解につながり、ひいては、語用論意識を向上させることにつながるのではないかと考えた。英語で表現される世界は、自分には関わり合いのない別世界と思っているのか、指導中、日本の常識と同じ考え方で対応できる問題でも対応できない例が散見される。他の状況を考えることによって、英語の世界と自分の世界につながりがあることを認識するようになり、さらには自分事として考えることによって、TOEIC の正答率向上と語用論的能力向上の両方が期待できるのではないかと考えた。

3. 研究の目的と課題

従って、本研究の目的は、TOEIC を扱う英語科目において、TOEIC Part 2 を使った応用活動を通じて、語用論的意識を向上させることとした。また、L1 話者にあり得る応答例とその文脈を、人間の L1 話者の代わりに ChatGPT から得ることとした。本研究での研究課題は下記の通りである。

- 1) イラスト+説明で前提情報(文脈)を提供すると、内容を想像したり、自分事として考えたりするか？
- 2) 実際に起こり得る他の状況を検討することにより、内容を想像したり、自分事として考えたりするようになるか？
- 3) ChatGPT から得た情報を L1 話者の情報として利用することが可能か？

4. 研究方法

4.1 研究対象

対象は、都内四年制大学 1 年次必修の英語 (TOEIC) の習熟度別クラスのうち、期初の TOEIC 換算スコア平均が 504.9 点 (標準偏差 8.2) のクラス (Group 1: 34 名) と同 387.4 点 (標準偏差 5.7) のクラス (Group 2: 34 名)、そしてコントロールグループとして同 473.9 点 (標準偏差 8.4) のクラス (35 名) の計 3 クラスである。なお、後期は、コントロールグループも対象のグループと同様の活動を行った。

4.2 使用教材、ツール

使用教材は、北尾他 (2017a) の『一步上を目指す TOEIC listening and reading test: Level 2 -Intermediate-』で、表 1 の通り、前期は Unit 1-6、後期は Unit 7-12 を扱った。各 Unit では、Part 2 の設問が 4 問ずつ出題されている。ここで注意が必要なのは、TOEIC で扱うテーマが社会人向けであり、前期の会議や人事はまだ想像可能な範囲と思われるものの、後期は Advertisement で広告宣伝を行う側、Office Work、Business の他、Traffic の出張精算、Finance and Banking の渡航や両替など、学生が知識や経験をあまり持ち合わせていない内容が多くなっていることである。

表 1

テキストが扱うテーマ

Unit	前期テーマ	Unit	後期テーマ
1	Eating Out	7	Advertisement
2	Travel	8	Daily Life
3	Amusement	9	Office Work
4	Meetings	10	Business
5	Personnel	11	Traffic
6	Shopping	12	Finance and Banking

注. 北尾他 (2017a) の目次より各 Unit のタイトルを筆者が表にしたものである。

研究課題 1 のイラストは、イラストダウンロードサービス『イラスト AC』(AC ワークス, 2023) を利用した。研究課題 3 に利用した ChatGPT は、GPT-3.5 (OpenAI, 2022、以下単に「GPT-3.5」) で、プロンプトは、以下の通りである。

Give me ten possible diverse answers to the question, “[copied from the textbook]” and their corresponding contexts in the form of a table.

Give me ten possible diverse responses to the statement, “[copied from the textbook]” and their corresponding contexts in the form of a table.

2 つのプロンプトは、質問かどうかなによって使い分けた。また、Unit 3 から表形式を指定するなど、初回から多少の調整を行っている。学習効果の確認には、respon (レスポンス, 2018) のアンケート機能を利用した。

4.3 指導手順と効果確認方法

前期の初めに Part 2 の問題に取り組む際の取り組み方についてのアンケートを実施した。

通常の授業は予習を前提とし、授業内に、グループで、自然なやり取りとなる正答はどれかとその根拠についてディスカッションさせた後に、正答でのやり取りに関する全体シェアと補足・解説を行う。使用したテキストの Part 2 の設問には、語用論研究で典型的に扱われるやり取りが少なかったが、対象のグループか否かにかかわらず、学生が自分事として、自身の経験と照らし合わせたり、自分がその場にいることを想像したりすることを要求した。

本研究では、事前活動として、各 Unit の予習前に、初見で、正答を想定した背景知識や前提情報をイラストと説明で提示して、回答させた。通常の授業内容の後に応用活動として、同じ質問/発言に対して、様々な状況・場面を想像させ、その回答/応答を考えさせた。当初は、英語で答えさせたり、GPT-3.5 から得た応答例から読み上げさせたりもしたが、限られた時間のため、日本語でペアあるいはグループで 1 分/30 秒間ディスカッ

ションさせ、数人の案をクラス全体でシェアした後、GPT-3.5 から得た英語でのあり得る回答/応答とそれぞれの文脈を紹介する方法に落ち着いた。

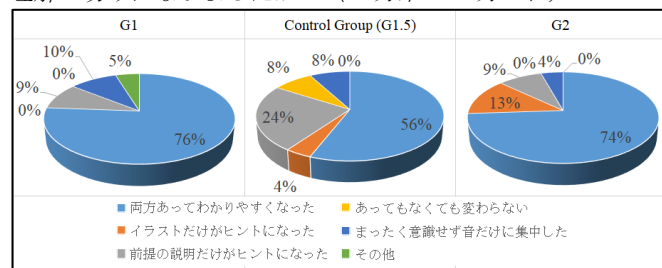
効果の確認には、前期の初めと終わり、さらに最終の Unit 12 の事前活動を行った 11 月末～12 月 1 日に Part 2 の問題への取り組み方についてのアンケートを実施した。

5. 結果と考察

5.1 事前活動：イラストと説明による文脈の提示

まず、11 月末～12 月 1 日に行った事前活動における文脈提示の効果についてのアンケート結果を図 1 に示す。内容によってイラストのみ示した回もあったが、図 1 を見ると、概ね「両方あってわかりやすくなった」と答えている。しかし、少数派をよく見ると、「まったく意識せず音だけに集中した」が 4-10%いたことがわかる。これは視覚情報と聴覚情報を同時に処理することを避けた結果と考えられる。また、コントロールグループ (Group 1.5) が他のグループ (Group 1、Group 2) と異なって「前提の説明だけがヒントになった」が多いのは、このグループでは状況に経験がないことの多かった後期でのみ事前活動を行った影響が大きいと考えられる。

図 1 前提情報を紹介した上で Part 2 に取り組んでももらいましたが、理解の助けになりましたか？ (11 月末～12 月 1 日)



注. 回答数は、G1: 21 名、G1.5: 25 名、G2: 23 名。

しかしながら、実際の正答率を見てみると、各設問でのクラス間での高低は様々で、表 2 のように前期、後期、通期ごとに正答率を平均すると、コントロールグループを設けた前期も含め、すべてで習熟度レベル順という結果となった。従って、文脈情報でわかりやすくなったとの声が多かったものの、正答率を上げるまでには至らなかったことになる。この理由は、後述する 5.3 の Part 2 の問題への取り組み方で明らかになった結果から、英語力が、状況を理解し、その場面で使われる可能性のある語彙や表現を知識として持っているレベルにまで達していなかったためと考えられる。

一方、前期に扱った Unit 1-6 で、コントロールグループの正答率が際立って低かった例があった(表 3 参照)。以下、これらの設問について、具体的に指導した例を扱う。これらの設問は、下記の通りである。これらは、テ

キストの音声¹及び教授用資料の Script (北尾他, 2017d) にのみ載っているものである。

表 2

実際の正答率 (Unit1-12 平均)

Group 期間	Group 1	Control Group (G1.5)	Group 2
通期	53.89%	49.56%	43.25%
前期	50.75%	48.06%	43.95%
後期	57.03%	51.06%	42.55%

注. 通期は Unit 1-12、前期は Unit 1-6、後期は Unit 7-12 の初見正答率平均。Control Group (Group 1.5) は、後期のみ、対象のグループと同じくイラスト・前提情報ありでの実施。

表 3

コントロールグループの正答率が際立って低かった例

Unit テーマ No. 質問概要	Group 1 正答率	Control Group 正答率	Group 2 正答率
Unit 2 Travel No.6 持ち込み手荷物	36.67%	15.15%	31.25%
Unit 5 Personnel No.4 面接前	54.17%	34.38%	48.39%
Unit 6 Shopping No.4 お使い依頼	73.08%	37.93%	41.38%

注. 各 Unit において、Part 2 の設問は No. 3-6 の 4 問である。

持ち込み手荷物 (トラック 07 1:07-20/p. 172)

6. How much carry-on baggage do you have?

- (A) Only this one.
- (B) 50 dollars.
- (C) I can carry it by myself.

面接前 (トラック 19 0:30-45/p. 178)

4. Has the candidate showed up yet?

- (A) I have to show them the plan.
- (B) They finally worked out a solution.
- (C) Yes. He's waiting for you.

お使い依頼 (トラック 23 0:30-45/p. 180)

4. Can you pick up some groceries on your way home?

- (A) They sell fresh vegetables.
- (B) What exactly do you want?
- (C) Sure. Would you like some more coffee?

事前活動では、スライド (紙面の関係上、5.2 で示す応用活動のスライド図 2-4 参照) を提示し、文脈に関する文字情報を教員が読み上げてから音声再生した。回答は、respon (レスポンス, 2018) で入力させた。

持ち込み手荷物では、例外的にイラストのみ提示したが (図 2 参照)、航空機を利用したことがある者なら馴染みのある状況だったため、正答がわかりやすかつ

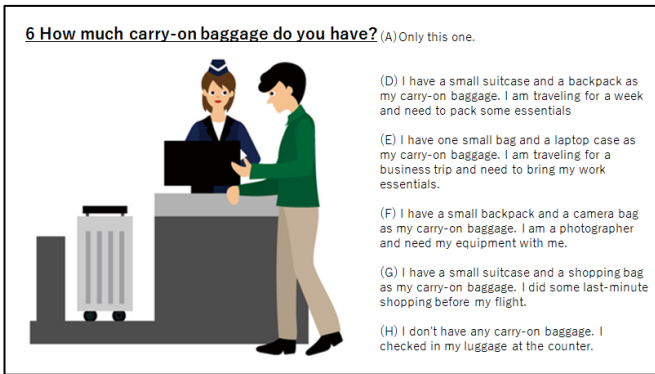
たとえられる。面接前とお使い依頼では、日本語による簡単な文脈情報とイラストを提示した(図3-4参照)。面接前には、2人のうち立っている男性に「You」という表示をした。面接については受験やアルバイト応募、お使い依頼では日常生活の中で、学生がこれまで経験した可能性の高いものであり、文脈情報によって効果的に理解を助けることとなったと考えられる。

5.2 応用活動：他の状況・場面の検討と例の紹介

ここでは、応用活動のために GPT-3.5 から得た情報、それに対して問い合わせや確認をした内容、応用活動で使用したスライドを示す。事前活動時のスライドには文脈情報とイラストのみを提示していたが、応用活動では、これらを小さくして端に寄せ、1人目の発言と2人目の正答となっている選択肢、GPT-3.5 から得たあり得る例を追加した。

図2の持ち込み手荷物では、(A)が実際の正答であり、そこに GPT-3.5 から得た 10 種類の回答の中から、選択した 5 種類²を例として合わせて紹介した。GPT-3.5 からの情報では、手荷物の量を聞かれているだけであるにもかかわらず、事細かに説明している傾向が見られたため、追加で問い合わせを行った。Is it common to explain their reasons for the amount of baggage when checking in? との問いかけには、「It's not necessary (中略) but it's not uncommon (後略)」と回答があった。「I don't have any carry-on baggage.」を含む回答について、まれなケースか問い合わせたところ、「actually quite common (前後略)」との返答を得た。しかし、GPT-3.5 からの情報だけでは、生成 AI 利用時の注意として問題となるハルシネーションの可能性を完全に排除できない。吉田 (2023a, スライド 31) は、「専門知識を持つ、1 次情報にあたる、情報の信頼性を検討することが重要」としている。従って、授業内では、これらの GPT3.5 から得た情報と合わせ、この内容が本当に英語の L1 話者に一般に当てはまるかどうかは確実ではないことを説明した。

図 2
Unit 2 No.6 持ち込み手荷物提示スライド

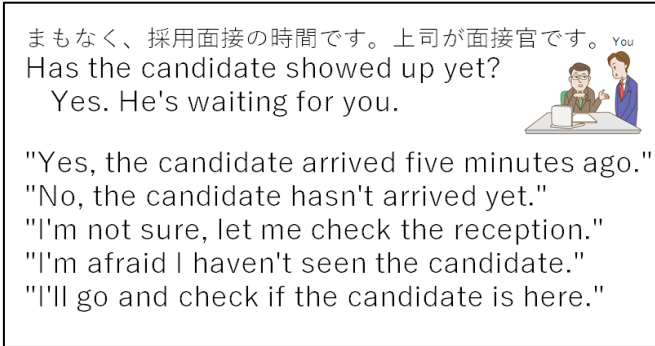


注 イラストは、ACworks (n.d.)。

図3の面接前でも、GPT-3.5 から得た 10 種類のうち

5 種類を提示した。紹介したのは、既に到着している、まだ来ていない、見ていない、確認するなどだが、他には、遅れる、来られないと連絡があった、などがあつた。この時、あるクラスの授業内で指名された学生は「骨折して来れない」と答えた。非常にまれながらあり得るケースであり、学生が様々な状況を想像し始めたことが確認できる発言であつたと言える。


図 3
Unit 5 No.4 面接前提示スライド



注 イラストは、すずしろ (n.d.)。

図4のお使い依頼は、語用論研究で最も多い依頼である。GPT-3.5 で得た情報は量が多く見づらかったため、まず2人の関係の欄を作るよう要求し (Could you add a column which shows the respective relationships between the speaker and the addressee?), さらにポライトネスレベル順に並べ替えるよう要求した (Could you rearrange the order based on politeness level?).

図 4
Unit 6 No.4 お使い依頼提示スライド

お先に電話がかかってきました。 Can you pick up some groceries on your way home? What exactly do you want?		
1. "I'm happy to help. Can you provide me with the grocery list?"	Friend/Neighbor	
2. "I'd love to assist you. Can you tell me where your preferred grocery store is?"	Friend/Relative	
3. "Sure, do you have a specific grocery list or any preferences?"	Family member	
4. "Of course, is there anything specific you need?"	Friend/Acquaintance	
5. "Sure, do you have a budget in mind for the groceries?"	Friend/Relative	
6. "I can definitely pick up groceries for you. Any specific brands or quantities?"	Colleague/Friend	
7. "I'm sorry, I won't be able to stop by the store today."	Colleague/Friend	
8. "Unfortunately, I won't be passing by any grocery stores on my way home."	Colleague/Friend	
9. "I'm sorry, I'm not available to pick up groceries today."	Colleague/Friend	
10. "I'm sorry, I'm not familiar with the grocery stores in that area."	Friend/Neighbor	

注 イラストは、きのこ (n.d.)。

しかし、図4にもあるその順番が納得できるものではなかったため、ポライトネスレベルの基準の説明を要求した。それらのプロンプトと回答の抜粋は下記の通りである。

Could you give me the specific criteria including what politeness is?

Politeness refers to the act of displaying respectful and

considerate behavior (中略) **It's important to note that politeness is subjective and can vary depending on cultural norms, individual preferences, and the specific context of the interaction.** These criteria provide a **general guideline** for understanding politeness, but the perception of politeness may differ among individuals. (太字、下線は筆者追加)

GPT-3.5 の言っている「ポライトネスが文化的規範や個人の好み、具体的な文脈によって変わる(筆者訳)」ことは、Roever (2022, pp. 2-3) のペンを借りるための 5 種類の表現例と、それらの状況による適切性に関するコメントに一部表れている他、関連する研究として “the ‘social distance’ (D),” “the relative ‘power’ (P),” “the absolute ranking (R) of impositions in the particular culture” を用いた式でフェイスを脅かす行為 (FTA) の重さを表そうとした Brown and Levinson (1987, p. 74, 76) や、文脈を 7 要因の 5 段階レベルで定義してその依頼行動のポライトネスレベルを日米と男女で比較した Kuraya (2016) などがある。また、生成 AI を「別の視点を提供」する「チームメイト」として活用する方法の報告もある(吉田, 2023b, スライド 43: Mollick & Mollick, 2023 より吉田訳)。このように元々適切な回答が曖昧である内容に関する GPT-3.5 から得た回答例は、ただ、こういう考え方もあるかもしれないと知るだけの目的であれば学習に使用可能であり、なぜそのような回答となったのかを学生に考えさせるきっかけにもなり得る。この内容は、学生たちに特にじっくり検討させたいものであったが、残念ながらこの Unit ではあまり時間を割くことができず、教員の一方的な説明に終始した。

5.3 Part 2 の問題への取り組み方

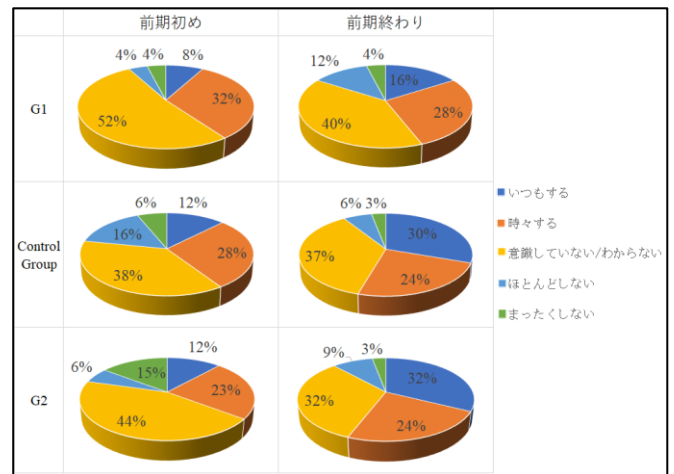
コントロールグループを設けた前期の初めと終わり、全クラスで同じ活動を行った後の 11 月末～12 月 1 日に Part 2 の問題への取り組み方についてアンケートを行った。

図 5 は、自分がそのやりとりに参加しているつもりで聞くことができるかどうかの質問である。前期のコントロールグループでは、初見回答時にイラストも背景知識や前提情報も提示しておらず、応用活動としての他の状況の検討や GPT-3.5 からの例の紹介も行っていないが、通常の設問の解説で、日本でも同様のことがあり得ることを思い起こさせるような指示や説明を行っていたため、コントロールグループも含めて「いつもする」「時々する」が増えていた。

図 6 は、内容を自分の経験と照らし合わせようとするかどうかを尋ねた結果である。こちらは、終わりに「前より意識してするようになった」単独も「前からいつもする」「前から時々する」「前から時々経験があると気付く」を加えた割合も、コントロールグループより、Group 1 (G1)、Group 2 (G2) の方が多くなっている。

図 5

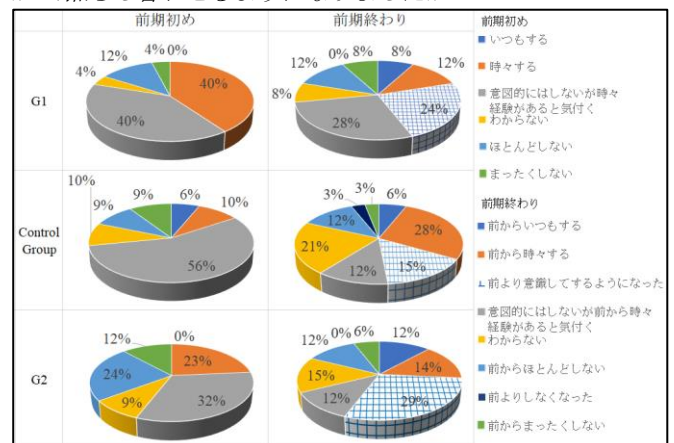
自分がそのやりとりに参加しているつもりで聞くことがありますか？



注 回答数は、G1: 25 名、CG: 33 名、G2: 34 名。

図 6

内容を自分の経験と照らし合わせようとすることがありますか？/照らし合わせるようになりましたか？



注 回答数は、G1: 25 名、CG 33 名、G2: 34 名。

さらに、自分がその状況にいると想像しない／できない理由を自由記述で聞いた。似た内容ごとにラベリングを行い、集計したものが表 4 である。

どのクラスも「経験がない」が多くなっているが、その他では、習熟度が少し上の G1 とコントロールグループは「余裕がない」、コントロールグループと G2 は「英語力不足」が多くなっている。これらの組み合わせにより、想像できるか否かには習熟度が関係していると考えられる。「意識していない」も一定数いることがわかる。一方、コントロールグループの「教わっていない」は、コントロールグループならではの回答と言える。

また、G1 では、意図的に意識することを回避していると思われる「主観を入れない」の割合が多いのが特徴的であった。これに含まれていた具体的な回答は、以下の通りである。

表 4

自分がその状況にいると想像しない／できないのはなぜだと思いますか？（前期末、自由記述）

Group 1 (回答数 15)		Control Group (回答数 20)		Group 2 (回答数 25)	
余裕がない	20.00 %	余裕がない	25.00 %	経験がない	32.00 %
経験がない	20.00 %	経験がない	20.00 %	英語力不足	20.00 %
主観を入れない	20.00 %	英語力不足	20.00 %	意識していない	12.00 %
意識していない	20.00 %	意識していない	10.00 %	余裕がない	8.00 %
		想像するよう教わっていない	10.00 %	想像力不足	8.00 %

注 前期最終授業内での自由記述をラベリングしたもの。2名以上の項目のみ。

「主観を入れない」にカウントした実際の回答例³

- ・自分の感覚と問題の感覚にはズレが有ることがあるため。
- ・受験期に、学校の先生から、「入試問題を解く際、主観を入れてはならない」と言われ、その癖が今でも残っているから。
- ・問題内容に入り込みすぎると自分の固定観念で「自分だったらこうするだろう」と解いてしまう可能性があるから俯瞰して問題を解いた方が入り込みすぎの間違いを無くせる。
- ・俯瞰しているので。あと仕事の話が多くて関連性が低いと感じたから。

このような考え方はL1 転移の問題を回避するためには有効であるが、英語というだけで自分とは別世界と考えてしまっていると感じられることの方がむしろ多い。俯瞰する意義を認めつつ、まずは現実であることを認識させる方向に導く必要があると考える。

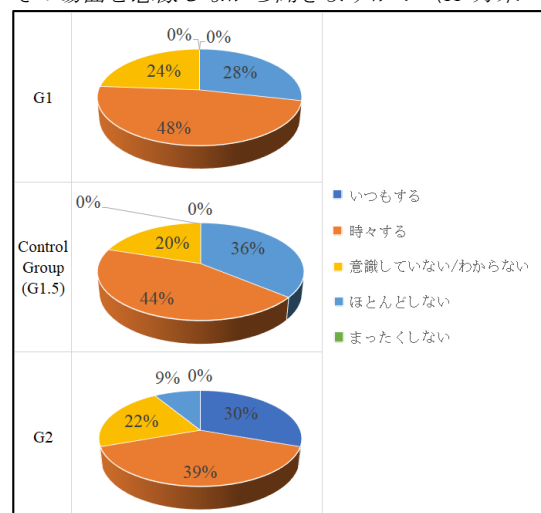
図 7 は、最終の 11 月末～12 月 1 日にその場면을想像するかを尋ねた結果である。この質問は、前期初めと終わりに聞いた、図 5 の「自分がそのやり取りに参加」では、学生には「経験がない」会社でのやり取りなどを考えづらいことを考慮して変更したものである。すると、どのクラスも 3/4 前後の学生が「いつもする」か「時々する」と答えていた。これは、本研究の活動の結果なのか、質問を変えた結果なのか特定できないため、これだけでは、本研究の活動の効果と断定はできないが、場面を想像している学生が多いことは、評価できるだろう。

最後に、応用活動についての気付きや感想を自由記

述で調査した。前述の表 4 と同じく、似た回答が同じになるようにラベリングしてまとめたものが表 5 である。

図 7

その場面を想像しながら聞きますか？（11 月末～12 月 1 日）



注 回答数は、G1 2 月 1 日名、G1.5: 25 名、G2: 23 名。

表 5

Part 2 の各質問/発言に対して、毎回他の状況を考えてもらいましたが、その活動はどうでしたか？気付いたことや感想（形容詞程度で可）を教えてください。（11 月末～12 月 1 日、自由記述）

Group 1 (回答数 15)		Control Group (Group 1.5) (回答数 17)		Group 2 (回答数 22)	
英語の学びにつながる	26.67 %	いろいろな答えがあることがわかった	29.41 %	いろいろな答えがあることがわかった	40.91 %
楽しい/おもしろかった	20.00 %	楽しい/おもしろかった	23.53 %	楽しい/おもしろかった	22.73 %
いろいろな答えがあることがわかった	20.00 %	英語の学びにつながる	17.65 %	難しかった	13.64 %
		難しかった	17.65 %	その他	9.09 %
		おもしろくなかった	11.76 %		

注 11 月末～12 月 1 日の授業内での自由記述をラベリングしたもの。2 名以上の項目のみ。

結果として、どのクラスも「いろいろな答えがあるこ

とがわかった」と答えた学生が多く、「楽しい」「おもしろかった」と好意的な印象が続いた。一方、「英語の学びにつながる」とした意見が習熟度の高いクラスであり、回答者1名で表内では提示しなかったものにも「英語で活動したい」や「不十分なので、もっと学びたい」という意見があった。

それぞれの具体的な意見の例は、以下の通りである。

- いろいろな答えがあることがわかった
（「想像の幅が広がった」も含む）
 - ・ たまになるほどなと思う回答があった。
- 楽しいおもしろかった
 - ・ 教科書の解答は真面目すぎるので、みんなからの意見は個性が出て面白い
- 英語の学びにつながる
 - ・ 想像した方が聞き取りやすくなる
 - ・ 実際に自分が陥った状況などをモデルにして考えたため取り組みやすかった。
 - ・ TOEIC でいかせるとおもった
 - ・ 英語表現を増やすいい機会になったと思う
 - ・ writing をしないといけないときに役立つかもしれないと思った
 - ・ スピーキングのテストでは役に立つ時があるかもしれないと思った。
- おもしろくなかった
 - ・ 蛇足

学びに関する意見は、TOEIC Part 2 で事前に予想できる回答のバリエーションが増える、スピーキングやライティングの勉強になる等、TOEIC に関するものだけでなく、英語のアウトプットの活動に幅広く応用できる可能性を示唆するものだった。一方、少数意見ではあるが、「おもしろくなかった」とラベリングしたものの例は、「蛇足」とのことで、TOEIC の設問に直接的に取り組む活動だけで十分という意見と思われる。

6. 結論

本研究では、TOEIC Part 2 の問題に取り組むのに際し、事前活動、応用活動を行うことにより、語用論的意識を高めることを目指した。

事前活動における、研究課題1の「イラスト+説明で前提情報（文脈）を提供すると、内容を想像したり、自分事として考えたりするか？」に関しては、わかりやすく感じたとの声が多いが、これらの情報で確実に正答率が上がるとまでは言えず、全体的には正答率は元々の英語習熟度によるという結果となった。また、音に集中しようとする人もいて、入る情報を制限して聴覚を研ぎ澄ませようとしたと考えられる。わかりやすく感じた学生も、提示した視覚情報や文脈情報によって状況は理解したが、英語自体は聞き取れていなかった、あるいは、英語の内容理解にまでは及ばなかった場合が多かったと言わざるを得ない。ただ、「わかりやすく感

じた」のは、状況を理解することによって、聞こえてくる英語を理解しようとする姿勢が整ったと言えるのではない。また音や語彙など、英語が理解できなかった原因を特定することによって学習の次のステップにつなげることが可能になる。音に集中したり文法を強く意識したりするよりも、目で見えて状況を想像しながら内容を理解しようとする方が実践に近いことから、引き続き頭の中でのシミュレーションを誘導し、語用論的受容能力の向上を目指すのが良いと考える。

応用活動における、研究課題2の「実際に起こり得る他の状況を検討することにより、内容を想像したり、自分事として考えたりするようになるか？」については、答えは1つではないことに気づき、様々な状況があることを楽しんでいるように見受けられた。応用活動における想像が設問への取り組み方法に確実に繋がっているという確認はできていないが、机上の学習における単なる問題への取り組みではなく実践的なコミュニケーションとして理解し、応用しようという、語用論的運用へ意識を向けるところまで、つまり語用論的意識への第一歩を踏み出すことはできたと言えるだろう。

研究課題3の「ChatGPT から得た情報を L1 話者の情報として利用することが可能か？」は、まず英語表現について、第6回 JAAL in JACET (2023) の基調講演において、英語の L1 話者である Tom Gally 氏が「ChatGPT の英語は素晴らしい」「指定しない限り平易な普通のアメリカ英語」と述べており、L1 話者の英語として参照することに問題はないと言える。内容については、授業準備や授業内活動を通じての教員の主観的な判断ではあるが、多くの一般的な事柄についての回答は、妥当、あるいは、あり得るとして納得できるものであり、学生にとってはそれらの英語表現に触れる良い機会となっていた。ただし、吉田 (2023a) が注意を喚起していた通り、得た情報が正しいとは限らないと常に疑い、事前に確認することが望ましい。特に、日本を出ると異なる反応があり得ることは常に念頭に置いておく必要があり、L2 語用論で問題となるような、国や文化によって規範が変わる内容については、文献等で改めて確認する必要がある。それでも、ChatGPT を利用することにより、教員 1 人が思い浮かべられる例だけを紹介するよりもはるかに多様で意義のある回答例を準備できるため、活用すべきと考える。

想定したコミュニケーションにおいて予想外の回答、オチがあると盛り上がり、学習の定着や学習への動機付けにもつながることが期待できる。本研究における特に ChatGPT を利用した TOEIC Part 2 の他の状況を検討する応用活動によって語用論的運用に向けたアクティブにより実践的な英語の学びが期待できると言える。

7. 今後の課題

今後の課題として TOEIC の設問に取り組ませる上で最も重要なことは、学生が知らない、あるいは経験のな

い内容をしっかり伝えることであろう。筆者は毎年度初回授業において、英語だけができる人よりも仕事ができる英語ができない人が英語の勉強をする方が英語でその仕事ができるようになるのは早いと伝えている。業務上のやり取りの多い TOEIC の問題に取り組むに当たり、個別の設問に取り組む前に、各 Unit のテーマに関する大まかな状況を説明し、TOEIC の設問を通して学生がどのような場面に出会う可能性があるのかを理解できるようにするのが良いだろう。また、状況が理解できたのに正答を特定することができなかった理由が何かを学生自ら考えさせるのも良いかもしれない。

GPT-3.5 やその他生成系 AI を活用する方法としては、グループワークあるいは予習の一部として学生に問い合わせをさせ、結果を発表させる活動が考えられる。その際、合わせて回答の真偽を確認させたり、適切性を議論させたりすると自ら考えるより深い学びにつながる。授業内での時間の制約を考えると、予習に組み込む方が授業中の活動により多くの時間を割くことができる。

英語科目としての本来の活動を増やす方法としては、生成系 AI で得た「実際にあり得る例」を参照しながら英語でやり取りを練習させることが望ましい。参考資料として配布する場合は著作権に注意する必要がある。

L2 語用論に関しては、L1 語用論の転移が問題になるが、現時点では TOEIC Part 2 で耳にするやり取りは自分にも経験があるかもしれないと認識できていない学生が多いため、転移を恐れるのはまだ早いと感じられる。TOEIC Part 2 の設問では、L2 語用論における典型的なやり取りが少なく、扱う機会が少ないが、英語を指導する上では、学生が L2 語用論に関する事柄にもしっかり気付きを得られるような機会を与えたい。

本研究で行った教育実践により、学生たちが楽しみながら実践的コミュニケーションに向けて準備を進められることが確認できた。限られた授業時間の中でここに挙げた全活動を行うことは難しいが今後もこれらの活動を発展させながら続けていきたいと考えている。

注

¹ 音声には、非売品の教員向け Disc (北尾他, 2017c) と、自習用にオンライン上で公開されている学習用音声・動画ストーリーミングがある (北尾他, 2017b)。

² 10 種類から何種類を紹介するかは、GPT-3.5 から得た情報による。問い合わせを繰り返して 20–30 種類から選ぶこともあった。後期は全 Unit の全問において 10 種類を紹介した。

³ 一部の回答において、言葉遣いに修正を加えている。

謝辞

本研究のイラストは、利用規約中、利用目的を問わず利用可能であり、公開制限のある「イラストデータを含む二次的著作物が主要コンテンツ」ではないものとして、『イラスト AC』で公開されているものを使用させていただいております。

参考文献

AC ワークス. (2023). 「イラスト AC」 [Clip arts].
<https://www.ac-illust.com/>

- ACworks. (n.d.). 「イラスト素材：チェックインカウンター」 [Clip art]. 『イラスト AC』 . <https://www.ac-illust.com/main/detail.php?id=2000800&word=%E3%83%81%E3%82%A7%E3%83%83%E3%82%AF%E3%82%A4%E3%83%B3%E3%82%AB%E3%82%A6%E3%83%B3%E3%82%BF%E3%83%BC>
- Brown, P., & Levinson, S. (1987). *Politeness: Some universals in language usage*. Cambridge University Press.
- Crystal, D. (2003). *A Dictionary of Linguistics and Phonetics* (5th ed). Blackwell.
- Gally, Tom. (2023, December 2). 「言語と生成 AI」 [Keynote Speech]. 第 6 回 JAAL in JACET 学術交流集会, お茶の水女子大学.
- きのこ. (n.d.). 「イラスト素材：白人ポニーテール女性-310-バースト」 [Clip art]. 『イラスト AC』 . https://www.ac-illust.com/main/detail.php?id=1533049&word=%E7%99%BD%E4%BA%BA%E3%83%9D%E3%83%8B%E3%83%BC%E3%83%86%E3%83%BC%E3%83%AB%E5%A5%B3%E6%80%A7-310-%E3%83%90%E3%82%B9%E3%83%88&data_type&from_order_history
- 北尾泰幸・西田晴美・林姿穂・ブライアン・コバート. (2017a). 『一步上を目指す TOEIC listening and reading test: Level 2 - Intermediate-』. 朝日出版社.
- 北尾泰幸・西田晴美・林姿穂・ブライアン・コバート. (2017b). 「一步上を目指す TOEIC listening and reading test: Level 2 - Intermediate-」 [Audio]. 『朝日出版社 学習用音声・動画ストーリーミング配信ページ』 . <https://text.asahipress.com/free/player/index.html?bookcode=215595>
- 北尾泰幸・西田晴美・林姿穂・ブライアン・コバート. (2017c). 『一步上を目指す TOEIC listening and reading test: Level 2 - Intermediate- Disc 1 [Audio]』 . 朝日出版社.
- 北尾泰幸・西田晴美・林姿穂・ブライアン・コバート. (2017d). 『一步上を目指す TOEIC listening and reading test: Level 2 - Intermediate- 教授用資料』 . 朝日出版社.
- Kuraya, N. (2016). *Comparison of politeness levels in American English and Japanese requests: Analysis of behavioral and linguistic data from American and Japanese film scenes* [Doctoral dissertation, Nihon University]. 日本大学リポジトリ. <https://repository.nihon-u.ac.jp/xmlui/handle/11263/782>
- 大野仁寛. (2011). 「教員を目指す学生の語用論的意識の向上に向けて」. 『JACET 関東支部学会誌』, 8, 27–36. <https://dl.ndl.go.jp/pid/10495657>
- OpenAI. (2022). GPT-3.5: Generative Pre-trained Transformer 3.5 [Online app]. <https://chat.openai.com/>
- レスボン. (2018). respon [Mobile app / Online app]. <https://respon.jp/>
- Roever, C. (2022). *Teaching and testing second language pragmatics and interaction: A practical guide*. Routledge.
- すずしろ. (n.d.). 「イラスト素材：上司と部下」 [Clip art]. 『イラスト AC』 . <https://www.ac-illust.com/main/detail.php?id=535426&word=%E4%B8%8A%E5%8F%B8%E3%81%A8%E9%83%A8%E4%B8%8B>
- Taguchi, N. (2019). Second language acquisition and pragmatics: An overview In N. Taguchi (Ed.), *The Routledge handbook of second language acquisition and pragmatics* (pp. 1–14). Routledge.
- 吉田 壘. (2023a, May 13). 「教員向け ChatGPT 講座：基礎から応用まで」 [Online event]. Zoom ウェビナー.
- 吉田 壘. (2023b, June 24). 「学習者の学びを深めるために ChatGPT を如何に取り入れるか」 第 29 回関西大学 FD フォーラム [Online event]. Zoom ウェビナー.

JACET