慶應義塾 外国語教育研究

第13号

2016

目 次

慶應義塾大学外国語教育研究センター
Keio Research Center for Foreign Language Education

JOURNAL
of
FOREIGN LANGUAGE EDUCATION

Vol. 13　2016

CONTENTS

KEIO RESEARCH CENTER FOR FOREIGN LANGUAGE EDUCATION

The Role of TBLT in Promoting Success in Adult L2 Learning

Masashi Otake

Abstract

This literature review investigates the efficacy of Task Based Language Teaching (TBLT) in achieving a high level of proficiency in adult L2 learning. First, the learner variables that trigger successful adult L2 learning are identified and classified into two categories: cognitive and non-cognitive variables. Then, the possible roles that TBLT play on these variables are examined. This paper maintains that even though TBLT cannot directly stimulate people's cognitive variables, it presumably promotes two non-cognitive variables for successful adult L2 learners: explicit learning, and extremely high motivation. In addition, TBLT temporarily surrounds the learners with L2 input, which is similar to the immersion experience that all the successful adult L2 learners have had. These findings suggest that when certain conditions are met, TBLT offers what the successful adult L2 learners have done in the process of achieving their high proficiency. This paper concludes that it is possible for adult learners to achieve a practical, high-level L2 competence through TBLT.

Introduction

With the increasing professional opportunities due to globalization, many second language (L2) learners are aiming for a high level of L2 fluency. Their pragmatic objectives are to be able to negotiate deals, persuade others, and resolve disputes to name a few. The focus of language teaching has been shifting from accuracy to fluency (Brumfit, 1984), and Communicative Language Teaching (CLT), especially the strong form of CLT, Task Based Language Teaching (TBLT), has gained the worldwide attention of researchers and analysts. (Van den Branden, Bygate, & Norris, 2009).

In TBLT, the learners discover new language structures by themselves through the use of the language in 'meaning negotiation'. They exchange knowledge and ideas with others

to solve pedagogical tasks. While engaging in these tasks, the learners' focus shifts from the meanings of words to their language structures in order to make their meaning negotiations more smooth and effective (Ellis, 2003; Long, 2014). The proponents are not only SLA researchers, such as Ellis (2003), Skehan (1998b), and Long (2014), but also teachers and educationalists, such as Prabhu (1987), Willis (1996), and Nunan (1989). These TBLT advocates suggest TBLT nurtures learners' practical high-level of L2 competence for not only everyday basic language use, but also for educational and professional language use, which many current adult learners are aiming for in their L2 learning (Long, 2005).

This literature review, therefore, aims to investigate the relationship between TBLT and the learner variables that successful adult L2 learners possess in order to predict the efficacy of TBLT in achieving a high level of proficiency in adult L2 learning. The research questions are the following: (1) What are the learner variables that trigger successful adult L2 learning? (2) Does TBLT promote the development of these variables? In order to answer these questions, the remainder of the paper will, first, review TBLT and discuss what L2 learning success is. Then, the focus will fall on the cognitive and non-cognitive variables that several empirical studies on successful adult L2 learners have discovered. Finally, the paper will discuss the possibility of TBLT and address whether TBLT promotes adult L2 learning in terms of these variables.

What is TBLT and adult L2 learning success?

Many SLA researchers believe L2 learning happens when learners are richly exposed to the target language and engage in meaningful interaction. TBLT approaches create such an environment and elicit learners' implicit incidental L2 learning. The tasks in TBLT provide opportunities for learners to engage in meaningful communications, and to elaborate their linguistic structures, while its focus-on-form instruction accelerates the development of their grammar ability (Ellis, 2014; Long, 2014). The criteria of TBLT tasks that TBLT proponents all agree with are the following: meaning based, goal oriented, outcome-evaluated, and authentic (Skehan, 1998a). When it comes to "authenticity," it does not need situational authenticity, but at least it needs interactional authenticity. In other words, tasks themselves do not have to mirror what they do in real life, but they have to be able to prepare the learners for the language use in real life (Nunan, 1989; Ellis, 2003; Long, 2014).

TBLT supporters have slightly different conceptions of TBLT pedagogy, especially on pedagogical interventions (Skehan, 1998b; Robinson, 2001; Ellis, 2003; Long, 2014). TBLT

proponents all agree on the importance of learners paying attention to their language forms during meaning negotiations, but how to attract the learners' attention on these language forms differ for several researchers. Ellis (2003) and Long (2014) maintain learners' attention to forms are mainly induced by instructors' weak intervention using focus-on-form instructions. Focus-on-form instructions are just-in-time grammar instructions that immediately shift the learners' attentions to linguistic forms whenever they make errors in meaning negotiations (Ellis, 2014; Long, 2014). On the other hand, Skehan (1998b) claims that learners' attentions to form should be implicitly induced by manipulating task characteristics and conditions. Robinson (2001) possesses a neutral position in which manipulating tasks characteristics and conditions and teachers' interventions with focus-on-form instructions are both essential.

Both Ellis (2003) and Long (2014) believe that teachers' explicit or implicit focus-on-form instructions bring efficiency and accelerate the learners' attainment of the target linguistic forms. Long says, without teachers' interventions, learners might not be able to notice their errors in their interlanguages because some erroneous forms that learners make do not hinder their communications. Ellis focuses on explicit intervention of focus-on-form instruction. He argues that explicit instructions promote learners' learning in two ways. First, explicit instructions, especially negative feedback, allow learners to clearly notice the gaps between their linguistic forms and others' forms, which in turn promotes modification of their output and accelerates their mastering the target language structures. Second, explicit instructions help teachers clearly understand which linguistic forms learners are trying to master and their progress. Long stresses implicit intervention of focus-on-form instruction, especially recasts, where teachers repeat or reformulate learners' linguistic errors in their utterances. Long argues that recasts focus on conveying only necessary information about the target language in context, which promote learners' form-function mapping efficiently and promote learners' noticing correct forms. Recasts also provide motivation and attention to the fact that their messages are at stake and they already understand some portions of the message.

On the other hand, Skehan (1998b) believes manipulating task complexity induces learners' attentions implicitly and brings great progress in the learners' language performance. Skehan's (1998b) trade-off hypothesis argues learners' attentional resources are provided from only one resource pool, and fluency, complexity and accuracy cannot develop at the same time. He argues that without teachers' interventions, learners' attentions on each aspect of linguistic performance can be manipulated by controlling task conditions and characteristics. For example, providing more familiar topics and clearly structured tasks improve learners' accuracy

and fluency, and providing more interactive tasks augments their accuracy and complexity (Skehan & Foster, 1997). In terms of teachers' interventions, Skehan (1998b) believes that teachers' interventions hinder learners' automatization, in which learners' spontaneous noticing of gaps and restructuring their own interlanguages control their L2 acquisition.

Robinson (2001, 2005) stands between these interventionists and noninterventionists. Like Skehan (1998b), Robinson (2001) believes that manipulating task complexity can improve learners' language performance, but his concept of learners' attention deployment on language fluency, complexity and accuracy is different from Skehan's (1998b) understanding. Robinson's (2001) 'Cognition Hypothesis' argues that there are multiple attentional resource pools, and learners' language complexity, accuracy, and fluency never compete as long as attention resources are provided from different pools. According to Robinson (2001, 2005), task complexity is manipulated in two dimensions: resource-dispersing and resource-directing. In the resource-dispersing dimension, task complexity is manipulated by controlling the amount of the attentional resource from each attentional resource pool, and in the resource-directing dimension, task complexity is manipulated by directing learners' attentions to a specific attentional resource pool and controlling the competitions of these attentions. Robinson (2001, 2005) believes that increasing complexity in the resource-directing dimension stimulates learners' attentional and memory resources and leads to noticing gaps between target input and learners' output, which develops their L2s. In addition, Robinson (2001) claims teachers' focus-on-form instructions are essential. It compresses a long period of time that is needed for learners to be exposed to the target language input extensively and to notice the gaps between their input and interlanguages.

As mentioned above, unlike the basic concepts of TBLT tasks, there is no consensus on the pedagogical interventions in TBLT. Further investigation of how L2 learners' automatization of input processing happens will help expose the efficacy of teachers' interventions in TBLT. The following section will define adult L2 learning success.

Adult L2 learning success can be defined in numerous ways because every L2 learner has a different concept of L2 learning success based on his or her future L2 needs. In the difficulty of conceptualizing L2 learning success, the concept of the success this paper adopts is the one commonly used in several empirical studies focusing on their successful adult L2 learning participants: near-nativeness or nativelikeness. (Birdsong, 1992; Ioup, Boustagui, Tigi, & Moselle, 1994; Moyer, 1999; Bongaerts, 1999; DeKeyser, 2000; Abrahamsson, & Hyltenstam, 2008). The nativelikeness described in these studies was measured in two

different dimensions: (1) L2 production features and (2) L2 knowledge.

L2 production features are measured and analyzed by several different methods, such as error analysis, obligatory occasion analysis, frequency analysis, interactional analysis, conversation analysis, and so on (Ellis & Barkhuizen, 2005). In the studies that depicted the adult L2 learners achieving near-native level of proficiency, the qualitative coding of data and the quantitative analysis of learners' complexity, accuracy, and fluency are often adopted. Qualitative data are normally gathered through naturalistic observation, open-ended interviews, introspection/retrospection, field notes, and life-stories. In the analysis of learners' complexity, accuracy, and fluency, accuracy was measured by examining the number of errors, complexity was measured by counting the number of subordination, and fluency is measured by calculating the rate of production and the number and length of pauses and false starts (Ellis & Barkhuizen, 2005).

Explicit and implicit language knowledge that L2 learners possess is also often measured to analyze L2 proficiency. In this measurement, L2 proficiency is believed to be the accumulation of linguistic building components that the learners possess. In Purpura's (2013) conceptualization of language knowledge, grammatical knowledge plays a key role. He argues that form-meaning mappings occur in grammatical knowledge, which allows learners to use L2 at a pragmatic level. According to Purpura, several empirical studies provide the evidence of this process. That is, the learners' knowledge of grammatical form interacted with their knowledge of the semantic meaning, which conveyed literal and pragmatic meanings in the learners' utterances. The Grammaticality Judgment Test (GJT) focuses on learners' grammatical knowledge to measure adult learners' L2 proficiency and has been administered in several studies (Johnson & Newport, 1989; DeKeyser, 2000; Ellis, 2009). GJT examines learners' knowledge of syntax and morphology, and the examinees are asked to distinguish between grammatical and ungrammatical sentences. The following section will introduce the empirical studies that depict the adult L2 learners' near-nativeness or nativelikeness from these two dimensions and examine the learner variables that presumably caused their success.

What variables contribute to adult L2 learning success?

This paper categorizes the learner variables found in the L2 learners who have attained a near-native level of proficiency into two types: cognitive variables and non-cognitive variables, adopting the definitions from Farkas's (2003) study. According to Lee and Shute (2009), Farkas's definitions of cognitive and non-cognitive variables have been adopted in many

studies in economics and sociology, as well as in psychology and education. According to Farkas, cognitive variables are the traits, behaviors, and skills that are measured by objective test scores, and non-cognitive variables are ones that are assessed by more subjective rating systems, such as surveys or observations. According to Hall (2011), there are several different cognitive and non-cognitive variables that allegedly influence L2 development, such as age and aptitude for cognitive variables, and motivation, attitudes, personality, learning styles, and learning strategies for non-cognitive variables. This paper, however, will only focus on the variables that were observed in successful adult learners who have achieved near-native or native-like L2 proficiency. The subsequent sections will describe these cognitive and non-cognitive variables that presumably affected adult L2 learners' success.

According to several researchers (Schneiderman & Desmarais, 1988; Ioup et al., 1994; DeKeyser, 2000; Abrahamsson & Hyltenstam, 2008), some of the successes of adult L2 learning are attributed to a cognitive variable: high language-learning aptitude. Language-learning aptitude is human's innate linguistic information processing ability indicated by the scores of aptitude tests (Doughty, 2014).

Language-learning aptitude was originally theorized by Carroll and Sapon (1959), and it consists of four factors: phonetic detecting, syntax comprehension, memorization, and inductive learning. Carroll's 'Modern Language Aptitude Test' (MLAT) has since played an increasing role in assessing adult language-learning aptitude (Abrahamsson & Hyltenstam, 2008). The updated version of Carroll's model, Hi-LAB (High-Level Language Aptitude Battery) was introduced by Doughty, Campbell, Mislevy, Bunting, Bowles and Koeth (2010). This model takes into account the understanding of the human memory system: working memory and long-term memory. It can predict adult learners' L2 acquisition especially at the advanced level.

DeKeyser (2000) and Abrahamsson and Hyltenstam (2008) both investigated how language-learning aptitude works in L2 learning. DeKeyser tested Bley-Vroman's (1988) Fundamental Difference Hypothesis, which posits that child language acquisition happens implicitly by simply accessing Universal Grammar (UG), whereas adult language acquisition instead needs learners' focuses on the linguistic structures of the L2 inputs. DeKeyser did a replication study of Johnson and Newport (1989) and tried to find the relationship between learners' language-learning aptitude and their L2 proficiency as well as the aptitude differences between adult and child L2 learners. DeKeyser administered a grammaticality judgment test (GJT) and the Hungarian adaptation of the MLAT to 57 Hungarian learners of

English. All the participants had been the residents in the United States at least for 10 years, and their age at arrival was from 1 to 40. As predicted in his hypothesis, DeKeyser found that the 5 out of 6 adult L2 learners who achieved near-native proficiency possessed high aptitude. In addition, he found a significant correlation between the learners' aptitude and their L2 proficiency. Yet, he could not find any correlations between child learners' aptitude and their L2 proficiency. DeKeyser concluded that reaching an exceptionally high level of L2 proficiency after puberty needed activation of language-learning aptitude. Conversely, child L2 learners attain a high level of L2 proficiency only by activating innate implicit language-learning mechanism.

Abrahamsson and Hyltenstam (2008) also investigated the role of language-learning aptitude on attaining L2 proficiency and the relationship between learners' ages and the scores of the test in a different language. In the experiment, GJT and MLAT were administered to 42 near-native Swedish speakers whose L1 is Spanish and assessed the correlation between their L2 proficiency and language-learning aptitude. Their average length of residence in Sweden was 25 years. The result confirmed DeKeyser's (2000) finding. The near-native adult L2 learners had a higher language aptitude score than the near-native child L2 learners. However, unlike DeKeyser's study, Abrahamsson and Hyltenstam found a correlation between child L2 learners GJT scores and their aptitude test scores. Based on the fact that language aptitude was activated during the children's early L2 acquisition, they concluded that language aptitude is essential not only for adult learners, but also for young learners for their L2 learning.

Both DeKeyser (2000) and Abrahamsson and Hyltenstam (2008) found the necessity of high language-learning aptitude for adult learners to attain a high level of L2 proficiency, and the finding of Abrahamsson and Hyltenstam suggests the importance of language-learning aptitude in any L2 learning. Besides language learning aptitude, no other cognitive variables were found in the studies that depict adult learners' high level L2 proficiency. However, DeKeyser (2000) and Abrahamsson and Hyltenstam (2008) instead found several non-cognitive variables that allegedly contributed to the high level of proficiency of adult L2 learners in their studies.

Non-cognitive variables, which are measured not by test scores, but by subjective rating systems, such as surveys or observations, can also play key roles in learners' L2 attainment. The studies on achievement of near-native level or native-like proficiency by adult L2 learners have observed three of these variables in successful learners (Ioup et al. 1994; Moyer, 1999;

Bongaerts, 1999; DeKeyser, 2000; Birdsong, 2007; Abrahamsson & Hyltenstam, 2008). One of them is explicit learning. Explicit learning is an extraneous non-cognitive activity, which according to Ellis (1994), is "a conscious operation where the individual makes and tests hypotheses in a search for structure." DeKeyser and Hall (2009) maintain that it is indispensable for adult L2 learners, since it compensates for the implicit learning ability that adults have lost after puberty. The explicit grammatical rules these learners receive in their learning provide shortcuts, and it accelerates the speed of their L2 learning. The subsequent sections will describe how explicit learning becomes effective comparing the studies of Ioup et al. (1994) and Schmidt (1983).

Ioup et al. (1994) described two cases of attainment of near-native L2 proficiency in their study. Both learners, Julie and Laura, are English speaking L2 learners of Arabic, who were chosen for this case study because native speakers could not recognize them as non-native. Julie came from Britain to Cairo, Egypt when she was 21 years old, and had since then been living there for 26 years. She had never received any formal language instruction of Arabic, but constantly communicated with her Egyptian husband and two children in Arabic. In the first four months after coming to Cairo, she kept a diary about her Arabic learning and also received continuous explicit feedback from native speakers when she made errors in communication. Laura, who was American, had been living in Morocco for 10 years. She had received many years of formal instruction in Arabic. She started learning Arabic when she was an undergraduate and continued studying it until a Ph.D. Program in the U.S. She moved to Cairo in the middle of the program to pursue the acquisition of spoken Arabic and had been an Arabic teacher since then. She also had been married with an Egyptian man for 10 years. When comparing these two learners, Julie showed higher fluency than Laura, even though both of them had achieved an exceptionally high level of proficiency. Ioup et al. (1994) attributed Julie's success to her explicit learning, specifically grammar-focused self-teaching with native speakers' immediate feedback. Compared to Laura, who had kept taking formal language classes to improve her L2 skills, Julie had actively engaged with native speakers in informal settings. She also received error feedback from native speakers whenever she made errors in everyday conversations. This was further enhanced by her self-study, as she kept learning notes on the learned grammatical structures, especially morphological variations.

Schmidt's (1983) case study of an adult L2 learner also describes the efficacy of explicit learning on adult L2 learners. Schmidt (1983) conducted a three-year longitudinal case study of Wes, a 33-year-old Japanese man who achieved high English communication skills in only

three years without any formal language instructions. Wes had moved to Hawaii because of his profound interest in and great affection for Hawaii and Hawaiian people. The comparison of this study with Ioup et al.'s (1994) case of Julie provides an insight into the role of explicit learning. Like Julie, Wes also constantly interacted with native English speakers and received error feedback from them. However, Wes's language had little accuracy in grammar. Unlike Julie, who always listened to the feedback she received and made tremendous efforts in acquiring unseen linguistic elements of the target language, Wes ignored the feedback he received from native speakers, which resulted in his inability of attaining the L2 accuracy that native speakers could not recognize as nonnative. This comparison shows that explicit learning may become more effective when the learners are actively involved in grammar-focused self-study and try to internalize the feedback they receive from native speakers.

The study of Ioup et al. (1994) and Schmidt (1983) depict two other non-cognitive variables that helped promote the success of the adult L2 learners: immersion experiences and extremely high motivation for acquiring the target languages. In the study of Ioup et al. (1994), Julie immersed herself in an Arabic speaking community for 26 years and her motivation was being able to communicate effectively with her husband, two children and people in the community in Arabic. In the study of Schmidt (1983), Wes lived in Hawaii for three years, and his motivation was also becoming part of the Hawaiian community.

A short or long period of immersion experience of adult L2 learners who attained a near-native level or native like proficiency has been observed in several studies (Ioup et al. 1994; Moyer, 1999; Bongaerts, 1999; DeKeyser, 2000; Birdsong, 2007; Abrahamsson & Hyltenstam, 2008). The immersion experiences that these studies describe are mostly studying abroad and immigration to other countries. The efficacy of immersion experience on adult L2 proficiency development was focused in the study of Lindseth (2010) and Tschirner (2007).

Lindseth (2010) studied the improvement of the oral proficiency of adult German learners during a semester of studying abroad in Germany. The use of two specific advanced level grammar structures by 38 graduate students, who were majoring in German, were assessed by a well-known oral proficiency interview offered by ACTFL (ACTFL OPI). This interview was conducted by the professional ACTFL raters. As a result, 31 students improved their L2 proficiency after the one-month immersion experience including two students who reached an advanced level and seven students who improved their L2 proficiency by at more than one sublevel on the ACTFL Proficiency Guidelines for Speaking.

Tschirner (2007) also investigated the efficacy of immersion experience. His study investigated not only the effectiveness of immersion experience on improving learners' oral proficiency, but also how to maximize the immersion experience. Tschirner's (2007) study assessed the oral performance of 15 adult American learners of German before and after the four-week intensive immersion program in Germany. Two different instruments were used to measure the learners' oral proficiency: the Simulated Oral Proficiency Interview (SOPI) for the pretest and the Oral Proficiency Interview (OPI) for the posttest. The learners were also asked to fill in questionnaires to provide their biographical information and the head teacher provided written, detailed comments to each student. The study showed that at the end of the program, 11 of 15 participants improved their oral proficiency by at least one sub-level on the ACTFL Proficiency Guidelines for Speaking. Tschirner (2007) also found out the five following components that maximize people's immersion experience: understanding how to develop oral proficiency; intensive speech instructions with many opportunities to interact with different native speakers; homestays that provide large amounts of language input and speaking opportunities, and extremely high motivation. The only drawback of Tschirner's (2007) study was that two different instruments were used to measure oral proficiency: the SOPI for the pretest and the OPI for the posttest. He argued that he administered different tests because of the practical, scheduling and cost reasons. Even though Tschirner (2007) maintained that the scores of these two tests are highly correlated, it is problematic to compare the different test scores to observe the progress during the immersion program because they are not exactly the same measurements.

Even though both Lindseth (2010) and Tschirner (2007) discovered the efficacy of immersion experience on improving learners' oral proficiency, neither of these studies nor other studies delved deeper into what elements in immersion experience stimulate the learners' L2 development. However, Long's (2014) Interaction Hypothesis, and Robinson's (2001, 2005) Cognition Hypothesis corroborate the reasons immersion experiences accelerate adult L2 learning. Long's (2014) interaction hypothesis contends that, in addition to comprehensive input, modifying their interlanguages in meaning negotiation facilitates learners' L2 acquisition. As Tschirner (2007) mentioned, effective immersion experiences provide a lot of meaning negotiation opportunities in the interactions with native speakers. Robinson (2001, 2005) and Robinson and Gilabert (2007) argue that increasing task complexity in the resource-directing and resource-dispersing dimensions provides more opportunities of interactions and meaning negotiations in learners, which promotes their

noticing of new language forms. According to Robinson and Gilabert, learners experience a lot of task complexity in both resource-directing and resource-dispersing dimensions when they are exposed in the L2 environment. Tschirner argues that the ideal immersion experiences provide the learners with such an environment, where they engage in large amounts of speaking and interaction opportunities. As these theories suggest, immersion experiences promote adult L2 development.

The third non-cognitive variable is motivation. According to Skehan (1989), the exact nature of motivation is uncertain, and there is no consensus on the definition of motivation among SLA researchers. Moyer (1999) also mentioned the difficulty of measuring motivation and the scarcity of studies on the relationship between motivation and ultimate L2 attainment. Despite the ambiguity in SLA research on motivation, what is known about it is depicted as the crucial factor for attaining a near native level of or native-like L2 proficiency in many studies (Ioup et al. 1994; Moyer, 1999; Bongaerts, 1999; DeKeyser, 2000; Birdsong, 2007; Abrahamsson & Hyltenstam, 2008). Particularly, the studies of Bongaerts (1999), Birdsong (2007), and Moyer (1999) indicated the importance of motivation on attaining a near-native level of L2 pronunciation by adult learners.

Bongaerts (1999) conducted three experiments in his study: two with Dutch learners of British English and one with Dutch learners of French, aiming to find if adult L2 learners can achieve the L2 pronunciation that is indistinguishable from that of native speakers, and to discover what factors helped the learners achieve such a high-level of attainment. The first study compared the English pronunciation between three groups of participants: 5 native speakers; 10 adult Dutch learners of English whose proficiency levels are highly advanced according to English as a Foreign Language (EFL) experts; and 12 adult Dutch learners of English in different levels of proficiency. All participants were asked to do the following two tasks: (1) to talk about their recent holiday experiences and (2) to read 10 English sentences and 25 words. Then, four linguistically inexperienced native English speakers rated the accents in these speeches and reading samples. The result showed that the judges could not tell the pronunciation differences between the highly advanced L2 learners and the native speakers. The same results were found in the experiments conducted on the supraregional accent of British English and standard French; the judges could not distinguish the sounding differences between the highly advanced L2 learners and the native speakers. In this study, Bongaerts discovered the three common variables in the learners who achieved a near native level of pronunciation skills: receiving a large amount of L2 input from native speakers after

entering university, taking intensive perceptual and pronunciation trainings for acquiring accurate pronunciation of the target language, and exhibiting extremely high motivation in attaining native-like pronunciation. One drawback of Bongaerts's (1999) study is the lack of scrutiny. The participants' pronunciations were not analyzed at the segmental level, which made the accuracy of the result skeptical.

Birdsong (2007) reexamined Bongaerts's (1999) study. He replicated Bongaerts's study and analyzed successful adult L2 learners' pronunciations in both sentence level and segmental level. Twenty-two English speaking adult learners of French, who had experienced immersion in the French speaking countries, were assessed on the durations of their French vowels, voice onset time of their plosives and the sentence level pronunciations of three passages from French literature. He also investigated the relationship between the productions of the two levels. The recorded read-aloud of adult learners of French were compared with the ones of native French speakers who had grown up in France. The result showed 2 of the 22 participants achieved an exceptionally high-level of pronunciation in both segmental and sentence levels, and their pronunciation was indistinguishable from the native French speakers. He also found that the sentence level production predicts the segmental level production but not vice versa. Finally, he found that both of the successful learners who achieved near-native level of pronunciation in segmental and sentence levels had had formal phonetic trainings, had spent more than 20 years in France, and had been highly motivated in acquiring proper French pronunciation.

The studies of Bongaerts (1999) and Birdsong (2007) found the effect of high motivation of adult L2 learners on attaining a near-native level of L2 pronunciation, but they did not directly test the effect of motivation in their studies. Moyer (1999), instead, focused on the correlations between phonological proficiency of adult L2 learners and extrinsic variables: motivations, the starting ages of immersion and receipt of instruction, and the types and the amounts of instruction and feedback. Moyer found the negative correlation between the participants' phonological performance and the starting ages of immersion and receipt of instructions and the positive correlation between the participants' phonological performance and the amount of phonological instructions on stress, rhythm and intonation. While no correlations were found between the participants' phonological performance and other extrinsic variables including their motivations on acquiring the language, Moyer found one of the participants who achieved extremely high scores in all the pronunciation tasks and was rated as native-like by the all the four judges exhibited an exceptionally high motivation in

learning the German language.

Moyer also found there are two other non-cognitive variables that were mentioned earlier behind this participant's L2 success: explicit learning and immersion experience. He not only had an extremely high interest in Germany and the German language, but also self-taught through communication with his German friends, and had joined a one-year exchange program in Germany at the age of 22, which, according to him, brought rapid growth of his German skills. Successful learners' explicit learning and immersion experiences were also discovered in the studies of Bongaerts (1999) and Birdsong (2007). In Bongaerts's study, the successful learners had taken intensive perceptual and pronunciation trainings and had immersion-like experiences, in which they had received a massive L2 input from native speakers on a daily basis.

In the previous sections, the studies that had portrayed the adult L2 learners who attained a high-level of L2 proficiency were reviewed and the cognitive and non-cognitive variables that presumably caused a high-level of L2 competence were examined. The subsequent section will focus on the relationship between these cognitive and non-cognitive variables and TBLT.

Does TBLT enhance the learners' cognitive and non-cognitive variables and promote their L2 learning?

A cognitive variable, language learning aptitude, and non-cognitive variables, explicit learning, immersion experiences and extremely high motivation were identified as the stimulants that have been described in the previous studies as ones that could trigger adult L2 learning success. Whether or not TBLT can promote these learner variables will be discussed further.

The cognitive variable, which was allegedly contributed to adult L2 learners' high proficiency, was language-learning aptitude. Language-learning aptitude is, however, believed to be an innate fixed talent for learning languages in many researchers (Abrahamsson & Hyltenstam, 2008). If it is not developmental or does not fluctuate under external influences, the role of TBLT in successful language learning is uncertain. Since there were only a few studies that focused on the development or fluctuation of language-learning aptitude, it is hard to understand the mechanism of language-learning aptitude. The study of Harley and Hart (1997), which sought the relationship between the language-learning aptitude of 65 11th grade French learning students in immersion programs and the ages of first language exposure, could not find the evidence of the development of language-learning aptitude.

McLaughlin (1990), however, claims that language-learning aptitude is developmental because he believes people's prior successful language-learning experiences unconsciously influence the following new language learning. Doughty (2014) also argues that at least one component of language-learning aptitude, working memory, is developmental and can be trained. If the language-learning aptitude or a part of it is developmental as McLaughlin and Doughty claim, there is a possibility that TBLT can influence its development.

In addition, TBLT also might have a possibility to trigger the activation of language-learning aptitude. Carroll (1964) investigated the contributions of motivation, instruction qualities, time and intelligence on learners' language aptitude in one-week intensive language course and in a yearlong extensive language course. He found that when learners' motivation is low or the quality of instruction is low, their aptitudes were not activated. However, when they consider that the language courses are beneficial or entertaining, the learners were more motivated and their language learning aptitude started to make effects. Carroll's (1964) findings suggest that if TBLT stimulates learners' motivation, TBLT might also be able to stimulate adult L2 learners' language learning aptitude. Further research on aptitude development and activation is needed to discover the role of TBLT on promoting language-learning aptitude. In the subsequent section, the roles of TBLT on non-cognitive variables will be investigated.

Non-cognitive language learner variables discussed in this paper were explicit learning, immersion experience, and extremely high motivation. Firstly, TBLT provides effective explicit learning. In the study of Ioup et al. (1994), Julie received corrective feedback from native speakers whenever she made grammatical errors in everyday conversations. This is what TBLT's focus-on-form approach offers to the learners. As mentioned earlier, in the focus-on-form approach, a just-in-time grammatical instruction is given whenever the learners need it in meaning negotiation (Ellis, 2000; Robinson, 2001; Long, 2014). Focus-on-form instructions help develop the learners' language abilities because the learners become psycholinguistically the most susceptible to new forms when learners' attention is shifted from meanings to linguistic elements due to perceiving problems in comprehension and production (Prabhu, 1987; Long, 2005).

Secondly, when it comes to immersion experiences, there are no studies on the explicit relationship between TBLT and immersion experiences. However, TBLT provides similar learning environments to what actual immersion experiences offer during class periods. According to Révész (2009), TBLT aims to create the environment where the learners engage in implicit and explicit learning, receive rich exposure to the target language, and meaning

based communication. This notion matches with the five components that Tschirner (2007) argued for maximizing people's immersion experience. In addition, Long's (2014) Interaction Hypothesis and Robinson's (2001) Cognition Hypothesis suggest TBLT can create an immersion-like environment. The input and output processing through interactions with peers and instructors in classrooms expose the learners in comprehensive input and promote them meaning negotiation and output modification, which accelerate the learners' L2 acquisition. TBLT learners can also engage in many cognitively demanding tasks in both resource-directing and resource-dispersing dimensions. It provides them with ample opportunities of interactions and meaning negotiations with teachers and other learners in the target language, which allows them to notice new language forms and promotes their L2 complexity development. What TBLT lacks when compared to the actual immersion experiences are the quality and the quantity of input the learners receive. Unlike actual immersion experiences, the quality of the input that TBLT provides is not as high as that of the input that actual immersion experiences offer. TBLT learners receive native level input only from teachers and from the students who grew up abroad unlike being surrounded by real natives in the actual immersion experiences. The amount of the L2 input that TBLT learners can receive is limited, and it is normally only the time in the classroom. On the other hand, in the actual immersion experiences, learners receive the L2 input for a long period of time, sometimes all day long. Even though these differences from actual immersion experiences exist, TBLT still provides immersion-like experiences to its learners, which possibly accelerates their L2 attainment.

Finally, TBLT positively works on learners' motivation. According to Krashen's (1982) Affective Filter Hypothesis, negative affective factors, such as negative emotions, low self-confidence, and anxiety become psychological obstacles and hinder language learners' full absorption of available comprehensible input. However, TBLT adopts needs analysis, which eliminates these negative affective factors during the task design and helps teachers avoid psychological obstacles that hinder the learners' L2 attainment in the classroom. TBLT's needs analysis also takes account of learners' present or future communicative needs or interests (Long, 2014), which stimulates the learners' motivation and interests in the L2 development (Robinson, 2001; Tamponi, 2004). In the following studies, Chen and Brown (2012) have found the efficacy of TBLT on the learners' motivation.

Chen and Brown (2012) investigated whether or not task-based computer-mediated communication (CMC) approach facilitates adult English learners' motivation toward ESL writing. Six adult learners (18 to 33 years old) who study in the ESL program at a university

were assigned a task: to create informative websites using CMC tools. Their perceptions of their progress in English writing and the motivation for improving it were analyzed. The data was derived from the participants' interviews, observation notes and finished websites. The study found that the CMC supported TBLT facilitated the learners' motivation. The learners mentioned that relaying the knowledge they had in writing articles on the website made them feel like being actual professional web journalists. Since they had to open their websites to the public, they not only perceived the task as an authentic, purposeful, and meaningful activity, but also made tremendous efforts to improve the quality of their writing in order not to "lose face". In addition, many learners often visited the websites that the other students made, even though it was not the requirement of the task. Chen and Brown (2012) mentioned that this fact indicates how much many of the learners were intrinsically motivated in this task. They concluded that the task removed the learners' anxiety and generated their high motivation. The result of this study clearly demonstrates the efficacy of TBLT on motivation. It reduced the learners' anxiety and increased their motivation. However, there are limitations in this study. First, the sample size was small, which compromised the validity. The number of the participants was six, which was too small to generalize the findings to a bigger group of people. Second, the tasks were mediated through the computer, and there was a possibility that the learners were simply interested in computer-based activities, compromising the construct validity. In spite of these limitations, this study still deserves attention, as few studies have focused on the relationship between TBLT and L2 learners' motivation.

Conclusion

The aim of this paper was to investigate the roles of TBLT approach in promoting success in adult L2 learners. For the first research question, multiple cognitive and non-cognitive variables that promoted success were indicated in the reviewed studies. When it comes to cognitive variables, language-learning aptitude allegedly played a role in promoting a high level of adult L2 proficiency. As for non-cognitive variables, three different variables were identified as the stimulants for successful adult L2 learning: explicit learning, immersion experiences, and extremely high motivation. Then, for the second question, the possible role of TBLT on each variable was discussed. For the cognitive variables, even though TBLT cannot directly stimulate people's language-learning aptitude, it may be able to activate the aptitude by stimulating the learners' motivation for learning L2 (Carroll, 1964). On the other hand, TBLT presumably promotes two of the three non-cognitive variables of the successful adult

L2 learners: explicit learning, and extremely high motivation. When it comes to immersion experiences, TBLT also offers immersion-like experiences in which the learners only receive low quality L2 input. These findings suggest that over all, TBLT offers what the adult L2 learners who have achieved a near-native level of proficiency have experienced in their life. Provided that the tasks are always given in the areas of learners' needs and interests, effective feedback is provided instantly from instructors, and the target languages are meaningfully and constantly used by both instructors and learners in the process of accomplishing these tasks, it is possible for adult learners to achieve practical high-level of L2 competence through TBLT.

In addition, the studies reviewed above proved that explicit learning, immersion experience, and extremely high motivation were often observed together in the successful L2 learners who achieved a high level of proficiency (Ioup et al. 1994; Moyer, 1999; DeKeyser, 2000; Birdsong, 2007; Abrahamsson & Hyltenstam, 2008). This finding implies explicit learning, immersion experiences, and extremely high motivation might interact with each other to stimulate adult L2 learners and promote their L2 to a high level. In future studies, whether or not these three non-cognitive variables interact with each other and how TBLT promotes this interaction need to be investigated.

As Ellis (2003) and Gonzalez-Lloret and Nielson (2015) mentioned, the studies on the evaluation and the efficacy of TBLT program are scarce and the existing studies on the evaluations of TBLT programs are mostly qualitative and only focus on the experience and perception of the learners on the gained language skills (McDonough & Chaikitmongkol, 2007; Towell & Tomlinson, 1999). The efficacy of TBLT on promoting success in adult L2 learning also should be further discussed in future research.

References

Abrahamsson, N., & Hyltenstam, K. (2008). The robustness of aptitude effects in near-native second language acquisition. *Studies in Second Language Acquisition, 30*(04), 481-509.

Birdsong, D. (1992). Ultimate attainment in second language acquisition. *Language, 68*, 706-755.

Birdsong, D. (2007). Nativelike pronunciation among late learners of French as a second language. *Language experience in second language speech learning*, 99-116.

Bley-Vroman, R. (1988). The fundamental character of foreign language learning. *Grammar and second language teaching: A book of readings*, 19-30.

Bongaerts, T. (1999). Ultimate attainment in L2 pronunciation: The case of very advanced late L2 learners. *Second language acquisition and the critical period hypothesis*, 133-159.

Brumfit, C. (1984). *Communicative methodology in language teaching: The roles of fluency and accuracy* (Vol. 129, p. 33). Cambridge: Cambridge University Press.

Bygate, M. (2009). *Task-based language teaching: a reader*. John Benjamins Publishers.

Carroll, J. B., & Sapon, S. M.(1959). *Modern Language Aptitude Test*, New York: The Psychological Corporation.

Carroll, J. B. (1964). The Prediction of Success in Intensive Foreign Language Training. In R. Glaser (Ed.), *Training research and education*. Pittsburgh: Univer. Pittsburgh Press, 1962. Pp. 87-136

Chen, J. C., & Brown, K. L. (2012). The effects of authentic audience on English as a second language (ESL) writers: a task-based, computer-mediated approach. *Computer Assisted Language Learning, 25*(5), 435-454.

DeKeyser, R. (2000). The robustness of critical period effects in second language acquisition. *Studies in second language acquisition, 22*(04), 499-533.

DeKeyser, R., & Larson-Hall, J. (2009). What does the critical period really mean. *In Handbook of bilingualism: Psycholinguistic approaches*, 88-108.

Doughty, C., Campbell, S., Bunting, M., Mislevy, M., Bowles, A., & Koeth, J. (2010). Predicting near-native L2 ability. *Proceedings of the 2008 Second Language Research Form*. Cascadilla Press.

Doughty, C. J. (2014). Assessing aptitude. In Kunnan (Ed.), *The Companion to Language Assessment*, 25-46.Wiley-Blackwell.

Ellis, N. C. (1994). Implicit and explicit language learning. *Implicit and explicit learning of languages*, 79-114.

Ellis, R. (2000). Task-based research and language pedagogy. *Language teaching research, 4*(3), 193-220.

Ellis, R. (2003). *Task-based language learning and teaching*. Oxford: Oxford University Press.

Ellis, R. (2014). Taking the critics to task: The case for task-based teaching. *Knowledge, Skills and Competencies in Foreign Language Education*, 103-177.

Ellis, R., & Barkhuizen, G. P. (2005). *Analyzing learner language*, 51-275. Oxford: University Press, USA.

Ellis, R. (2009). *Implicit and explicit knowledge in second language learning, testing and teaching* (Vol. 42). Multilingual Matters.

Farkas, G. (2003). Cognitive skills and noncognitive traits and behaviors in stratification processes. *Annual review of sociology*, 541-562.

González-Lloret, M., & Nielson, K. B. (2015). Evaluating TBLT: The case of a task-based Spanish program. *Language Teaching Research, 19*(5), 525-549.

Hall, G. (2011). *Exploring English language teaching: Language in action*, 123-142. Taylor & Francis.

Harley, B., & Hart, D. (1997). Language aptitude and second language proficiency in classroom learners of different starting ages. *Studies in Second Language Acquisition, 19*(03), 379-400.

Ioup, G., Boustagui, E., El Tigi, M., & Moselle, M. (1994). Reexamining the critical period hypothesis. *Studies in second language acquisition, 16*(01), 73-98.

Johnson, J. S., & Newport, E. L. (1989). Critical period effects in second language learning: The influence of maturational state on the acquisition of English as a second language. *Cognitive Psychology, 21*(1), 60-99.

Krashen, S. (1982). *Principles and practice in second language acquisition* (Vol. 2). Oxford: Pergamon Press.

Lee, J., & Shute, V. J. (2009). The influence of noncognitive domains on academic achievement in K-12. *ETS Research Report Series, 2009*(2), i-51.

Lindseth, M. U. (2010). The development of oral proficiency during a semester in Germany. *Foreign Language Annals, 43*(2), 246-268.

Long, M. (2005). Current issues in task-based language teaching. Paper presented at the University of Illinois, Urbana-Champaign

Long, M. (2014). *Second language acquisition and task-based language teaching*. Malden, MA: Wiley-Blackwell

McDonough, K., & Chaikitmongkol, W. (2007). Teachers' and learners' reactions to a task-based EFL course in Thailand. *Tesol Quarterly*, 107-132.

McLaughlin, B. (1990). The relationship between first and second languages: Language proficiency and language aptitude. *The development of second language proficiency*, 158-178.

Moyer, A. (1999). Ultimate attainment in L2 phonology. *Studies in Second Language Acquisition, 21*(01), 81-108.

Nunan, D. (1989). *Designing tasks for the communicative classroom*. Cambridge: Cambridge University Press.

Prabhu, N. S. (1987). *Second language pedagogy*. Oxford: Oxford University Press.

Purpura, J. E. (2013). Assessing grammar. In Kunnan (Ed.), *The Companion to Language Assessment*, 100-124.Wiley-Blackwell.

Révész, A. (2009). Task complexity, focus on form, and second language development. *Studies in Second Language Acquisition, 31*(03), 437-470.

Robinson, P. (2001). Task complexity, task difficulty, and task production: Exploring interactions in a componential framework. *Applied linguistics, 22*, 27-57.

Robinson, P. (2005). Cognitive complexity and task sequencing: Studies in a componential framework for second language task design. *IRAL-International Review of Applied Linguistics in Language Teaching, 43*(1), 1-32.

Robinson, P., & Gilabert, R. (2007). Task complexity, the Cognition Hypothesis and second language learning and performance. *IRAL-International Review of Applied Linguistics in Language Teaching, 45*(3), 161-176.

Schmidt, R. (1983). Interaction, acculturation, and the acquisition of communicative competence: A case study of an adult. *Sociolinguistics and language acquisition, 137*, 174.

Schneiderman, E. I., & Desmarais, C. (1988). The Talented Language Learner: Some Preliminary Findings. *Second Language Research, 4*(2), 91-109.

Skehan, P. (1989). *Individual differences in second-language learning*. London: Edward Arnold.

Skehan, P. (1998a). Task-based instruction. *Annual Review of Applied Linguistics, 18*, 268-286.

Skehan, P. (1998b). *A cognitive approach to language learning*. Oxford: University Press.

Skehan, P., & Foster, P. (1997). Task type and task processing conditions as influences on foreign language performance. *Language teaching research, 1*(3), 185-211.

Tamponi, A. R. (2004). Content and language integrated university course: A task-based approach. Paper resented at the Navigating the New Landscape for Language Conference, University of London.

Towell, R., & Tomlinson, P. (1999). Language curriculum development research at university level. *Language teaching research, 3*(1), 1-32.

Tschirner, E. (2007). The development of oral proficiency in a four-week intensive immersion program in Germany. *Die Unterrichtspraxis/Teaching German*, 111-117.

Van Den Branden, K., Bygate, M., & Norris, J. (2009). *Task-based language teaching: a reader*, 2-19. John Benjamins Publishers.

Willis, J. (1996). *A framework for task-based learning*. Harlow: Longman

Difficulty levels of English /r/ and /l/
for Japanese learners: An acoustic analysis

Aya Kitagawa

Abstract

The present study aimed to examine phonetic and phonological learning with a specific focus on production of /r/ and /l/ and to identify the difficulty level of these approximants by defining /r/ and /l/ as easy, learnable or difficult items for Japanese learners of English. Acoustic analyses were conducted, where the third formant was measured for /r/ and /l/. The variables obtained from the analyses were submitted to three statistical tests, a cluster analysis, a multivariate analysis of variance and a discriminant analysis. The results showed that both approximants were difficult for Japanese learners to learn to produce. Despite the difficulty in learning them, it was also found that there was an individual preference about which approximant was learned faster. Some Japanese learners of English were learning /r/ faster than /l/ whereas others were learning /l/ faster than /r/. Further studies will be required to explore what creates these individual differences, and how positively the preference for learning /r/ or /l/ can affect the learning of both approximants.

1. Introduction

1.1 Purpose of the study

Previous studies have pointed out the difficulty of Japanese learners learning to perceive and produce English /r/ and /l/ (Flege, Guion, Akahane-Yamada, & Yamada, 2004; Goto, 1971; Yamada, 1995). This is partly because of the difference in the phonological inventory of approximants between the two languages: English has four voiced approximants, /r, l, j, w/, whereas Japanese has only two approximants, /j, w/. Japanese does not have a voiced alveolar approximant, /r/, and a voiced alveolar lateral approximant, /l/, unlike English (English /r/ should be transcribed as [ɹ] if based on IPA, but /r/ will be used in this paper following the conventional transcription).

It should also be noted that although previous research often has found that Japanese learners of English tend to perceive and produce the English syllable-initial /r/ and /l/ as Japanese /ɾ/, English [r] and Japanese [ɾ] are phonetically different. Japanese /ɾ/ is not even classified as an approximant. It is a flap, labelled as [ɾ] or [ɽ], which is articulated with a brief closure immediately before the following sound by quickly contacting the tip of the tongue with the alveolar ridge (Kent & Read, 2002). It rather sounds similar to an English flap used in American English, as in *better*. Japanese /ɾ/ and English /r/ thus have very different phonetic qualities.

The contrastive phonetics and phonology above suggest that Japanese learners of English, less experienced learners in particular, would have difficulty in producing /r/ and /l/ in a native-like manner. There seems to be a lack of consensus as to which approximants are more likely to be learned, however, especially concerning production. Some argued that /r/ was learned with more ease (Aoyama et al., 2004; Hazan, Sennema, Iba, & Faulkner, 2005), while others maintained that both could be equally learned (Flege, Takagi, & Mann, 1995; Slawinski, 1999). This study therefore analyzed the production of these two approximants by Japanese learners with no experience of living in an English-speaking country, using acoustic analyses, in order to define the level of difficulty for these approximants.

1.2 Learning of L2 /r/ and /l/

The difficulty of discriminating perceptually between English /r/ and /l/ for Japanese learners was empirically examined (Goto, 1971; Hallé, Best, & Levitt, 1999; Yamada, 1995). Goto (1971) conducted an experiment, where American and Japanese participants read a list of words including /r/ and /l/ tokens, and then identified the tokens as /r/ and /l/ by listening to their own recorded samples or the other participants'. He found that both proficient and less proficient Japanese participants discriminated between /r/ and /l/ in perception poorly. Yamada (1995) is in accordance with Goto (1971), but claimed that the experience of living in the U.S. affected accuracy in perceiving these approximants.

Whereas the findings of Guion, Flege, Akahane-Yamada, and Pruitt (2000) showed that no Japanese sounds, including vowels and consonants, were similar to /r/ and /l/ perceptually, they reported that /l/ was closer to /ɾ/, a Japanese flap. In their experiment, highly experienced Japanese learners and moderately experienced Japanese learners performed better in discriminating between English /r/ and Japanese /ɾ/ than between English /l/ and Japanese /ɾ/. Guion et al. regarded this as an indication that English /l/ sounded closer to Japanese /ɾ/ than

English /r/ to Japanese learners.

According to Goto (1971), Japanese participants with higher English proficiency discriminated between /r/ and /l/ well in production. Therefore, whereas prior studies generally agreed that both /r/ and /l/ were perceptually difficult for Japanese learners to discriminate, these approximants may be more likely to be learned in production. Aoyama et al. (2004), however, found a better performance of /r/ in learners' production, as in their perception. They carried out experiments, where adult and child speakers of Japanese discriminated between /r/, /l/ and /w/ in perception and production. They reported that the children improved the production of /r/ and /w/ more than /l/, while the adults showed only a minimal improvement in learning these approximants. Although both children and adults performed better in producing /l/ than producing /r/ and /w/ at the first session, they concluded that there was a greater improvement in the production of /r/ and /w/, highlighting a relative improvement. Hazan et al. (2005) also found some differences between /l/ and /r/. According to their results, whereas the production test did not produce a significant difference in the effect of the perceptual training with audiovisual stimuli between learning /l/ and learning /r/, the rating task showed that the production of /r/ was better rated as an authentic token due to the effect of the training. That is to say, the participants performed slightly better in producing /r/ than /l/ as a whole.

In contrast, Slawinski (1999) argued that the learning of /r/ and /l/ was parallel in the production. She carried out experiments of both perception and production of /r/ and /l/ by Japanese children and adults to investigate the effect of spoken proficiency on the production of /r/ and /l/. Four groups of Japanese children and three groups of adults participated in the experiments. The production test was aimed at examining how the participants would use temporal and spectral cues in discriminating between /r/ and /l/, the results of which indicated that they improved the second formant (F2) and third formant (F3) transitions of both /r/ and /l/ with age and exposure to English. The experimental groups did not differ significantly, except that the adult late learners used a longer cue for /l/ at a significantly different level from the other adult groups. Flege, Takagi et al., (1995) also found that Japanese learners of English could learn to produce both /r/ and /l/ accurately, as they became more experienced.

1.3 Acoustic measurements of approximants

The approximants targeted in the current study, /r/ and /l/, are similar to vowels in that

they have clear formant frequencies. This provides them with vowel-like features, making these sounds acoustically distinct from other consonants. The sonority of these sounds is thus higher than other consonants, including plosives, fricative and affricates. Espy-Wilson (1992) noted these characteristics, and analyzed all four English approximants in terms of a decrease in energy at low frequencies, abrupt amplitude change, mid-frequency energy and the first four formants.

F3 is recognized as one of the primary cues used to distinguish between /r/ and /l/ of all measurements. The decrease of F3 is a characteristic feature of /r/, and the lowering of F3 in /r/ has been measured in previous studies (Flege, Takagi et al., 1995; Iverson et al., 2001; Saito & Lyster, 2011). Flege, Takagi et al. (1995) highlighted the importance of lowering F3 for English /r/, noting that a higher F3 value led to more perceived foreign accentedness. Figure 1.1 depicts F3 of [r] in rats and [l] in *learn*. The boxed portion in each spectrogram corresponds to the whole /r/ and /l/. The horizontal line immediately below the arrow is F3, measured in hertz (Hz).

(a) [r]   (b) [l]
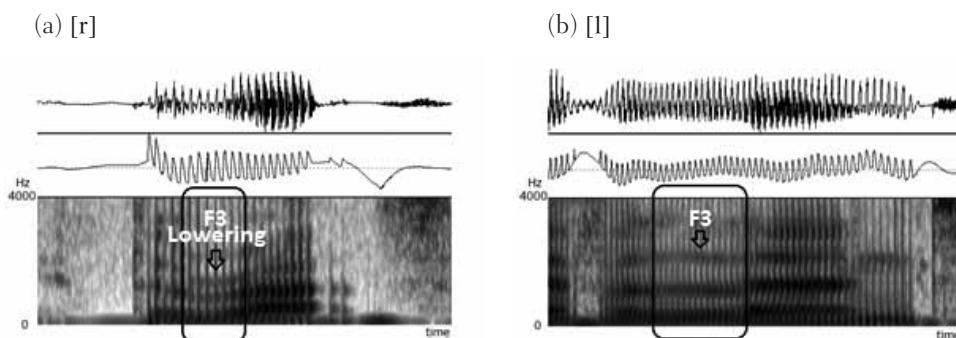
Figure 1.1. **The spectrogram of [r] and [l]: (a) [r]; and (b) [l].**

A relatively wider range of F3 values have been reported for English /r/, which reflects a different degree of r-coloring. Some initial /r/ was produced as low as 1240 Hz when articulated with a strongly curled tongue; in contrast, because of the lesser degree of r-coloring, F3 in the intervocalic /r/ is lowered only to a smaller extent (Ladefoged, 2003). Ladefoged (2003) described the intervocalic /r/ in *berry* as having the F3 value of 2100 Hz. Saito and Lyster (2011) maintained that the degree of F3 lowering affected native speaker judgment, reporting that /r/ tokens with F3 values ranging from 2200 Hz to 2300 Hz were judged as good examples of English /r/ of all tokens produced by Japanese learners of English.

In contrast, as evident in a comparison with F3 of the sounds preceding and following /l/ in Figure 1.1, F3 of /l/ has no abrupt change. Flege, Takagi et al. (1995) found that the average F3 value of /l/ produced by native speakers of American English was 2854 Hz. Saito and Lyster also noted that authentic /l/ would be produced with F3 of 2800 Hz.

Another acoustic cue of /l/ to be noted is the presence of anti-formants (Kent & Read, 2002). The spectrogram of [l] in Figure 1.1, surrounded by vowels, shows that formants of /l/ are also generally weaker than those of the adjacent vowels. Because /l/ is articulated with the tip of the tongue on the alveolar ridge, the air flow goes out through the side(s) of the oral cavity. This blockage causes the energy to radiate, which is reflected as anti-formants on the spectrogram, as with nasals.

1.4 Research question and hypotheses

This study aimed to examine the learning of /r/ and /l/ with a specific focus on production and to identify the difficulty level of these approximants by defining /r/ and /l/ as easy, learnable or difficult items for Japanese learners of English. To address this research question, hypotheses were formed as follows.

It was hypothesized that /r/ and /l/ would be learnable and difficult, respectively. Both approximants do not have any exact counterpart in the Japanese phonological inventory as noted in Section 1.1, which suggests that Japanese learners have to learn both phones from scratch. Previous research shows that training (Saito & Lyster, 2001) and exposure or age (Slawinkski, 1999) would be key factors in promoting learning, and therefore, these findings imply no strong ground for hypothesizing that these approximants would be easy items. Additionally, some researchers claim that the degree of difficulty varies between /r/ and /l/ as far as production is concerned: /r/ is more likely to be learned (Aoyama et al., 2004; Hazan et al., 2005). Although others argue that both approximants could be learned (Flege, Takagi, & Mann, 1995; Slawinski, 1999), it is not a major argument that /r/ is easier than /l/ for Japanese learners to learn to produce. Thus, /r/ was predicted to be learnable and /l/ to be difficult.

2. Methods

2.1 Participants

Ninety-one speakers participated in this study: 72 were Japanese learners of English (JL), 12 were native speakers of British English (BN) and 7 were native speakers of American

English (AN). The JL participants were third-year students at boys high school well reputed for its high academic level. They had diverse academic backgrounds, but none of them had the experience of living in an English-speaking country. The JL's performance data were compared against those of BN/AN obtained from publicly available databases, the UCL Speaker Database (Markham & Hazan, 2002) and the Audio Archive (Merfert, 1997). The data of these speakers with two different accents were used because both are the most common accents of English taught in the classroom in Japan and spoken around the world. It is known that gender affects the absolute values of formant frequencies, and thus, this study only collected data from male speakers.

2.2 Materials

Phonetically-balanced passages, *The Story of Arthur the Rat* and *Arthur the Rat*, were used. Data for the BN and JL participants were collected using the former passage and those for the AN participants, using the latter. These passages were slightly different in some words used, but they follow exactly the same story line.

Target words were selected from these passages. Voiced approximants, /r/ and /l/, occurred either at the word-initial position or the word-medial position as follows. The number within the square brackets in the list presents the number of repetitions, when provided:

> BN/JL data
> > /r/: *rat* [5]*, rainy, room, roof* [2]*, right, rode, carry* and *hurry*
> > /l/: *learn, looked, lived, last, later, lying, Helen* and *Nelly*

> AN data
> > /r/: *rat* [6]*, rainy, room, roof, right, rode, carry* and *hurry*
> > /l/: *like, look, loft, little, long, line, Helen* and *Nelly*

2.3 Recording and procedure

All recordings of the JL participants were made using a digital recorder, Roland-09, and a condenser microphone, SONY ECM-MS957. Their data were recorded at a sampling rate of 44.1 kHz, 16 bit. The recording level was first checked and adjusted to each speaker.

The material, printed on one side of A4 paper, was distributed to each participant between 3 days and 30 minutes prior to the recording. A summary of the story was presented in

Japanese on the other side of the paper, with the intention of helping them to grasp the gist of the story. Although the participants were allowed to look up the pronunciation of unfamiliar words in a dictionary before the recording session, no instructions were given by the experimenter as to phonetic and phonological features.

2.4 Acoustic analyses

Based on the finding of Saito and Lyster (2011) that only F3 values predicted whether the native listeners would perceive a produced sound as /r/, F3 was acoustically analyzed in this experiment. F3 was measured for /r/ and /l/, after the spectrogram and formant track were specified on Praat (Boersma & Weenink, 2011, 2015). Because the low F3 value characterizes /r/ as described in Section 1.3, the lowest F3 values at the beginning of the upward slope and the steady-state F3 value were measured for each token. The values were obtained in Hz.

One thing that should be considered when analyzing the speech sample collected from non-native speakers is that the F3 values cannot be measured when another sound is substituted for the target /r/ and /l/. There were two cases of this; one was the substitution of a flap-like sound and the other was that of a vowel-like sound. The flap-like pronunciation is evident from the presence of a hold phrase in most cases. When the presence of the hold phrase could be confused with the presence of anti-formant for /l/, a durational cue was applied to judge whether the token was /l/ or a flap-like sound, referring to the duration of a flap obtained by Rimac and Smith (1984). Both durations and F3 values were therefore recorded with the candidates for the flap-like tokens that had an unclear hold phrase. The articulation rate, calculated using the script (de Jong & Wempe, 2009) on Praat (Boersma & Weenink, 2011, 2015), was also obtained so as to take into account the difference in the speaking rate between the BN/AN participants and the JL participants. The other type of substitution, vowel-like pronunciation, was due to an incomplete articulation of /l/. The tokens were classified into this type of error, when characteristics of anti-formant were not visually evident in the spectrogram and waveform.

2.5 Variables for /r/ and /l/

Two variables were applied to the statistical tests on the production of /r/ and /l/: score for the /r/ and score for the /l/ tokens, which corresponded to the number of tokens produced as the intended sound as in Table 2.1. These variables were obtained by judging every target token as /r/, /l/ or other sounds with reference to the BN/AN data, and counting the number

of tokens identified as intended.

Table 2.1

Variables for the Analysis of /r/ and /l/

Variable	No. of variables	Level of measurement	Scale
Score for the /r/ tokens	1	Interval	0-8
Score for the /l/ tokens	1	Interval	0-8

After the F3 values were measured, they were converted from Hz to mel using Equation 1 (Fant, 1968):

$$\text{Mel} = 1000/\log_2 10 \times \log(1+F/1000) \qquad (1)$$

where F represents the frequency value. The thresholds of the F3 mel value to distinguish between /r/ and /l/ and that of the durational value to separate /l/ from a flap-like sound were then set based on the data of BN/AN participants. As for the F3 threshold to classify the tokens into /r/ or /l/, all tokens of initial /r/ and /l/ obtained from the BN/AN participants were ranked according to F3, and the F3 mel values whose z-scores fell at 2 SD and -2 SD were defined as the thresholds for /r/ and /l/, respectively. The durational threshold for a flap-like sound, on the other hand, was set using the average duration of American flaps for adult speakers, reported by Rimac and Smith (1984), 33 ms, as a reference. Non-native speakers are likely to speak more slowly than native speakers (Munro & Derwing, 1995); therefore, a modified threshold for the JL participants was calculated by multiplying 33 ms by the ratio of the average articulation rate of the JL participants to that of the BN/AN participants.

Every item was scored in terms of whether they were produced as intended or not, with reference to the threshold values above. First, by comparing the duration of the candidates for the flap-like tokens against the threshold of the durational value for a flap-like sound, tokens longer than the threshold were considered not to be flap-like. These tokens were submitted to the subsequent scoring process to judge whether or not the tokens were /r/ or /l/. The /r/ tokens with F3 lower than threshold F3 value for /r/ and the /l/ tokens with F3 higher than the threshold F3 value for /l/ were judged as intended. The two variables were obtained on a

scale of 0 to 8 by adding the scores for the eight items of /r/ and /l/ each. The target word *rat* was repeated in the passage five times for the BN and JL data and six times for the AN data, and *roof* was repeated twice for BN and JL data. These words were scored as intended when more than one token was judged as intended for *rat* and when at least one token was judged as intended for *roof*.

2.6 Statistical analyses

Three statistical analyses were performed using the variables above: a cluster analysis, a multivariate analysis of variance (MANOVA) and a discriminant analysis. The first statistical test, a cluster analysis, was carried out to group the participants into clusters, based on similarities in the input variables. This analysis was conducted using the entire sample, including the BN/AN participants, which made it possible to form the groups of JL participants depending on similarities in their performances. The cutoff point was selected on two criteria. One is that at least one of the clusters consisted of as many BN/AN participants as possible, called a BN/AN cluster. This is based on the theoretical hypothesis that the BN/AN participants would be grouped together. The other was that the JL participants formed four clusters at most, called as JL clusters, considering the balance of the sample size of each cluster for the subsequent analyses. The second statistical test, a MANOVA, was carried out using the clusters generated by this analysis as the between-subjects independent variables. It revealed whether there was a statistical difference among the clusters or not. The third statistical test, a discriminant analysis, was performed to identify clusters that differed in the variables at a statistically significant level. This judgment was based on the distance displayed in the canonical discriminant function plots and the location with reference to the group centroids indicated by positive and negative signs. A discriminant analysis additionally showed which variables discriminated them and to what degree these variables contributed to the discrimination. This was judged based on the structural matrix of the correlations between the variables and each of the discriminant functions. There is no decisive standard in the interpretation of the correlations, but those higher than .33 were interpreted to suggest variables contributing to the discrimination, following the convention provided by Tabachnick and Fidell (2007).

2.7 Criteria of learning

In order to define the difficulty level of /r/ and /l/ as easy, learnable or difficult, the results

were discussed by comparing the BN/AN cluster(s) with the JL clusters. The definition was then given against the following criteria. The first criterion was whether the target item discriminated between the BN/AN cluster(s) and the JL clusters. The items that did not discriminate between the JL clusters and the BN/AN cluster(s) were interpreted as easy for Japanese learners of English. The second criterion was how many JL participants were discriminated from the BN/AN cluster(s) when the target items differentiated between the JL clusters and the BN/AN cluster(s). The items that discriminated more than half the JL participants from the BN/AN cluster(s) were defined as difficult items. The items that discriminated some JL participants from the BN/AN cluster(s), but not more than half the JL participants, were defined as learnable items.

3. Results

Before presenting the results of the score for the /r/ tokens and the score for the /l/ tokens, the threshold value of duration for a flap-like sound and the threshold values of F3 will be reported. As described in Sections 2.4 and 2.5, the threshold value of duration was computed to judge some /l/ tokens as /l/ or a flap-like sound. The articulation rate of the BN and AN participants was 4.44 syllables per second and that of the JL participants was 3.43 syllables per second on average, and the threshold was thus defined as 43 ms for the JL participants by multiplying 33 ms (Rimac & Smith, 1984) by 1.29. The 33 ms threshold of duration was applied to the BN/AN participants, and as a result of this threshold, six tokens produced by BN/AN participants were judged as a flap-like sound.

The tokens defined as either /r/ or /l/ were then submitted to the scoring procedure to determine whether /r/ and /l/ were produced as intended using the threshold values of F3 for /r/ and /l/. The results of the BN/AN data showed that the F3 value of initial /r/ at 2 SD and that of initial /l/ at -2 SD were 1665 mel Hz and 1671 mel Hz, respectively, and these values were set as the threshold value of F3 for each approximant. Three /r/ tokens out of the 113 and two /l/ tokens out of the 111 that the BN/AN participants produced were identified as unintended when these thresholds were applied.

Table 3.1 and Figure 3.1 present the descriptive statistics of the variables for the BN, AN and JL groups. Figure 3.1(a and b) illustrates these variables, the average number of the /r/ and /l/ correct tokens out of eight and the average number of tokens for each error type, respectively. In Figure 3.1(a), the items are indicated on the x-axis and the score for the /r/ and /l/ tokens on the y-axis. In Figure 3.1(b), the error types and the average number of errors

are shown on the x-axis and y-axis, respectively. The errors were broadly categorized into three types, as displayed in the figure: the substitution of /l/ for /r/ and vice versa, that of a flap-like sound for /r/ and /l/ and that of a vowel-like sound for /r/ and /l/.
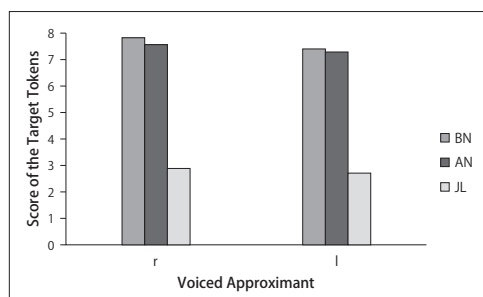
Table 3.1

Descriptive Statistics of /r/ and /l/ for BN, AN and JL Groups

	BN (*n* = 12)				AN (*n* = 7)				JL (*n* = 72)			
	M	*SD*	*Max*	*Min*	*M*	*SD*	*Max*	*Min*	*M*	*SD*	*Max*	*Min*
r	7.83	0.39	8.00	7.00	7.57	0.79	8.00	6.00	2.86	2.62	8.00	0.00
l	7.42	0.67	8.00	6.00	7.29	0.76	8.00	6.00	2.69	2.34	8.00	0.00

Note. For /r/ and /l/, the highest possible value is 8, corresponding to the number of items.

(a) Scores for the /r/ and /l/ tokens
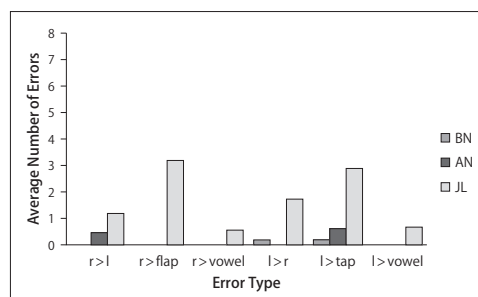
(b) Number of errors

Figure 3.1. **Score for the /r/ and /l/ tokens and average number of errors for BN, AN and JL groups: (a) the score for the /r/ and /l/ tokens; and (b) the number of errors for six error categories. r>l = substitution of /l/ for /r/; r>flap = substitution of a flap-like sound for /r/; r>vowel = substitution of a vowel-like sound for /r/; l>r = substitution of /r/ for /l/; l>flap = substitution of a flap-like sound for /l/; l>vowel = substitution of a vowel-like sound for /l/.**

The BN and AN groups and the JL group primarily differed in the number of correct tokens as in Table 3.1. The JL group achieved lower scores for both /r/ and /l/ (*M* = 2.86, *SD* = 2.62 for /r/; *M* = 2.69, *SD* = 2.34 for /l/) than the BN group (*M* = 7.83, *SD* = 0.39 for /r/; *M* = 7.42, *SD* = 0.67 for /l/) and the AN group (*M* = 7.57, *SD* = 0.79 for /r/; *M* = 7.29, *SD* = 0.76 for

/l/). Figure 3.1(a) displays no substantial difference in the scores between /r/ and /l/ for each group, and also shows how low the scores of the JL group were, compared to those of the BN/AN groups. Figure 3.1(b), furthermore, reveals that the low score of the JL groups were attributed to frequent substitutions of both /r/ and /l/ for a flap-like sound. Substitutions of /r/ for /l/ and vice versa came next.

A cluster analysis was carried out to profile the JL participants, using the z-scores of the scores for the /r/ and /l/ tokens calculated based on the mean and standard deviation of the entire sample. Figure 3.2 shows the dendrogram output by the analysis.
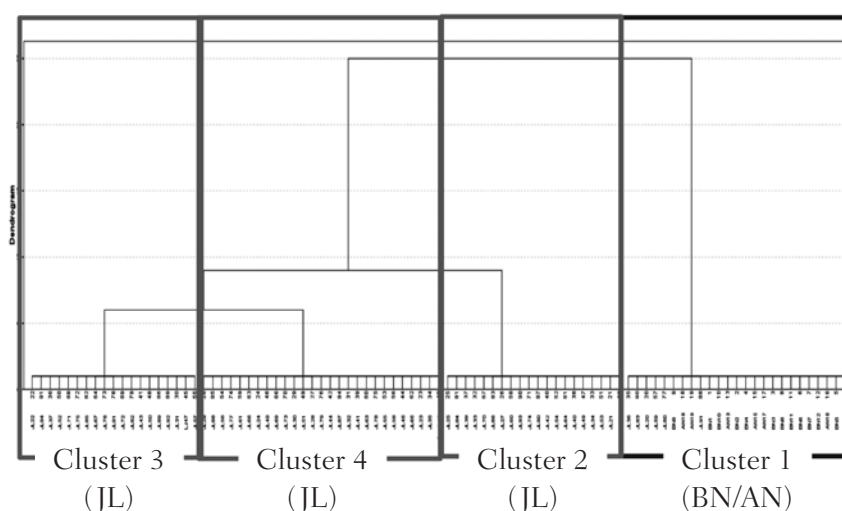

Figure 3.2. **Dendrogram for /r/ and /l/**

All participants were separated into four clusters at the earliest stage of the clustering process. The four clusters were thus selected for the statistical analyses that followed. Cluster 1 consisted of 12 BN participants, 7 AN participants and 6 JL participants. This cluster was considered to represent native speakers, and termed as a BN/AN cluster (Figure 3.2). Clusters 2, 3 and 4 were comprised of 20 JL participants, 19 JL participants and 27 JL participants, respectively, each of which was termed as a JL cluster (Figure 3.2).

Table 3.2 shows the descriptive statistics for the four clusters formed by the cluster analysis, where the valid F3 values averaged across participants were also presented. Some participants whose tokens were judged as neither /r/ nor /l/ failed to provide any valid F3 values; therefore, the number of participants from whom the F3 values were measured is also

given in the table. Figure 3.3(a and b) visually presents the scores for the /r/ and /l/ tokens, and the average number of errors for each error type, respectively. The results are summarized in Figure 3.3(a and b) in the same style as Figure 3.1(a and b).

Table 3.2

Descriptive Statistics of /r/ and /l/ for Four Clusters

	Cluster 1 (n = 25)		Cluster 2 (n = 20)		Cluster 3 (n = 19)		Cluster 4 (n = 27)	
Valid F3 *n*	25 /r/ / 25 /l/		20 /r/ / 18 /l/		13 /r/ / 19 /l/		14 /r/ / 19 /l/	
	M	*SD*	*M*	*SD*	*M*	*SD*	*M*	*SD*
r	7.40	1.15	5.75	1.41	1.21	1.03	1.11	1.22
l	7.28	0.74	1.55	1.23	4.95	1.22	1.00	0.78
F3 Hz [r]	1735.33	156.60	1819.72	102.23	1789.36	161.11	1861.69	156.89
F3 Hz [l]	2546.50	131.71	2562.12	176.40	2497.90	140.04	2565.37	115.93
F3 mel [r]	1489.06	69.40	1491.40	52.81	1460.20	81.78	1487.68	142.50
F3 mel [l]	1824.57	52.31	1830.76	70.51	1784.35	56.42	1833.13	46.98

Note. The number given on the third row shows the number of participants who provided a valid F3 value of /r/ and /l/. For /r/ and /l/, the highest possible value is 8, corresponding to the number of tokens.

The four clusters had different patterns of performance for the production of /r/ and /l/, as clearly shown in Table 3.2 and Figure 3.3(a). The target tokens that the participants in Cluster 1, the BN/AN cluster, produced were most frequently judged as intended for both /r/ and /l/ (*M* = 7.40, *SD* = 1.15 for /r/; *M* = 7.28, *SD* = 0.74 for /l/). In contrast, Cluster 2 obtained a higher score for /r/ than /l/ (*M* = 5.75, *SD* =1.41 for /r/; *M* = 1.55, *SD* =1.23 for /l/), Cluster 3 achieved a higher score for /l/ than /r/ (*M* = 1.21, *SD* =1.03 for /r/; *M* = 4.95, *SD* = 1.22 for /l/) and Cluster 4 performed poorly for both /r/ and /l/ (*M* = 1.11, *SD* =1.22 for /r/; *M* = 1.00, *SD* = 0.78 for /l/). These patterns are reflected in the pattern of errors that the participants in the JL clusters made. As can be seen in Figure 3.3(b), Cluster 2, which performed better in /r/, substituted /r/ for /l/ most frequently, Cluster 3, which performed better in /l/, tended to substitute /l/ for /r/ or a flap-like sound, and Cluster 4, which performed most poorly in both /r/ and /l/, substituted a flap-like sound for /r/ and /l/ more often than the other clusters.

(a) Scores for the /r/ and /l/ tokens
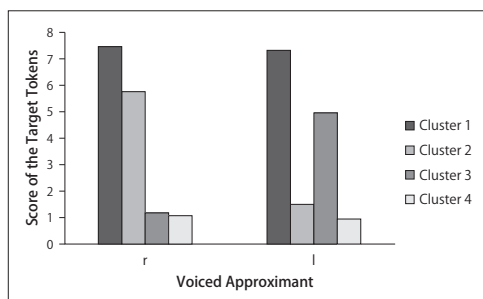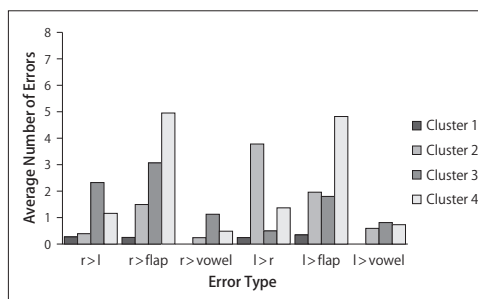
(b) Number of errors

Figure 3.3. Score for the /r/ and /l/ tokens and average number of errors for four clusters: (a) the score for the /r/ and /l/ tokens; and (b) the number of errors for six error categories. r>l = substitution of /l/ for /r/; r>flap = substitution of a flap-like sound for /r/; r>vowel = substitution of a vowel-like sound for /r/; l>r = substitution of /r/ for /l/; l>flap = substitution of a flap-like sound for /l/; l>vowel = substitution of a vowel-like sound for /l/.

A one-way MANOVA was conducted with the scores for the /r/ and /l/ tokens as dependent variables, and the four clusters as the independent variables, in order to determine whether these visually detected differences were significant or not. Table 3.3 shows that two variables, the score for /r/ the tokens and that for the /l/ tokens, were moderately correlated. This suggests that a MANOVA was estimated to work well with these variables, as in Tabachnick and Fidell (2007), who states that the MANOVA does not perform well when the variables have an extremely high or low correlation.

Table 3.3

Correlation between the Variables for /r/ and /l/

Variable	1
1. /r/	—
2. /l/	.48**

** $p < .01$.

The sample size of the largest cluster was less than 1.5 times as large as that of the smallest cluster, so that the α level was set at .05 (Stevens, 2007). Pillai's trace yielded a significant difference among the clusters, $F(6, 174) = 140.91$, $p < .001$, $\eta_p^2 = .83$.

A post-hoc discriminant analysis was performed to identify the differences among the clusters, which was found by the MANOVA. Two discriminant functions were found to discriminate between the clusters. The first function accounted for 76.7% of the variance, canonical $R^2 = .91$, and the second function accounted for 23.3% of the variance, canonical $R^2 = .75$. When combined, these functions significantly differentiated the clusters from each other with the Wilk's lambda value of .02, $\chi^2(6) = 328.64$, $p < .001$. After removing the first function, the second function was able to discriminate between the clusters at a significant level with the Wilk's lambda value of .25, $\chi^2(2) = 120.70$, $p < .001$. These results mean that the differences among the clusters can be explained by these two functions with the first function accounting more for the differences. The group centroids in Table 3.4 and the discriminant plot in Figure 3.4 show that the four clusters were well discriminated by the functions. The first function distinguished Clusters 2, 3 and 4 from Cluster 1. Cluster 1 and Cluster 4 were separated maximally. The second function differentiated between the JL clusters, where Cluster 2 and Cluster 3 were discriminated most.

Table 3.4

Group Centroids for /r/ and /l/

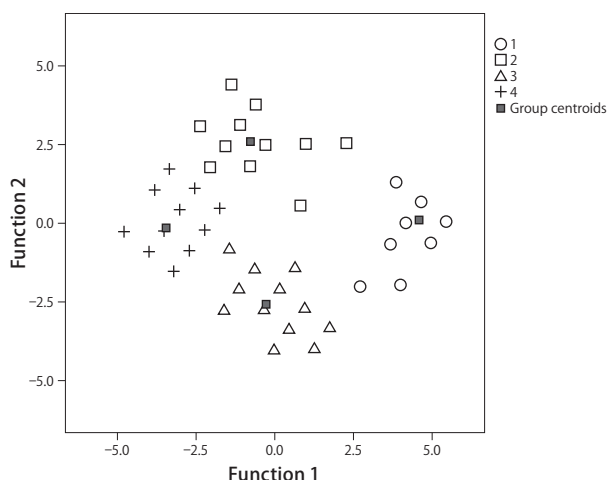| Cluster	Function	
	1	2
1	4.58	0.09
2	−0.78	2.58
3	−0.28	−2.59
4	−3.47	−0.17

Figure 3.4. Canonical discriminant function plot for /r/ and /l/.

Table 3.4 is a structural matrix to show the correlations between the variables and the two functions. The results revealed that the score for the /r/ tokens loaded on the first function most highly ($r = .81$), and that of /l/ also loaded on it ($r = .62$). As noted above, the first function discriminated Cluster 1 from the other clusters, especially between Cluster 1 and Cluster 4. Taken together, the results suggest that both the score for the /r/ tokens and the score for the /l/ tokens contributed to discriminating Clusters 2, 3 and 4 from Cluster 1, Cluster 4 from Cluster 1 in particular, as seen in Table 3.2 and Figure 3.3(a). The participants in Cluster 1 performed best for /r/ and /l/ of all clusters, obtaining the highest scores ($M = 7.40$, $SD = 1.15$ for /r/; $M = 7.28$, $SD = 0.74$ for /l/), whereas those in Cluster 4 performed most poorly, achieving the lowest scores for both /r/ and /l/ ($M = 1.11$, $SD = 1.22$ for /r/; $M = 1.00$, $SD = 0.78$ for /l/). Cluster 2 also achieved the lower scores for both target approximants ($M = 5.75$, $SD = 1.41$ for /r/; $M = 1.55$, $SD = 1.23$ for /l/) than Cluster 1. This was true of Cluster 3 ($M = 1.21$, $SD = 1.03$ for /r/; $M = 4.95$, $SD = 1.22$ for /l/). Thus, none of the JL clusters failed to attain the level of the BN/AN clusters for the production of /r/ and /l/.

Table 3.5

Structural Matrix for the Correlations between the Variables for /r/ and /l/ and the Two Discriminant Functions

	Function	
Variable	1	2
Score for the /r/ tokens	**.81**	**−.59**
Score for the /l/ tokens	**.62**	**.79**

Note. The variables with the absolute value of correlations with the corresponding functions of .33 and above were highlighted in bold.

The second function concerned the discrimination among the JL clusters, as noted earlier. According to the structural matrix in Table 3.5, this function was identified by both the score for the /l/ tokens ($r = .79$) and the score for the /r/ tokens ($r = -.59$). As in the values presented in Table 3.2, Cluster 3 gained the higher score for the /l/ tokens ($M = 4.95$, $SD = 1.22$) than Cluster 2 ($M = 1.55$, $SD = 1.23$). In contrast, Cluster 2 obtained the higher score for the /r/ tokens ($M = 5.75$, $SD = 1.41$) than Cluster 3 ($M = 1.21$, $SD = 1.03$). The participants in Cluster 4 failed to achieve such higher scores than those in the other JL clusters for both approximants ($M = 1.11$, $SD = 1.22$ for /r/; $M = 1.00$, $SD = 0.78$ for /l/) as in Table 3.2. Accordingly, these results demonstrated that the second function particularly highlighted the differences between Cluster 2 and Cluster 3. The average number of errors in Figure 3.3(b) also emphasized these differences among the JL clusters, as described above. Cluster 2, which showed better performance in /r/, produced /r/ even for the /l/ tokens more often than Clusters 3 and 4. Cluster 3, obtaining the higher score for the /l/ tokens, were likely to substitute /l/ for /r/ more frequently than Clusters 2 and 4. Cluster 4 performed more poorly in both /r/ and /l/ than Cluster 2 and Cluster 3, which would be reflected in the most frequent substitution for a Japanese consonant, a flap-like sound.

4. Discussion

4.1 Findings

This study measured F3 values produced by the JL participants and BN/AN participants to examine the difficulty level of learning two English approximants /r/ and /l/. The results showed that a majority of the JL participants produced both consonants less accurately than

the BN/AN participants. At the same time, two patterns of learning were found for the JL participants: one is the pattern that /r/ is learned faster than /l/ and the other is the one that /l/ is learned faster than /r/. The results of the experiment and the definition of the difficulty level will be discussed in more detail below.

To judge the target tokens as intended /r/ or /l/, the threshold values of F3 were defined for /r/ and /l/, respectively, as follows: 1665 mel Hz and 1671 mel Hz. These thresholds of /r/ and /l/ are equal to 2177 Hz and 2185 Hz, respectively, when mel was converted back to Hz. The threshold of /r/ in this study was close to the F3 value of /r/ in Saito and Lyster (2011), who reported the F3 value between 2200 Hz and 2300 Hz. This value defined as the threshold for /r/ in the present study would therefore be reasonable. On the other hand, the threshold of /l/ defined was lower than the value that Saito and Lyster reported as F3 value of /l/, 2800 Hz.

Iverson and Kuhl (1996) also showed the F3 values for /r/ and /l/, which helped discuss whether the thresholds of this study were reasonable or not. They investigated the perceptual similarity underlying /r/, /l/ and /w/ within the framework of the native language magnet model, and found that an F3 value of the best exemplar for /r/ was 1473 Hz, and that for /l/ was 3478 Hz for one group, and 3329 Hz for the other. Compared with these values, the threshold values /r/ and /l/ in the present study seemed higher and lower, respectively. As Iverson and Kuhl noted, however, the best exemplars tended to be more extreme. The stimuli were created by synthesizing female speech, and were tokens that constituted a simple syllable structure, CV, in Iverson and Kuhl. In contrast, the F3 values that this study set were not F3 values for good exemplars or averages, but thresholds. The values were defined based on the speech sample collected from male speakers, and a passage was used to collect the data. In addition, the phonetic boundary of F3 between /r/ and /l/, found by Iverson and Kuhl, was somewhere between 2067 Hz and 2523 Hz. The thresholds of this study fell within this range. Taken together, the thresholds of F3 defined in the present study would be acceptable.

The cluster analysis was carried out using the variables obtained based on these yardsticks, and it generated one BN/AN cluster, consisting of 25 participants. All BN/AN participants were classified into this cluster. The fact that only six JL participants were grouped into the same cluster as the BN/AN participants suggests the overall difficulty of learning /r/ and /l/.

The cluster analysis also formed three JL clusters, each of which comprised of 20 JL participants, 19 JL participants and 27 JL participants. All these clusters were discriminated from the BN/AN cluster in production of both /r/ and /l/. The results suggest that /r/ and /l/ were difficult items for the JL participants to learn to produce according to the criteria

described in Section 2.7 because more than half the JL participants were differentiated from the BN/AN cluster.

At the same time, it was found that there was a difference in the performances of both /r/ and /l/ among the JL participants, although none of the JL clusters reached the level of BN/AN cluster. This points to some potential for learning by Japanese learners of English. The major difference was that the JL cluster of 20 participants performed better in producing /r/, that the JL cluster of 19 participants performed better in /l/ and that the JL cluster of 27 participants performed poorly on both /r/ and /l/. This pattern was further supported by the results of the pattern of the errors that they made. The JL cluster of 20 participants produced /r/ even for the /l/ tokens, and /r/ was thus an easier item than /l/ for them to learn. In contrast, the JL cluster of 19 participants tended to substitute /l/ for the /r/ tokens more often, which suggests that /l/ was an easier item than /r/ for them to learn, unlike the JL cluster of 20 participants. The cluster of 27 JL participants were most likely to replace both /r/ and /l/ tokens with a flap-like sound of all clusters. This indicates that they did not have any preference for learning /r/ or /l/, but rather had been learning neither /r/ nor /l/. The statistical tests including a MANOVA and a discriminant analysis confirmed these results that some of the JL participants had been learning /r/ or /l/, while the others had not. However, evidence of learning these approximants was only found in less than half of the JL participants. Accordingly, both target approximants were defined as difficult items.

4.2 Definition of the difficulty level

It was hypothesized that /r/ and /l/ would be learnable and difficult for Japanese learners of English, respectively. According to the results, more than half of the JL participants significantly differed from the BN/AN participants in production of both target approximants, and fewer than half of the JL participants improved their articulation of /r/ and /l/. Twenty-seven JL participants, nearly half the JL participants, even performed poorly for both /r/ and /l/, and showed a high frequency of substitution of a flap-like sound for /r/ and /l/. Both approximants were thus identified as difficult items for Japanese learners of English to learn to produce, which rejected the hypothesis of /r/ and upheld that of /l/.

The present research formed the hypothesis for /r/ based on the findings of Aoyama et al. (2004) and Hazan et al. (2005) that /r/ could be learned faster than /l/. One of the possible reasons that this study failed to support this would be related to the proficiency level of the participants. Previous studies have found the difficulty for less experienced learners to attain

the level of native speakers in production of /r/ and /l/ (Flege, Takagi et al., 1995; Goto, 1971), and have suggested a different status of /r/ and /l/ in the L2 phonological space of Japanese learners of English. Considering that the participants in this study were less experienced learners, who had no experience of living in an English-speaking country, the findings here were consistent with these past studies. Yamada (1995) pointed out that the experience of living in the U.S. could affect the perception of these approximants. Saito and Lyster (2011) reported that with training, Japanese learners of English could improve the production of /r/.

It was also found that there were /r/ preference and /l/ preference for learning, which was unexpected. The participants in one JL cluster showed that they were going through the learning process of /r/, while those in another JL cluster were learning /l/. This gives an indication that there are individual differences in the way of learning. It should be emphasized that some JL participants were learning /l/ faster than /r/, in particular. This is against the prediction of one of the learning models for production, the Speech Learning Model (SLM; Flege, 1987, 1995), which proposes that the newer L2 phones are easier for learners to learn while the more similar L2 phones to an L1 phone are more difficult. Guion et al. (2000) claimed that the difference between English /r/ and Japanese /ɾ/ would be more salient than that between English /l/ and Japanese /ɾ/, although no Japanese sound was perceptually similar to /r/ and /l/. From the articulatory perspective, /r/ could also be assumed to be newer than /l/. The major phonetic features of articulating English /r/, such as retracted tongue or lip rounding, are not prominently used in Japanese. On the other hand, English /l/ and Japanese /ɾ/ differ in that the former is continuant and the latter is not, but both require the tip of tongue as an active articulator and the alveolar ridge, or around this region, as a passive articulator. Thus, the newness is higher for /r/. When the SLM is applied here, the higher degree of salience of /r/ can facilitate Japanese learners learning this approximant. However, this did not hold true of this study that demonstrated that some learners were learning /l/ faster. Thus, the findings in this study suggest the need to test one of the most influential learning models, the SLM, although it should be taken into account that the participants in the present study were not so experienced as the SLM requires learners to be experienced for its prediction.

5. Conclusion

The present study aimed to define /r/ and /l/, notoriously difficult for Japanese learners of English to learn to produce, as easy, learnable or difficult items. The results showed that

both approximants were difficult. However, it was also found that there was an individual preference about which approximant was learned faster. Some Japanese learners of English were learning /r/ faster than /l/ whereas others were learning /l/ faster than /r/. Further studies will be required to explore from what these individual differences arise, and how positively the preference for learning /r/ or /l/ will affect the learning of both approximants.

References

Aoyama, K., Flege, J. E., Guion, S. G., Akahane-Yamada, R., & Yamada, T. (2004). Perceived phonetic dissimilarity and L2 speech learning: The case of Japanese /r/ and English /l/ and /r/. *Journal of Phonetics, 32*(2), 233-250.

Boersma, P., & Weenink, D. (2010). Praat: doing phonetics by computer (Version 5.2) [Computer program]. Retrieved from http://www.praat.org/

Boersma, P., & Weenink, D. (2015). Praat: doing phonetics by computer (Version 6.0.05) [Computer program]. Retrieved from http://www.praat.org/

Espy-Wilson, C. Y. (1992). Acoustic measures for linguistic features distinguishing the semivowels /w j r l/ in American English. *Journal of the Acoustical Society of America, 92*(2), 736-757.

Fant, G. (1968). Analysis and synthesis of speech processes. In B. Malmberg (Ed.), *Manual of phonetics* (pp. 173-177). Amsterdam, the Netherlands: North-Holland.

Flege, J. E. (1987). The production of "new" and "similar" phones in a foreign language: Evidence for the effect of equivalent classification. *Journal of Phonetics, 15*(1), 47-65.

Flege, J. E. (1995). Second language speech learning: Theory, findings, and problems. In W. Strange (Ed.), *Speech perception and linguistic experience: Issues in cross-language research* (pp. 233-277). Timonium, MD: York Press.

Flege, J. E., Takagi, N., & Mann, V. (1995). Japanese adults can learn to produce English /ɹ/ and /l/ accurately. *Language and Speech, 38*(1), 25-55.

Goto, H. (1971). Auditory perception by normal Japanese adults of the sounds "L" and "R." *Neuropsychologia, 9,* 317-323.

Guion, S. G., Flege, J. E., Akahane-Yamada, R., & Pruitt, J. C. (2000). An investigation of current models of second language speech perception: The case of Japanese adults' perception of English consonants. *Journal of the Acoustical Society of America, 107*(5), 2711-2724.

Hallé, P. A., Best, C. T., & Levitt, A. (1999). Phonetic vs. phonological influences on French listeners' perception of American English approximants. *Journal of Phonetics, 27*(3), 281-306.

Hazan, V., Sennema, A., Iba, M., & Faulkner, A. (2005). Effect of audiovisual perceptual training on the perception and production of consonants by Japanese learners of English. *Speech Communication, 47,* 360-378.

Iverson, P., & Kuhl, P. K. (1996). Influence of phonetic identification and category goodness on American listeners perception of /r/ and /l/. *Journal of the Acoustical Society of America, 99*(2), 1130-1140.

Iverson, P., Kuhl, P. K., Akahane-Yamada, R., Diesch, E., Tohkura, Y., Kettermann, A., & Siebert, C. (2001). A perceptual interference account of acquisition difficulties for non-native phonemes. *Speech, Hearing and Language: Work in progress, 13,* 106-118.

de Jong, N. H., & Wempe, T. (2009). Praat script to detect syllable nuclei and measure speech rate automatically. *Behavior Research Methods, 41*(2), 385-390.

Kent, R. D., & Read, C. (2002). *Acoustic analysis of speech* (2nd ed.). Clifton Park, NY: Delmar.

Ladefoged, P. (2003). *Phonetic data analysis: An introduction to field work and instrumental techniques.* Malden, MA: Blackwell.

Markham, D., & Hazan, V. (2002). The UCL speaker database. *Speech, Hearing and Language: Work in Progress, 14*, 1-17.

Merfert, I. (1997). The alt.usage.english Audio Archive. Retrieved from http://alt-usage-english.org/index.shtml

Munro, M. J., & Derwing, T. M. (1995). Processing time, accent, and comprehensibility in the perception of native and foreign-accented speech. *Language and Speech, 38*, 289-306.

Rimac, R., & Smith, B. L. (1984). Acoustic characteristics of flap productions by American English-speaking children and adults: Implications concerning the development of speech motor control. *Journal of Phonetics, 12*(4), 387-396.

Saito, K., & Lyster, R. (2011). Effects of form-focused instruction and corrective feedback on L2 pronunciation development of /ɹ/ by Japanese learners of English. *Language Learning, 62*(2), 595-633.

Slawinski, E. (1999). Acquisition of /r-l/ phonemic contrast by Japanese children and adults. *Psycholinguistics on the Threshold of the Year 2000*, 583-590.

Stevens, J. P. (2007). *Intermediate statistics: A modern approach* (3rd ed.). New York, NY. Routledge.

Tabachnick, B. G., & Fidell, L. S. (2007). *Using multivariate statistics* (5th ed.). Boston, MA: Pearson Education.

Yamada, R. (1995). Age and acquisition of second language speech sounds perception of American English /ɹ/ and /l/ by native speakers of Japanese. In W. Strange (Ed.), *Speech perception and linguistic experience: Issues in cross-language research* (pp. 305-320). Timonium, MD: York Press.

APROXIMACIÓN A UN CURSO DE ESCRITURA CREATIVA

Beatriz Prieto Muñoz

ABSTRACT

When students hear the phrase "writing class", they usually anticipate boring classes with a great deal of solitary work. It is possible, however, to have fun and communicative "writing classes" that are much more in tune with the new educational times.

Based on relevant theoretical approaches about writing and creativity -and also taking into account the questionnaires that students filled out before the beginning of the course regarding their own interests- pupils exchange ideas, participate together in the creative process, and then share the final results in a personal blog that shows their work on both group and individual levels.

An evaluation system where students have to find and correct their own mistakes on their own is also applied, making the learning process much more meaningful.

In this article, therefore, we will analyze the creation process of a Creative Writing course in each of its stages: theory, class development and conclusion. The most popular activities will be shared as will the evaluation method and the technological tools that this relies on. Additionally, we will present the results of the final questionnaire, along with students' opinions about the course.

Keywords: Writing, creativity, activities, groups, fun.

1. INTRODUCCIÓN

Cuando a la autora de este artículo le asignaron un curso de Escritura Creativa en varios centros educativos, descubrió que no había muchos manuales sobre el tema y que los pocos que había no se ajustaban a los objetivos que se había propuesto para su clase, por lo que decidió crear el curso desde cero, basándose en un plan de organización que cubría, entre

otras, las siguientes áreas:

1. Investigar sobre la adquisición de la destreza escrita y factores que la favorecían.
2. Búsqueda de materiales para el curso: libros de texto, artículos de revistas, Facebook, materiales originales, etc.
3. Realización de una serie de encuestas al principio del curso sobre los temas de más interés entre los estudiantes así como sus expectativas, y al final del curso, para comprobar los resultados.

La palabra "escritura" no suele causar gran entusiasmo entre los estudiantes, ya que anticipan clases poco entretenidas y mucho trabajo dentro y fuera de la clase así que los objetivos del curso eran simples aunque no sencillos: conseguir que los estudiantes disfrutaran escribiendo en español y que pusieran en práctica lo aprendido a lo largo del curso anterior. De igual manera, como objetivo secundario aunque no por ello menos importante, se buscaba potenciar su creatividad e imaginación y, muy especialmente, su capacidad para generar ideas, algo que, todos los profesores de escritura coinciden en afirmar, es una de las partes que más dificultad entraña.

Se tratará también la forma de evaluación elegida para este curso, que tuvo que adaptarse a las diferencias en cuanto a facilidades y acceso a la tecnología de cada centro. En este artículo, pues, se describirá el proceso abierto de creación de un curso de escritura creativa en las tres fases de inicio, desarrollo y conclusión, se describirán actividades que resultaron exitosas en la clase y se compartirán los resultados de las encuestas realizadas durante el mismo.

2. FASE INICIAL
2.1 MARCO TEÓRICO

Antes de analizar la combinación de escritura y creatividad, decidimos profundizar primero en cada concepto por separado.

2.1.1 ESCRITURA. El origen de la palabra "escribir" lo encontramos en el vocablo latino *scribere*, el cual, según la primera acepción de la Real Academia Española, podemos definir en la actualidad como "representar las palabras o las ideas con letras u otros signos trazados en papel u otra superficie".

En la enseñanza de segundas lenguas, el papel de la escritura ha pasado por diferentes etapas: desde el enfoque tradicional, en el cual la escritura se utilizaba básicamente para hacer ejercicios de gramática y traducciones, pasando por ser un refuerzo a la oralidad en el método audio-lingual, hasta llegar a épocas más modernas donde se empezó a considerar un instrumento de comunicación en toda regla. (García Parejo, 1999: 25-27)

Es en estas últimas etapas donde se comienza a ver la destreza escrita como un proceso, no solo como un producto, y donde se incluyen tareas para trabajar las etapas intermedias del proceso de escritura: borradores, esquemas, etc. con el objetivo de dotar al estudiante de estrategias para escribir de forma creativa en función de sus necesidades comunicativas. El texto final, pues, es el resultado de un proceso en el que se planifica, se negocia con el profesor, se discute con otras personas y que se deja abierto a modificaciones posteriores, si fuera necesario (García Parejo, 1999: 33-34).

En esta misma línea, el Marco Común Europeo de Referencia define la práctica escrita como un instrumento que propicia la interacción comunicativa, aportando ejemplos de escritura reales que los aprendientes pueden necesitar en su vida cotidiana incluso en etapas muy básicas de su estudio, como por ejemplo, un formulario. (Consejo de Europa, 2002: 64)

2.1.2 CREATIVIDAD. Una vez recabada información sobre la adquisición de la destreza escrita, nos dispusimos a investigar sobre la creatividad.

El origen etimológico de la palabra "creatividad" se encuentra en el vocablo latino *creare*, que significa crear, hacer algo nuevo. En realidad lo que llegamos a calificar de creativo sería aquellas formas nuevas a partir de otras ya creadas a las que damos una nueva utilización (Vázquez, 2000).

En *La Estructura del Intelecto*, Guilford despertó un interés científico por el estudio de la creatividad, considerándola como una actividad intelectual que forma parte de lo que denominó "pensamiento divergente". Este pensamiento se define como aquel que, ante un problema específico, puede formular varias respuestas alternativas, en oposición a lo que sería el "pensamiento convergente", que ocurre cuando solo es posible una solución determinada. De esta manera, el pensamiento divergente implica utilizar el conocimiento previo de formas nuevas, y el convergente está relacionado con el pensamiento base, la reproducción y memorización de los aprendizajes y hechos. (Jiménez González, 2007)

A partir de los años 70 y el resurgir del paradigma cognitivo, el sujeto creativo se empezó a ver como un procesador activo de información, poseedor de una mente capaz de tener

actividad autónoma, y se reconoce que la mente dispone de representaciones de la realidad variadas, que se pueden entender como "módulos mentales". De esta forma, el pensamiento convergente sería el proceso utilizado para solucionar los problemas mediante procedimientos convencionales y predeterminados. En contraposición, el pensamiento divergente sería la operación que implica la producción de distintas respuestas o soluciones para un determinado problema. Así pues, la producción divergente reúne las cuatro características de la creatividad o factores propuestos por Guilford: fluidez, flexibilidad, originalidad y elaboración. Estos cuatro componentes que configuran el pensamiento creativo fueron definidos de la siguiente forma: a) fluidez: es la característica de la creatividad o la facilidad para generar un número elevado de ideas. Es decir, se trata de una habilidad que consiste en producir un número elevado de respuestas en un campo determinado, a partir de estímulos verbales o figurativos; b) flexibilidad: es la característica de la creatividad mediante la cual se transforma el proceso para alcanzar la solución del problema o el planteamiento de éste. Comprende una transformación, un cambio, un replanteamiento o reinterpretación. En definitiva, es la capacidad consistente en producir diferentes ideas para cambiar de un enfoque de pensamiento a otro y para utilizar diferentes estrategias de resolución de problemas; c) originalidad: es la característica que define a la idea, proceso o producto como algo único o diferente. Está referida a la habilidad para producir respuestas novedosas, poco convencionales, lejos de lo establecido y usual; d) elaboración: es el nivel de detalle, desarrollo o complejidad de las ideas creativas. Se trata de una capacidad para desarrollar, completar o embellecer una respuesta determinada. (Jiménez González, 2007)

Asimismo, Torrance (1974), creador del test de pensamiento creativo, definió la creatividad como el proceso de descubrir problemas o lagunas de información, formar ideas o hipótesis, probarlas, modificarlas y comunicar los resultados y le asignó a la creatividad un carácter de habilidad global.

2.1.3 ESCRITURA CREATIVA

Aunque existen muchas definiciones de escritura creativa, elegimos la de García Carcedo (2011): "La escritura creativa es aquella, de ficción o no, que desborda los límites de la escritura profesional, periodística, académica y técnica. Esta categoría de escritura incluye la literatura y sus géneros y subgéneros, en especial, la novela, el cuento y la poesía, así como la escritura dramática para el teatro, el cine o la televisión. En este modo de escritura prima la creatividad sobre el propósito informativo propio de la escritura no literaria."

Y la de Betancour (2000):

"Educar en la creatividad es educar para el cambio y formar personas ricas en originalidad, flexibilidad, visión, iniciativa, confianza; personas amantes de los riesgos y listas para afrontar los obstáculos y problemas que se les van presentando en su vida, tanto escolar y cotidiana. Además, educar en la creatividad es ofrecer herramientas para la innovación."

Seguidamente, encontramos también seis estrategias propuestas por Hal y Sweet (2008) para enseñar escritura creativa:

1. La estrategia del atelier, donde el maestro acepta a un discípulo y le enseña todas sus técnicas de escritor.

2. La estrategia del *imitatio auctoris*, donde el estudiante se empapa de las técnicas, formas y contenido de los clásicos y copia sus modelos de escritura.

3. La estrategia de la inspiración. Con el objetivo de activar la creatividad de los estudiantes, los profesores plantean una serie de ejercicios para vencer el bloqueo frente a la hoja en blanco.

4. La estrategia de la Techné, está centrada en el profesor, que, utilizando los preceptos y técnicas de grandes escritores para ilustrar conceptos, espera que los estudiantes aprehendan esa técnica y sean capaces de utilizarla en sus propios textos.

5. La estrategia del taller literario, en la que un grupo de aprendices son guiados por un solo maestro.

6. La comunidad de escritores, propone hacer trabajo colaborativo en pequeños grupos de tres o cuatro estudiantes con intereses literarios comunes que son asesorados por un profesor.

Nos pareció que la estrategia número seis era la que más se adecuaba a nuestro curso, ya que nuestro objetivo era que los estudiantes produjeran sus propias ideas, aunque proporcionáramos algunos modelos que les ayudaran cuando se produjera el bloqueo mencionado en la estrategia 3.

Así pues, teniendo en cuenta lo aprendido en nuestra investigación teórica, la preparación del curso de escritura debía cumplir unas condiciones básicas:

1. Trabajar el proceso de escritura tanto en la fase creativa como en la escrita, para lo cual se diseñaron actividades en las que los estudiantes tenían que trabajar en grupos

en alguna de las dos fases o incluso en las dos, creándose así contextos de interacción entre compañeros (Cassany, 1999: 20) y aplicando así la estrategia número seis ya mencionada: "Comunidad de escritores".

2. Crear el ambiente y las actividades adecuadas que permitieran al estudiante escribir de forma libre para desarrollar su pensamiento divergente, tratando de que el producto final reuniera fluidez, flexibilidad, originalidad y elaboración, las cuatro características de la creatividad.

3. Finalmente, nos propusimos que los estudiantes simplemente disfrutaran escribiendo o, citando a Concha Moreno: reivindicar el placer de escribir por escribir. Para ello, pensamos que era de vital importancia que los estudiantes escribieran sobre temas que les interesaran, por lo que el primer día de clase se realizaron unas encuestas en las que se les preguntó, simplemente, qué temas les interesaban, y cuyo resultado dilucidó que su vida y sus aficiones eran sus temas favoritos. También se trataron de incluir temas no tan populares pero que consideramos serían de utilidad en su vida profesional, por ejemplo, cómo escribir un curriculum vitae en español.

2.2 DESCRIPCIÓN DE LOS GRUPOS.

El curso de Escritura Creativa se impartió en los años 2013 y 2014 en varios centros educativos a nivel universitario y, por tanto, a estudiantes de entre 18 y 22 años. Cada universidad cuenta con un número de estudiantes diferente en clase, así como programas y metodologías diversos, pero las clases tenían entre 10 y 18 estudiantes.

Las encuestas se repartieron en tres grupos en concreto, de 10, 12 y 18 estudiantes, en total, 40 personas, de las cuales 23 eran mujeres y 17, hombres.

Asimismo, hemos trabajado con alumnos que estudiaban español como asignatura obligatoria y con otros que la tenían como asignatura opcional pero no hemos observado ninguna diferencia reseñable en los resultados de las encuestas, por lo que no hemos reflejado esta distinción en los resultados. En cuanto al nivel, en todos los casos correspondía a un nivel intermedio, alcanzado tras estudiar dos semestres de forma intensiva con 4 clases semanales de una hora y media.

Con respecto a la metodología que se ha usado en las clases de español, no se siguió ningún libro en particular, ya que se creó material original específico para esta clase. Además, los estudiantes tenían clases de gramática impartidas por un profesor japonés mientras que las clases de comunicación, tanto oral como escrita, estuvieron a cargo de un profesor nativo

hispanohablante.

Casi todos los estudiantes de estos cursos habían viajado a algún país de habla hispana como viaje de estudios con una estancia de entre 1 y 3 meses, lo que ayudó a que la mayor parte de ellos se encontrara bastante predispuesto y motivado de cara a estas clases.

2.3 MATERIALES

Para este curso se utilizaron actividades modificadas de varios libros de texto entre los que se incluyen: *Expresión Oral, Escritura Creativa, Nuevo ELE actual B1, Rápido rápido, Gente 2*, etc.

De ellos se tomaron ideas para realizar actividades referentes a cómo escribir un curriculum vitae, hacer una entrevista a su compañero o crear las leyes de un país imaginario, entre otros ejemplos.

De igual manera, se utilizó material auténtico tomado de artículos de revistas, blogs y Facebook, convenientemente modificados para adaptarlos al nivel de los estudiantes, claro está. Un ejemplo de ello es la actividad "Cosas que echas de menos de Japón".

La autora decidió también diseñar material original para recrear ideas que no aparecían en ninguna de las otras opciones y que incluían referencias culturales japonesas. Un ejemplo es la actividad "Un día normal en el mundo".

En cuanto a soportes digitales, se contempló también la posibilidad de crear un grupo en Facebook donde publicar lo que se iba escribiendo en clase pero se descartó la idea por la pérdida de control de privacidad y por la exposición a material auténtico no modificado que sobrepasaba el nivel de los estudiantes. En su lugar, se eligió la creación de un blog como opción que mejor se adaptaba a los objetivos del curso.

Uno de los centros donde se impartió este curso proporcionaba a los estudiantes un ordenador portátil que ellos siempre llevaban a las clases, por lo cual se les dio la opción de utilizar los últimos minutos de clase para escribir en el blog. Los otros grupos tuvieron que realizar esta tarea como deberes en su casa.

En el blog debían escribir todo el trabajo realizado en clase, una vez que éste había sido revisado por la profesora y modificado por los estudiantes, y se les indicó también que se valoraría el aspecto visual del mismo, como prueba de su esfuerzo y dedicación.

La profesora creó también su propio blog para que los estudiantes tuviesen una referencia pero decidió incluir solamente una entrada de presentación para dar más libertad creativa a los estudiantes.

3. DESARROLLO

En el seminario organizado por GIDE sobre Expresión escrita en junio de 2013 la profesora Concha Moreno empezó con una imagen del kanji japonés de escritura para hacernos reflexionar sobre su significado. Ese interesante comienzo, así como otras ideas del taller y de sus asistentes, resultaron de valiosísima ayuda para preparar el curso de escritura que nos ocupa e hicieron que en la primera clase también se comenzara con una pregunta de reflexión: "¿Qué escribimos?", siguiendo el modelo de Moreno et al. (1999:78). Para ayudarles a responder esa pregunta, se les pidió primero que hicieran una lista de todas las cosas que habían escrito en los últimos días: un mensaje en el móvil, una entrada en Facebook, un examen, etc. y se puso en común en la pizarra.

El siguiente paso a esta actividad fue hacer otra lista de cosas que les gustaría escribir o aprender a escribir, que conformó la base para preparar la mayor parte del curso y finalmente se les indicó que escribieran qué temas les resultaban más interesantes, donde las opciones más populares fueron hablar de su mundo: su vida, sus aficiones, viajes, etc.

Con este primer feedback en mente, revisamos y modificamos el plan del curso para ajustarlo a estos nuevos criterios.

3.1 ACTIVIDADES

En las actividades realizadas en este curso se trataron los siguientes géneros discursivos: entrevista, historia, entrada en un foro, lista justificada, test, carta informal, diario de viajes, biografía, leyes/normas, cuento, artículo periodístico, poesía, curriculum vitae; además pueden clasificarse en varias categorías:

1. Actividades donde se trabaja la generación de ideas en grupo pero se escribe individualmente.
2. Actividades donde se hace el proceso creativo de forma individual pero para escribir el texto se necesita la colaboración de los compañeros.
3. Actividades en las que tanto el proceso creativo como el escrito se hace en grupo.
4. Actividades en las que se da un modelo previo que se trabaja en gran grupo pero tanto el proceso creativo como el escrito se hace de forma individual.
5. Combinación de dos de las categorías anteriores.

En la mayor parte de las actividades se trabajó con los compañeros en alguna parte

del proceso, ya que el trabajo colectivo siempre suele producir mejores resultados que el individual, como demuestra la actividad "Supervivencia en la luna", incluida en la bibliografía, que la autora tuvo oportunidad de realizar en un seminario de formación de profesores.

En otras actividades, no solo se realizó alguna parte en grupo, sino la actividad completa, que luego cada estudiante podía modificar a su gusto y necesidad en la fase final.

A continuación pasamos a explicar la dinámica de las actividades más populares.

3.1.1 Un día normal en el mundo. Esta actividad, basada en una idea original de la autora, consistía en imaginar un viaje por el mundo en un solo día.

Vimos un gran potencial en este concepto, ya que los alumnos habían expresado su interés en los viajes en la encuesta inicial, por lo que decidimos escribir un texto original narrando un día normal de la autora en el mundo, como referencia para los estudiantes. Introdujimos la actividad con un componente cultural, preguntando a los estudiantes qué objeto mágico de Doraemon les gustaría tener, y esperando que mencionaran el famoso どこでもドア o "puerta a cualquier sitio". Desde ahí les propusimos que pensaran a dónde viajarían en un día si tuvieran ese objeto mágico y qué harían en el sitio elegido. Se propuso un número inicial de 10 lugares diferentes pero dependiendo del tiempo de clase o del ritmo de trabajo de los estudiantes, se modificó en consecuencia. Una vez pensados los 10 lugares y lo que harían allí, se les pidió que escribieran una memoria de su viaje, siguiendo el modelo que se les había proporcionado y del cual adjuntamos un par de párrafos a modo de ejemplo:

11:00 Me voy a tomar el sol a la playa de Bora Bora. Hace mucho calor pero es muy agradable sentir la brisa del mar mientras me tomo un zumo de mango y papaya. Después del zumo, me doy un baño laaaargo en el mar.

21:30 Cena con velas en París, a la orilla del río Sena. No soy fan de la comida francesa, pero me encanta el ambiente romántico de la noche parisina. ¿Con quién ceno? Eso es un secreto... ^^

Casi todos los estudiantes se esforzaron en elegir lugares repartidos por todo el mundo, destacando un estudiante que incluso viajó a la Luna.

3.1.2 Citas rápidas. Basada en una actividad que realizaron los profesores Mercedes Castro y Javier Fernández para el Instituto Cervantes de Tokio, el objetivo es que cada estudiante encuentre a su "pareja ideal".

Para ello, tras confirmar el significado de "citas rápidas" y preguntarles por su experiencia,

que en Japón podría equivaler al 合コン, se comienza dividiendo la clase en igual número de chicos y chicas. Si hay más de uno que de otro grupo, se pide que algunos estudiantes hagan el papel del sexo contrario, para equilibrar.

Conseguido el primer objetivo, se entrega a los estudiantes una foto boca abajo, es decir, que no pueden elegir, y se les pide que completen una ficha, que también entregaremos, con los siguientes campos: nombre y apellidos, nacionalidad, profesión, edad, idiomas que habla, gustos y qué tipo de persona busca.

Una vez completadas las fichas, se sienta a los estudiantes en parejas de sexo contrario y se les indica que disponen de 5 minutos para hablar con esa persona y escribir la información que consideren relevante o que les resulte de interés.

Transcurridos los 5 minutos, se cambia de pareja y cuando han terminado de hablar con todas las personas, ponen en orden sus notas y eligen a la persona que más les haya gustado o con la que piensen que encajarían mejor. Acto seguido, se pone en común para ver si la persona que han elegido también les ha elegido a ellos, es decir, si han conseguido formar "pareja". Este es, sin duda, el punto álgido de la actividad.

La tarea final consiste en escribir una entrada de foro con su información personal, que imaginaron en la primera tarea de la actividad, y una breve descripción de la persona que habían elegido como pareja ideal, indicando si su "amor" ha sido correspondido o no y cómo se sienten al respecto. Cuantos menos estudiantes haya en la clase, más fácil es que se forme alguna pareja, pero como promedio, casi siempre se forman dos o tres.

3.1.3 Carta informal sobre una fiesta hispana. Actividad original donde el protagonista vuelve a ser un viaje, aunque esta vez con más detalle y con un componente cultural.

Tras una lluvia de ideas para ver qué fiestas hispanas conocen, se les pide que elijan una y se agrupan con las personas que también han elegido esa fiesta. No importa si no la conocen bien, ya que parte de la actividad consiste en investigar sobre ella.

En grupos, buscan información (dónde, cuándo, en qué consiste, etc.) y diseñan un plan de viaje para experimentar cómo es la fiesta: cuántos días van, qué hacen allí, alguna anécdota especial, etc. Para esta fase de la actividad, se les permite el uso de sus dispositivos móviles y, si preocupa que se distraigan, se puede limitar a uno por grupo.

Tras revisar juntos el formato de una carta informal, del que se les han entregado un modelo, escriben una carta a una persona de otro grupo explicándole cómo fue su viaje y su experiencia con esa fiesta, pidiéndoles que presten atención a los tiempos del pasado, y

repasándolos si es necesario.

Ejemplos de fiestas elegidas por los estudiantes incluyeron las Fallas, la tomatina, el día de los Muertos, etc.

3.1.4 Cosas que echas de menos. Actividad basada en el artículo de Raquel Piñeiro: *22 cosas que echas de menos de España ahora que no vives aquí.*

Como actividad previa, se divide a los estudiantes en grupos y tratan de adivinar cinco cosas que echan de menos los españoles cuando no están en España, por ejemplo, el sol, el teclado del ordenador, etc. que forma parte del artículo.

Una vez que han entendido la actividad y el concepto de "echar de menos", se les pide que, también en grupos, hagan una lista de cosas que ellos echan (o echarían de menos) cuando no están en Japón.

Esa primera lista se pone en común en gran grupo y, ya de forma individual, tienen que elegir 10 cosas y escribir su lista particular, justificándola.

Es importante limitar el campo "comida" a una sola cosa. De lo contrario, para muchos de los estudiantes la lista estaría compuesta íntegramente por productos alimenticios...

Algunos ejemplos de cosas que los estudiantes incluyeron en sus listas fueron las tarjetas electrónicas como Suica o Pasmo, las tiendas de 24 horas, la costumbre de recibir agua o té gratis en los restaurantes o el modernísimo toilet japonés.

3.1.5 Cuento en versión moderna. Actividad inspirada en varios libros de texto y páginas web, consistente en reescribir un cuento clásico en una versión moderna.

Debido a la complejidad creativa de este ejercicio, se trabajó en grupos durante la totalidad de la clase. Como actividad previa, se visionó un vídeo de una moderna Cenicienta que encuentra a su príncipe en una red social, citándose después en una discoteca y se les proporcionaron expresiones relacionadas con los cuentos: "érase una vez", "colorín colorado, este cuento se ha acabado", etc.

Aunque el vídeo pareció interesarles mucho y enseguida se pusieron a trabajar, les costó bastante empezar a generar ideas, o por lo menos, ideas aceptadas por todo el grupo.
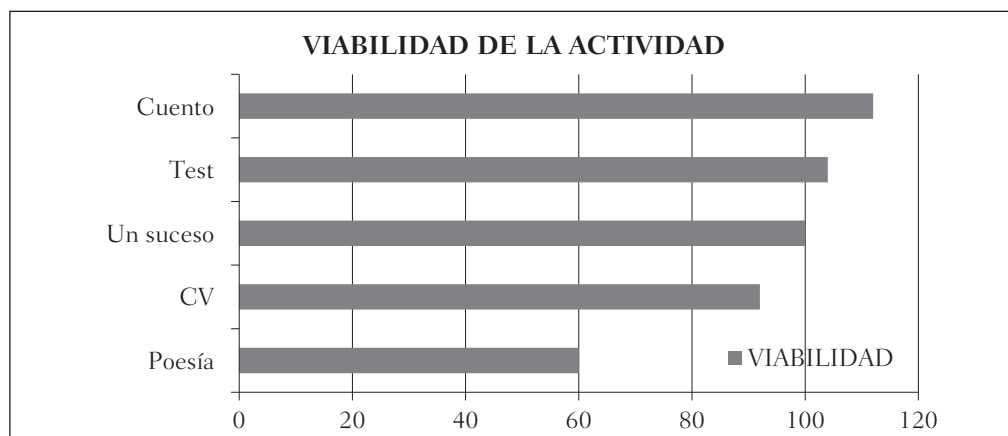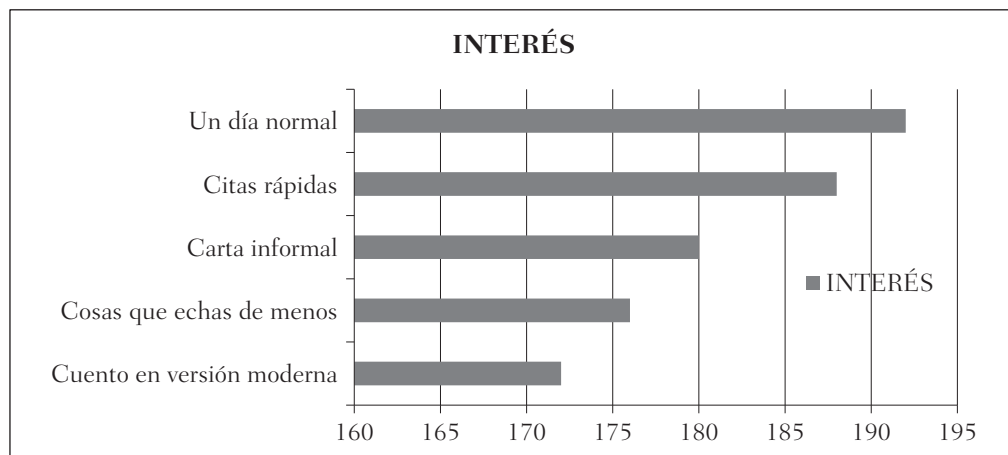
Lamentablemente el proceso creativo de esta actividad llevó tanto tiempo que nunca se pudo poner en común con el resto de la clase, lo que sin duda habría resultado muy interesante.

Algunos ejemplos de las historias creadas por los estudiantes fueron un Momotaro que

usa Facebook para buscar esposa, o una pequeña cerillera que, en vez de vender fósforos, se dedicaba a repartir pañuelos de papel en la estación de Shinjuku.

3.2 ENCUESTAS Y ANÁLISIS

En la encuesta de fin de curso, que se adjunta como anexo 1 al final de este trabajo, se pidió a los estudiantes que puntuaran cada actividad según dos parámetros: su interés y su dificultad, siendo 5 puntos el máximo en interés y 0 en mínimo. En cuanto a la dificultad, 0 correspondía a la mayor dificultad y 5, a la menor, por lo cual, cuanto más difícil les resultara una actividad, menos puntos le darían. El resultado se ilustra en los siguientes gráficos. En este trabajo, por limitaciones de espacio, solo vamos a hablar de las cinco actividades más interesantes y las cinco que les resultaron más difíciles.

INTERÉS

Actividad	Valor (≈)
Un día normal	192
Citas rápidas	188
Carta informal	180
Cosas que echas de menos	176
Cuento en versión moderna	172

(Eje: 160, 165, 170, 175, 180, 185, 190, 195 — ■ INTERÉS)

VIABILIDAD DE LA ACTIVIDAD

Actividad	Valor (≈)
Cuento	113
Test	105
Un suceso	101
CV	92
Poesía	60

(Eje: 0, 20, 40, 60, 80, 100, 120 — ■ VIABILIDAD)

Estas cantidades son el resultado de la suma de los puntos que cada estudiante había asignado a cada tarea:

"Un día normal" = 192 puntos. "Citas rápidas" = 188 puntos. "Carta informal" = 180 puntos. "Cosas que echas de menos" = 176 puntos. "Cuento en versión moderna" = 172 puntos.

"Poesía" = 60 puntos. "CV" = 92 puntos. "Suceso" = 100 puntos. "Test de la felicidad" = 104 puntos. "Cuento en versión moderna" = 112 puntos.

Como vemos en el gráfico, la actividad más popular fue "Un día normal en el mundo", que, ya cuando fue presentada por la profesora, obtuvo una recepción magnífica, con algunos estudiantes empezando a escribir sin terminar siquiera de escuchar las instrucciones. En los comentarios de las encuestas, los estudiantes escribieron que les había gustado mucho imaginar un viaje tan interesante y de hecho, fue la actividad donde más cantidad de material produjeron.

En cuanto a la segunda actividad más popular, que produjo un nivel de emoción y participación muy positivo, a los estudiantes les gustó mucho la idea de buscar pareja y además, hacerlo con fotografías, no sobre ellos mismos, les quitó cualquier presión personal. Solo hubo quejas de que algunos personajes de la fotografía no eran muy agraciados y la persona que había recibido esa fotografía pensaba que sería muy difícil ser correspondido. A destacar, sin embargo, la reacción de un estudiante que, tras recibir una de estas fotos "poco agraciadas", decidió hacer que su personaje fuera muy rico y con muchas aficiones, para asegurarse de que al menos tendría cosas en común con alguien. Le salió muy bien y fue elegido por dos personas.

De la carta informal sobre una fiesta hispana, los estudiantes resaltaron que había sido interesante aprender sobre estas fiestas y había sido divertido imaginar un viaje en grupo. De hecho, el trabajo de grupos fue algo que muchos estudiantes destacaron como positivo del curso, ya que les había permitido practicar la destreza escrita al mismo tiempo que la oral.

De la actividad "Cosas que echas de menos de Japón" los estudiantes escribieron que habían sentido mucha nostalgia de su viaje al extranjero y que esta actividad les había recordado los sentimientos que habían experimentado entonces.

Finalmente, sobre el cuento en versión moderna, los alumnos destacaron la originalidad de la idea y aseguraron que habían disfrutado imaginando y creando, aunque también comentaron que les había resultado bastante difícil.

En cuanto a las actividades en el top 5 de viabilidad o dificultad, la razón principal común

a todas las actividades fue que la generación de ideas les había resultado muy difícil, aun en grupos, lo cual hizo que nos planteáramos dar más modelos antes de la actividad y pautarla en más fases previas.

Mención especial para la actividad de poesía, que no solo les resultó difícil por el proceso creativo, sino porque la profesora cometió el error de seleccionar poesías demasiado complejas como modelo, con lo cual el tiempo de clase se pasó solamente intentando descifrar el significado. Por supuesto, también decidimos introducir abundantes modificaciones en esta actividad o incluso, viendo que el nivel de interés también había sido muy bajo, eliminarla del curso. Una de las opciones que pusimos en práctica en sesiones posteriores fue trabajar la rima de la poesía con alguna canción, donde numerábamos los huecos de la misma rima con el mismo número. Tanto la dinámica de clase como el resultado fueron mucho más exitosos aunque aún nos quedó pendiente la producción de la poesía misma que, concluimos, quizá deberíamos dejar para niveles más altos de español.

En cuanto a los comentarios generales de las encuestas, los más repetidos fueron que habían disfrutado la clase porque habían podido imaginar, ya que en otras clases no tenían esta oportunidad. Los estudiantes también comentaron que habían apreciado la oportunidad de escribir en español sobre temas que les habían resultado interesantes y, como ya hemos mencionado, destacaron además que les había gustado mucho trabajar en grupo ya que en otras clases trabajaban de forma más individual. Paradójicamente, a pesar de estos comentarios positivos hacia el trabajo en grupo, algunas de las actividades más populares fueron las que pertenecen al grupo donde se trabaja de forma individual. Las razones que indicaron fueron que los temas les habían resultado muy interesantes y habían disfrutado mucho imaginando.

En cuanto a los puntos que menos les habían gustado o que podían mejorar, hubo un estudiante que, además de no ser bueno usando tecnología y que, por lo tanto, parece que tuvo muchas dificultades para crear y mantener el blog, comentó que él quería una clase para aprender textos académicos, y solo le pareció útil e interesante la actividad de cómo escribir un curriculum vitae en español. No llegó a entender el objetivo de potenciar la creatividad porque no lo consideraba importante en su vida.

Durante el curso, también hubo alguna queja sobre el trabajo extra de escribir en el blog el texto corregido de clase, ya que lo entendían como una tarea innecesaria y les costó entender el objetivo de escribir el texto dos veces pero en las encuestas finales reconocieron que les había ayudado y sobre todo, que se sentían orgullosos de ver todo lo que eran capaces de

escribir en español.

4. EVALUACIÓN Y CRITERIOS DE CORRECCIÓN

En cuanto a los criterios de evaluación, que adjuntamos en el anexo 2 al final de este trabajo junto con el programa del curso, se decidió asignar un porcentaje de 50% al trabajo realizado en clase, basado en la observación y en el texto resultante de cada sesión, un 30% al blog y un 20% al examen final, impuesto por la universidad y que consistía en algunas pruebas del examen DELE B1.

Para la corrección de los textos que escribían o empezaban a escribir en clase, se decidió usar el siguiente código, basado en el que propone el libro de texto *Prisma B1* (2003: transparencia 12):

Error de gramática   _____   Error de vocabulario   ○

Error de ortografía ～～   Falta algo   V

Hay muchas teorías sobre el mejor momento para corregir un error pero, al menos en el caso de la expresión escrita, consideramos muy positivo ayudar al estudiante a descubrir y modificar sus fallos en el momento (Cassany, 1993a: 124), ya que con el paso del tiempo, se suele olvidar parcialmente (o completamente) tanto la forma como el contenido del texto que ha escrito.

De igual manera, nos parece completamente imprescindible que el estudiante vuelva a escribir su texto incorporando las correcciones sugeridas. De lo contrario, será más difícil que asimile su error y entienda cómo no cometerlo, de ahí que una de las tareas más importantes de la clase fuese el blog.

A pesar de la corrección formal, nos gustaría insistir en que este curso no consistía en aprender a escribir ni a corregir exclusivamente contenido léxico-sintáctico, sino un curso en el que los estudiantes pusieran en práctica de una forma lo más lúdica posible los conocimientos previos de español que habían adquirido en el año anterior. De tal manera, los criterios que primaron a la hora de realizar la evaluación, además del código ya mencionado, fueron los relativos a los componentes de la creatividad, aplicados de forma general basados en la definición de Guilford.

5. CONCLUSIONES

Comenzamos la preparación de este curso con varios objetivos: conseguir que los estudiantes disfrutaran escribiendo en español, que pusieran en práctica lo aprendido a lo largo del curso anterior, y que se potenciara su creatividad e imaginación, especialmente, su capacidad para generar ideas.

Basándonos en nuestra observación de la clase, pensamos que la mayoría de las actividades propuestas en el curso fueron del agrado de los estudiantes y que no les resultó pesado realizarlas. En algunas incluso, nos parece que realmente disfrutaron escribiendo e intercambiando ideas con sus compañeros, por lo cual, damos por cumplido el primer objetivo. Asimismo, las encuestas finales que pasamos a los estudiantes fueron en su mayoría muy positivas, destacando el hecho de que la clase les había parecido divertida, algo que aparentemente no consideraban posible en un curso de escritura y, de nuevo, apreciando la oportunidad de poder trabajar en grupos.

En cuanto a poner en práctica lo aprendido en etapas anteriores, aunque en muchas actividades fue así, hubo otras de vocabulario más específico que implicaron aprendizaje de vocabulario nuevo y difícil para los estudiantes. Consideramos eliminar las actividades de este tipo que no tuvieron un alto grado de interés y modificar las que sí les resultaron atractivas.

Con respecto al último objetivo y basándonos en los productos finales y en la observación de la clase en la etapa de producción de ideas, pensamos que al final de este curso los estudiantes estaban más acostumbrados a generar ideas sobre temas generales y específicos y que, como muchos mencionaron en las encuestas, el trabajo en grupo les había servido de gran ayuda, reconociendo que ese trabajo coral había mejorado el resultado final del texto y su propia creatividad.

De igual manera, nos ayudó mucho en la preparación del curso conocer la teoría sobre escritura, creatividad y escritura creativa, ya que nos permitió definir más claramente nuestros objetivos y nos guió en la tarea de preparar las actividades más relevantes para el curso. También resultó muy valioso este marco teórico en la etapa de evaluación de la creatividad, un tema del que, reconocemos, no teníamos mucha información previa.

Por supuesto, somos conscientes de que hay muchas cosas que revisar y mejorar, como también nos lo hicieron saber en las encuestas pero, gracias a estos valiosísimos comentarios de los estudiantes y del no menos valioso intercambio de ideas con otros colegas, seguiremos trabajando para hacer de este curso de Escritura Creativa un medio de aprendizaje a todos los niveles.

REFERENCIAS BIBLIOGRÁFICAS

Betancour, (2000), *Creatividad en la educación: educación para transformar.* Recuperado de http://psicopediahoy.com/creatividad-en-educacion/

Cassany, D. (1999), *Los procesos de escritura en el aula de E/LE*, Carabela, 46: p. 20

Cassany, D. (1999), *Construir la escritura.* Barcelona: Paidós

Consejo de Europa (2002), *Marco común europeo de referencia para las lenguas: aprendizaje, enseñanza, evaluación.* Madrid: Secretaría General Técnica del MEC, Anaya e Instituto Cervantes.

Equipo Prisma (2003), *Prisma Progresa B1: Libro del profesor.* Madrid: Edinumen.

García Carcedo, Pilar (2011), *Educación literaria y escritura creativa*, Granada, GEU

García Parejo, I. (1999), *Teoría de la expresión escrita en la enseñanza de segundas lenguas*, Carabela, 46: pp. 23-42.

Guilford, J. P. (1950), *Creativity*, American Psychologist, 5, pp. 444-454.

Guilford, J. P. (1967), *The nature of intelligence*, Nueva York: McGraw-Hill.

Hal, B., & Sweet, C. (2008), *The Writing Community: A New Model for the Creative Writing Classroom*, Durham: Duke University Press.

Jiménez González, Juan E. (et al) (2007), *Adaptación y baremación del test de pensamiento creativo de Torrance: expresión figurada. Educación Primaria y Secundaria*, Canarias: Consejería de Educación, Cultura y Deportes del gobierno de Canarias. Dirección General de Ordenación e Innovación Educativa.

Knox, Grahame. *Moon landing*. Recuperado de *http://insight.typepad.co.uk/moon_landing.pdf*

Moreno, C., Zurita, P. & Moreno, V. (1999), *A escribir se aprende escribiendo*, Carabela, 46: pp. 75-101.

Moreno, C. (2011), *Creación y explotación de materiales para la expresión escrita. Materiales, estrategias y recursos para la enseñanza del español como 2-L*, (pp. 373-414). Madrid: Arco Libros.

Piñeiro, R. (2013). *22 cosas que echas de menos de España ahora que no vives aquí.* Recuperado de http://www.traveler.es/viajes/mundo-traveler/articulos/22-cosas-que-echas-de-menos-de-espana-ahora-que-no-vives-aqui/4566

Real Academia Española, *Diccionario de la Lengua Española*. Recuperado de http://dle.rae.es/?id=GKCiBz2

Torrance, E. P. (1974), *The Torrance test of creative thinking: Norms-technical manual*, Bensenville's IL, Scholastic Testing Service.

Vázquez, M. (2000), *Apuntes sobre creatividad: origen del término y su pervivencia.* Revista Latina de Comunicación Social, 25, pp. 1-7.

ANEXOS

Anexo 1.

<div align="center">

ENCUESTA

</div>

1. Escribe tu opinión sobre las actividades que hemos hecho en clase en este curso, por favor, y puntúalas de 1 a 5:

1. Entrevista a un compañero.				
5 Muy interesante	4 Interesante	3 Normal	2 Poco interesante	1 No interesante
5 Muy fácil	4 Fácil	3 Normal	2 Difícil	1 Muy difícil

2. Historia colaborativa.				
5 Muy interesante	4 Interesante	3 Normal	2 Poco interesante	1 No interesante
5 Muy fácil	4 Fácil	3 Normal	2 Difícil	1 Muy difícil

3. Citas rápidas.				
5 Muy interesante	4 Interesante	3 Normal	2 Poco interesante	1 No interesante
5 Muy fácil	4 Fácil	3 Normal	2 Difícil	1 Muy difícil

4. Cosas que echas de menos.				
5 Muy interesante	4 Interesante	3 Normal	2 Poco interesante	1 No interesante
5 Muy fácil	4 Fácil	3 Normal	2 Difícil	1 Muy difícil

5. Test de la felicidad.				
5 Muy interesante	4 Interesante	3 Normal	2 Poco interesante	1 No interesante
5 Muy fácil	4 Fácil	3 Normal	2 Difícil	1 Muy difícil

6. Carta informal sobre una fiesta hispana.				
5 Muy interesante	4 Interesante	3 Normal	2 Poco interesante	1 No interesante
5 Muy fácil	4 Fácil	3 Normal	2 Difícil	1 Muy difícil

7. Un día perfecto en el mundo.				
5 Muy interesante	4 Interesante	3 Normal	2 Poco interesante	1 No interesante
5 Muy fácil	4 Fácil	3 Normal	2 Difícil	1 Muy difícil

8. Biografía de un hispano.				
5 Muy interesante	4 Interesante	3 Normal	2 Poco interesante	1 No interesante
5 Muy fácil	4 Fácil	3 Normal	2 Difícil	1 Muy difícil

9. Las normas de tu país.				
5 Muy interesante	4 Interesante	3 Normal	2 Poco interesante	1 No interesante
5 Muy fácil	4 Fácil	3 Normal	2 Difícil	1 Muy difícil

10. Cuento clásico en versión moderna.				
5 Muy interesante	4 Interesante	3 Normal	2 Poco interesante	1 No interesante
5 Muy fácil	4 Fácil	3 Normal	2 Difícil	1 Muy difícil
11. Un suceso.				
5 Muy interesante	4 Interesante	3 Normal	2 Poco interesante	1 No interesante
5 Muy fácil	4 Fácil	3 Normal	2 Difícil	1 Muy difícil
12. Escribir una poesía.				
5 Muy interesante	4 Interesante	3 Normal	2 Poco interesante	1 No interesante
5 Muy fácil	4 Fácil	3 Normal	2 Difícil	1 Muy difícil
13. Curriculum vitae.				
5 Muy interesante	4 Interesante	3 Normal	2 Poco interesante	1 No interesante
5 Muy fácil	4 Fácil	3 Normal	2 Difícil	1 Muy difícil

2. ¿Qué te ha gustado en este curso y por qué?

3. ¿Qué **NO** te ha gustado en este curso y por qué? ¿Qué podría mejorar?

¡MUCHAS GRACIAS!

Anexo 2.

ESCRITURA CREATIVA

1. PROGRAMA DEL CURSO:

CLASE 1	Presentación de la clase, explicación del contenido y de los criterios de evaluación. Presentación y entrevista. (Traer el ordenador para la próxima clase)
CLASE 2	Escribir una historia de forma colaborativa. Pretérito indefinido/imperfecto. Explicación del blog y creación del mismo en clase.
CLASE 3	Citas rápidas. Buscar a la pareja ideal y escribir una entrada en un foro.
CLASE 4	Crear una lista común y justificarla. "22 cosas que echas de menos cuando…"
CLASE 5	Elaborar un test para evaluar el nivel de felicidad de los compañeros de clase. También se crearán las puntuaciones y las conclusiones de ese test.
CLASE 6	Buscar información sobre una fiesta hispana y escribir una carta informal a un amigo contando la experiencia en esa fiesta.
CLASE 7	Diario de viajes: un día normal en el mundo. Escribiremos sobre la rutina diaria ideal si pudiéramos viajar por el mundo en un día.
CLASE 8	Biografía. Buscar información sobre un personaje hispanohablante y escribir su biografía.
CLASE 9	Crear las leyes de un país. Estructuras de opinión con subjuntivo. Leer las leyes de otros grupos y elegir el mejor país para vivir.
CLASE 10	Reescribir un cuento clásico desde la perspectiva actual. Estructuras para escribir cuentos.
CLASE 11	Artículo periodístico. Describir un suceso (real o imaginario)
CLASE 12	Escribir poesías en diferentes grupos con diferentes temáticas.
CLASE 13	Escribir un curriculum vitae en español.
CLASE 14	EXAMEN: Prueba de Comprensión lectora, Comprensión auditiva y Expresión escrita del examen DELE B1.

2. CRITERIOS DE EVALUACIÓN:

A. Trabajo y participación en clase: 50%

A	B	C	D
Ha participado muy bien en clase y ha hecho los deberes siempre.	Ha participado bien en clase y ha hecho los deberes normalmente.	Ha participado en clase regular y ha hecho los deberes algunas veces.	Ha participado poco en clase y ha hecho los deberes muy pocas veces o ninguna.

B. Blog: 30%

A	B	C	D
La presentación y el contenido son excelentes. Muy buen esfuerzo	La presentación y el contenido son muy buenos. Buen esfuerzo	La presentación y el contenido son buenos. Esfuerzo regular	La presentación y el contenido son insuficientes o no presentados

C. Examen final: 20%

A	B	C	D
90-100 puntos	70-89 puntos	50-69 puntos	Menos de 49 puntos

Extensive Reading
in a University EFL Classroom:
Issues of Fluency, Vocabulary, and Proximity

David P. Shea

Abstract

As universities move to offer content classes in the medium of English, students face the challenge to handle large amounts of reading. Research in English as a Foreign Language (EFL) suggests that extensive reading (ER) can strengthen fluency as well as improve attitudes toward reading. While most ER researchers recommend graded readers, sociocultural approaches suggest that shared texts may also prove effective. In this paper, I report on a qualitative investigation of a required first-year reading class which incorporated selective components of ER: a large quantity of interesting text, read for enjoyment with a focus on meaning, not translation. Students overwhelmingly endorsed ER, reporting faster reading speeds and more positive attitudes, yet a number of students felt that their vocabulary declined, due to skipping unknown words. However, diagnostic tests suggested that although academic vocabulary may have declined slightly, general vocabulary likely increased. While a modified ER approach may be effective, it may be helpful to supplement reading with vocabulary practice and to "sell" students on the idea of ER.

Introduction

With the spread of globalization, there is increasing pressure placed on universities in Japan to offer classes that deliver authentic academic content in English (Brown, 2014). The Ministry of Education has endorsed a university curriculum where "Japanese people can acquire the necessary English skills" to take classes with foreign students in the country (MEXT, 2012, p. 17), and the situation is arguably similar in other parts of Asia (Kirkpatrick, 2016). This increase of English medium instruction (EMI) raises questions about the balance between content and language skills, between studying English as a foreign language versus

studying an academic subject in English. There is some overlap, depending on the context. Content-based instruction (Lyster, 2007) places emphasis on academic subject matter, with a simultaneous concern for various elements of linguistic structure and discourse. Carty and Susser (2015) note, however, that with most EMI, there is usually little if any provision for language learning, which suggests that for many Japanese students, EMI is a sink or swim situation. Advanced "returnee" students may have the language proficiency to take classes in English, but "regular" students face an imposing hurdle.

Perhaps the biggest challenge for intermediate first-year university students who hope to enroll in upper level EMI courses is learning how to deal with the large amount of required reading (Spack, 1997). For many students at this stage, there is a strong tendency to rely on an ineffective and time consuming word-for-word approach to reading, which involves the "close study of short passages, including syntactic, semantic, and lexical analyses and translation into the L1 to study meaning" (Susser & Robb, 1990, p. 161). An intensive approach, however, simply does not give students the skills to deal with large amounts of reading material.

Although the character of English education is slowly changing, with the recent introduction of public international schools (Noguchi, 2015) and educational initiatives for high-level returnee education (Kanno, 2003), the goal of reading instruction in many high schools is typically word-for-word translation, with explanation of associated grammar patterns, often in the L1 (Nishino & Watanabe, 2008). Informal surveys of incoming students in my first-year classes regularly confirm that many have never read an entire book in English. The problem is compounded when students report that not only have they had little practice reading, they basically do not like to read. The challenge for instruction thus involves developing both more positive attitudes as well as more effective reading practices. As Lees and Althomsons (2015, p. 74) point out, "non-fluent readers may be trapped in a vicious circle; by reading slowly and with little comprehension, they do not enjoy what they read, which in turn causes them to read less and further stunts their reading attitude and skill development."

Extensive Reading

A good deal of educational research suggests that extensive reading (ER) is an effective approach that improves fluency as well as attitudes toward reading, giving students the skills to deal with large amounts of text. In Spack's (1997) study of the academic socialization of

a Japanese student at an American university, for example, one of the turning points was the student's decision to read extensively over the summer vacation, which improved confidence as well as proficiency.

One of the strongest proponents of ER, Nation (2015), has stated that the "single most significant change that a teacher could make to a language learning course would be to include a substantial extensive reading program" (p. 139). Nation (2001) has also argued that reading is "strongly related" to English L2 proficiency, pointing out that the amount of reading learners do outside class is the "most important direct contributor to TOEFL test performance" (p. 154). In reviewing empirical studies of "book floods," Elley (1991) found that engagement in reading programs using ER produced strong gains on text comprehension measures, as well as increased motivation to read. Bell (2001) noted that extensive reading produced higher comprehension than intensive methods focused on short texts with grammar and vocabulary tasks. Robb and Susser (1989) found that university students who were engaged in ER activities performed equal to students engaged in a traditional skills-based approach, but with greater interest and enjoyment. Other research (e.g., Day & Bamford, 1998; Dupuy, Tse, & Cook, 1996; Krashen, 2000) has provided compelling evidence that ER is a positive means of developing English proficiency.

Although there are common practices that define extensive reading (Stoller, 2015), ER is perhaps most widely associated with use of simplified graded readers controlled for vocabulary and grammatical complexity (Bamford & Day, 1997; Hill, 1997; Nation & Wang, 1999). Within a graded reader program, students typically choose books that they read individually, either inside or outside class. According to Day and Bamford (1998), students select only books in which they are interested, with "the freedom to stop reading when they want to, with no questions asked" (p. 27). Within this paradigm, reading is an independent activity with the teacher's role not to instruct, but to encourage students as a model reader, an "active member of the classroom reading community" (Day & Bamford, 1998, p. 8). Prowse (2002) has argued that most instructional follow-up activities such as tests and comprehension questions are inappropriate because they preclude the natural engagement with books that stems from intrinsic interest rather than compulsory assignment.

In endorsing ER, Nation (2015) points out that while reading for meaning provides a good opportunity to develop incidental acquisition of vocabulary, students need to understand 98% of running words if they are to effectively comprehend a text. A large amount of reading is needed to acquire new vocabulary incidentally, so that the repeated meeting of new words

strengthens the "small amounts of knowledge gained from previous meetings" (Nation, 1997, p. 15). Along similar lines, Waring and Takaki (2003) concluded that students need a good deal of repeated exposure if they are to successfully acquire vocabulary through reading for meaning. However, Waring and Takaki argue that the primary "aim" of ER is not vocabulary growth, but development of already known vocabulary. The point is that for ER to be successful, it is critically important that students understand the books they are assigned. Nation (2001, p. 21-22) recommended conducting a diagnostic vocabulary levels test as a "reliable, valid and very practical" means to determine word recognition ability.

Many ER approaches refer to Krashen's (1985) theory of acquisition via comprehensible input, which hypothesizes that reading itself improves achievement – as long as the text is understandable and interesting, so that the "more students read, the better they become at it" (Day & Bamford, 1998, p. 4). In the classroom, the focus on natural acquisition often revolves around sustained silent reading (Chow & Chou, 2000), an approach where students devote time in the classroom to reading, typically accompanied by the teacher.

A number of teacher-researchers, however, advocate a more social approach, seeking to balance learner development and collaborative interaction (Barfield, 2000). Within this view, the teacher adopts a proactive role, guiding interaction in class and facilitating understanding of the text through discussion and follow-up activities. Casanave (1993), for example, argued that students "should think, talk and write about what they read, before, during, and after" engagement with a text (p. 149), and she proposed that students write reflective journals in response to reading assignments, with the goal not to reproduce textual facts, but engage with ideas and generate thoughtful response. Renandya, Rajan, and Jacobs (1999) incorporated a number of post-reading tasks into classroom activities, especially group-based discussion and found that in-class peer interaction not only reinforced learning vocabulary, it also gave students a sense of progress and broader understanding. Similarly, Jacobs and Gallo (2002) integrated cooperative learning into ER classes, producing higher reading scores and more positive attitudes. Both Dupey, Tse, and Cook (1996) and Hill and Van Horn (1997) stress the importance of talking about books, discussing themes and critical issues, and addressing open-ended questions related to character motivations, plot features, and related personal experience. It is important, Dupey, Tse, and Cook (1996) argue, to give students relevant cultural and historical information to facilitate understanding, a point that seems especially relevant in Asian contexts.

While the rationale for integrating follow-up activities into an ER program often refers

to Swain's (2000) output hypothesis, which suggests that productive expression, both oral and written, enhances acquisition, Swain (1993) herself argues that output is not as much about pushing students toward more conscious awareness of increased grammatical accuracy, as it is about situating engagement with texts in productive activities where students are involved in expressing and exchanging opinions. A few ER advocates (e.g., Jacobs & Hannah, 2009) have grounded an extensive approach to reading in the Vygotskyian notions of scaffolding (Wood, Bruner, & Ross, 1976) and shared cognition (Moll, 1990) which view learning not as individual but collaborative social activity.

Methodology

This paper is a report of a small-scale, action research project (Burns, 2009) involving reflective practice (Walsh, 2011), which investigated the effectiveness of using extensive reading in a required first-year university English course. I was particularly interested in adopting ER, but given the administrative, institutional, and organizational difficulties in setting up a graded reader program, I opted to use shared texts (Rosszell, 2010) that students read together as class assignments. As a result, a modified ER approach was adopted: a large quantity of interesting but accessible English text, read for enjoyment with a focus on meaning, not translation or skill-based exercises.

There were two sections of the same class which met weekly over two semesters, with approximately 30 students in both sections (attendance varied significantly from week to week). The two required texts were *Holes* (Sachar, 1998) and *Molly Moon's Incredible Book of Hypnotism* (Byng, 2002) critically acclaimed young adult novels filled with action, humor, and adventure. In total, students read more that 600 pages of text over the course of the academic year, at a rate of approximately 20-25 pages per week. I provided a basic glossary for both novels that students could use as they read. At the beginning of each class, there was a brief short-answer quiz with the stated purpose of encouraging students to keep up with the reading. Quizzes were designed to review plot details, and students who did poorly on a quiz were allowed to write a short summary as a make-up exercise. In-class activities were organized around discussion of the reading, and students made short presentations explaining selected passages. They wrote journal entries that they sometimes read aloud. I gave mini-lectures explaining key vocabulary terms, central themes, and culturally specific background information. I also talked about my own response to the reading and asked students to do the same.

On the first class at the beginning of the academic year, I administered a vocabulary levels diagnostic (Nation, 2001, p. 416) to obtain an approximate idea of vocabulary recognition at the 2,000 (2K) and 3,000 (3K) levels, as well as the academic word list (AWL). Overall scores for all students on the 2K diagnostic averaged 27 on a 30 item scale (90%), suggesting a basic-intermediate reading proficiency. Scores on the AWL and 3K tests were slightly lower. Average scores measured 22/30 on the 3K (77%) and 18/30 on the AWL (60%).

Individually, two students scored low (below 20) on the 2K test. Five students scored below 20 on the AWL, and 12 students on the 3K. A preliminary estimate of the reading level of *Holes*, based on a selected excerpt, suggested that the book would be broadly accessible to a reader with a strong command of the 2,000 word level, with approximately 95% of the vocabulary at the 2K level, excluding proper names and words likely familiar to a Japanese reader. It was fairly clear that most students would not have 98% coverage of the vocabulary of the novels.

In terms of overall language proficiency, individual student scores for the reading segment of the TOEIC IP placement test ranged from 115 to approximately 300, reflecting noticeably disparate abilities, from low to mid-intermediate, a common situation in university EFL classes. Most students were diligent and attentive in class, although 5-6 students were absent three or more times per semester, and 4-5 students regularly showed up 10-15 minutes late. Two students had trouble staying awake in class, though they were attentive when not drifting off. Four students had lived abroad for periods of six to nine months as exchange students during high school.

Individual learning styles were noticeably mixed. For example, Aki (all names are pseudonyms) was attentive and consistently received A's on quizzes, but he scored low on the vocabulary levels diagnostic (73% at the 2K level). Similarly, Taka scored high on the vocabulary index (97% at the 3K level) but was often absent and rarely spoke in class. In contrast, Nori was a returnee who had what seemed to be near native proficiency coupled with the highest vocabulary level scores (100% at 3,000, 83% at 5,000 level) but, until the last third of fall semester, was disinclined to make an effort to participate in class discussion or keep up with assignments (i.e., he read in spurts, usually at the last minute). Kazuko, on the other hand, asked for extra reading after quickly finishing both texts and in fact read two far more advanced novels, including *Jurassic Park*. Most students were characteristically shy and reluctant to talk in English, even in small group contexts.

At the end of the academic year in January, I asked students to complete Form B of the

vocabulary levels diagnostic, to obtain a broad baseline comparison. Because of absences, late registration and other factors, only 50 students took both diagnostics (Form A in April and Form B January). Scores on the two forms, A and B, were compared to indicate broad changes in vocabulary recognition. I also asked students in January to complete a 38 item evaluation questionnaire using a 4 point Likert scale (strongly agree, agree, disagree, strongly disagree), with open-ended questions and space for comments. Responses were anonymous. The intention was not to find statistical significance, but to get an indication of the effectiveness of ER. Feedback from students was evaluated within a qualitative paradigm (Patton, 2015), using interpretative, inductive analysis (Hubbard & Power, 1993; Thomas, 2006) to generate a grounded understanding of student attitudes. I tabulated survey answers in terms of a simple calculation of percentage. I report the findings in the section below.

Results

First and foremost, there was a near unanimous endorsement of ER as an effective approach to English study. A clear majority of students (77%) thought that their reading improved because of ER. Almost everyone (88%) said that they read faster, with greater understanding (96%) and increased enjoyment (96%). Taken together, all students (100%) reported that they felt either speed or comprehension had improved. There was one student who thought his reading speed didn't increase, but his comprehension did, while six students responded that comprehension had not improved, but reading speed had. A large majority of students (87%) stated that they understood the gist of the narrative, defined as more than 75% of the story, and two of the eight students who said they only understood only 50% of the story actually made A's in the class, consistently scoring well on quizzes and the final exam.

Perhaps more significantly, most students felt that they had developed more positive attitudes toward reading because of ER. Though not all were unreservedly positive, either in evaluations or class participation, nearly every student asserted more favorable orientations toward English. The majority (75%) indicated that, after taking the course, they enjoyed reading more than at the beginning of the year. Nearly two thirds (62%) attributed their more positive attitudes toward English to ER. Most (80%) said that they would take another ER class if offered, and 57% stated that they were more likely to read an English book in their free time because of ER. Taken together, all but a handful of students (95%) asserted that they either enjoyed reading more, were more positive about English in general, or would take another ER course given the opportunity.

The change in attitudes among students who were either reluctant readers or negative about English was particularly striking. Eighteen students reported at the beginning of the academic year that they did not like English in high school. Of the 18, almost all (16) said that at the end of the ER course they enjoyed reading more or were willing to enroll in another class with similar reading requirements. Some of the attitudinal changes were dramatic. For example, Matosuke commented: "I hated English in high school because it was not fun. There were too many things to memorize and the teacher let us read only boring sentences. But we can read interesting novels in ER and I'm enjoying studying now." Ayami, another student who didn't like English in high school, said: "I really didn't like English. I knew it was useful, but I thought that studying was hard and boring. But now I enjoy reading more than before and I think it's because of Molly Moon! I really enjoy reading this book."

There was also a shift in attitudes among enthusiastic students. For example, Chikako, said that the way she thought about reading improved:

> *Until I took this class. I always had to translate each word. Studying for the entrance exam, learning grammar, reading difficult texts, I couldn't enjoy the content of what I was reading.… And at first, with the homework for this class, I translated every word as I had always done, but that amount decreased the more I read, and I was able to guess even words I didn't know before. Because of that, my reading speed became a lot faster and since I didn't have to stop reading in the middle of the story, I could enjoy the story much more.* [original in Japanese; translations are mine]

Similarly, Taka, a reluctant student who rarely demonstrated outward enthusiasm but who scored very high on the 3K vocabulary test (97%), said that he was "thankful" for the class because it gave him a new orientation toward English:

> *I like to read English books a little more than before. I will try to read English books for myself. This class is an opportunity to start to read English in private.*

Choice of Texts

Nearly all students found the two novels interesting. All but three (88%) said they liked *Holes*, and all but one (96%) said they enjoyed *Molly Moon*, suggesting that perhaps attitudes became more positive as reading proficiency increased over the year. Although there was

more to the class than simply reading, the overwhelming majority of students (96%) said that they enjoyed the course as a whole. The only student who said he enjoyed neither the class nor *Molly Moon* was Aki, who was nonetheless positive about ER. He replied that the class had been helpful, and said that his reading confidence and speed, as well as vocabulary and overall English proficiency, all improved by taking the class. Importantly, Aki's grade for the class was a solid A. He scored consistently high marks on the quizzes and participated actively and attentively in class discussion.

Only a minority of students (25%) expressed a desire to choose their own individual book to read. A full three-fourths of students said that they would rather have the entire class read the same book chosen by the instructor. Students who said they wanted to read different texts were divided into two fairly clear categories. Most (11 of 15) were strong readers with high vocabulary test scores who reliably completed reading assignments and participated actively in class. A few (4 of 15) were more passive in class and either had trouble with attendance or completing reading assignments. In other words, class readers were positively received by the majority, with little resistance to a shared text. In fact, a number of students noted that they were able to complete the reading precisely because of the structured nature of the assignments. A full 80% of students admitted that, were they not enrolled in the ER class, they would likely not have read as much or as regularly. Only 12 of the stronger, more independent students reported that they would likely be reading English even if they had not had the class assignments.

The weekly quizzes, designed to structure the reading in terms of pace and attention to details of plot and character, were not overly popular. At the same time, most students seemed to accept the value of the quizzes. Only slightly more than half (53%) agreed that they actually liked having quizzes. At the same time, 62% admitted that, without the quizzes, they would likely not have read as consistently or as carefully. By far, the large majority (92%) admitted that the quizzes served as a useful review of the story and were a good check of comprehension.

Vocabulary

In spite of the positive attitudes and strong endorsement of ER, almost half of the students (47%) thought that their vocabulary did not improve doing ER. Only a slim majority (53%) said their vocabulary increased. A few even felt that their vocabulary had actually decreased because of ER. Apparently, some students thought that the emphasis on quantity adversely

affected the quality of reading. Certainly, the practical exigencies of reading 20-25 pages per week pushed students to focus on understanding the bigger picture of the story rather than the particular details of word meaning. Given this tendency to read pragmatically and skip unknown words, some students seemed to feel that they were not acquiring new vocabulary incidentally. Further, the skepticism about vocabulary development did not seem to be related to not doing the reading. While eight students who stated that their vocabulary did not increase admitted that they sometimes skipped reading assignments, the other 20 students who said their vocabulary did not improve reported doing homework conscientiously.

Student impressions about vocabulary acquisition seem rooted, at least in part, in the practice of guessing, a reading strategy which appeared prevalent. Students almost universally (97%) reported guessing the meaning of unfamiliar words encountered in the course of reading. Although most (73%) seemed to feel that guessing actually helped to learn new words, many seemed to automatically assume that guessing was not productive and in fact, even constituted a kind of academic laziness. For example, one student explained that his vocabulary did not improve because, he said, "I skip a lot of words." Another student used almost exactly the same reasoning: "The disadvantage is that I skip words I don't know, so it's difficult to learn new words." Still another drew a distinction between comprehending the story and learning vocabulary:

> When I came across a word that I didn't know, I skipped over it and kept reading. More than before, I got used to reading a lot of text. On the other hand, I think my knowledge of vocabulary decreased as a result. – S22 [original in Japanese]

Apparently, some students felt that, when they skipped words (sometimes whole sections) in order to grasp the "gist" of the story without stopping to look up unknown vocabulary, they missed something important. In other words, students seemed to be assuming that they could not learn if they did not pay explicit attention, arguably an attitude rooted in the intensive approach to reading developed in high school English classes.

Student intuitions were not supported, however, by the vocabulary levels tests which offered a more complex picture, suggesting that students may have in fact gained vocabulary, at least the kind associated with reading novels. There may have been a slight decline in recognition of academic vocabulary terms that do not usually appear in fiction but that students typically study in preparation for university entrance exams, but the loss appeared to be small.

On the whole, average vocabulary scores for the three levels (2K, 3K and AWL) showed distinct trends. On the 2K test, average scores showed little decline over the course of the year. If anything, class averages suggest a slight increase. The average score in April was 27.4 of 30. The average score at the end of fall semester was 27.7, a largely insignificant change. Interestingly, however, four students showed changes of five or more points, and three of the four were positive gains. For example, Josaburo, who demonstrated the biggest shift, scored 18/30 (60%) on the 2K diagnostic in April, and 27/30 (90%) in January, which was a noticeable increase.

On the AWL diagnostic, the average score showed a slight decline over the year. Total scores for both classes dropped a total of 137 points, and the average individual score dropped from 23.8/30 to 21.1/30, a little more than two words per student. Among individual students who demonstrated a noticeable shift of 5 points or more, eleven students showed declines, while only one showed an increase. Etaya, for example, dropped from 24/30 (80%) to 15/30 (50%), which was one of the two largest shifts. The downward trend suggests that on the whole, students may indeed have lost some command of the academic vocabulary they studied prior to entering university, perhaps reflecting a decline that follows the intensive study associated with exam preparation. In the words of Sakio, "You may have a good ability to understand and read English when you enter college, but it declines day by day because you study less than before." The lower scores seem to reflect how students slowly lost some of the complex vocabulary they had memorized.

However, scores for vocabulary at the non-specialized 3K level of frequency demonstrated a markedly different trend. Diagnostic results showed an increase at the 3K level more pronounced than the decline in academic vocabulary. Scores for all students increased 151 points, while the average individual score showed an increase of more than three words per student, rising from 21.3 to 24.4. In particular, sixteen students demonstrated a shift of 5 or more points, with four students making gains of 10 points or more. Guriko, who lost 10 points on the AWL test, gained 23 points at the 3K level. Matosuke gained gained 10 points and Sasuko gained 11. In sum, the increases in recognition of non-specialized words at the 3K range surpassed the AWL decline, both in overall totals and individual scores.

The increase in vocabulary was particularly noticeable in students who scored low on the diagnostic test at the beginning of Spring semester. Two students scored below what might be considered a minimum criterion of 20 points (66%) on the 2K diagnostic, and 12 students scored below 20 on the 3K test. At the end of the year, all 14 students demonstrated

increases, all more than five points each. On the AWL diagnostic, six students scored below 20, and there was little noticeable change, though one student increased six points. All students who scored low on the AWL, however, showed increases on the 3K.

While the vocabulary levels tests provide only approximate measures of word recognition, they do suggest that students might not be losing vocabulary in the way they imagine. Explanations students made in open-ended comments are instructive on this point, in part because they illustrate a flawed dichotomous logic. For these students, it was either understanding content or skills development. For example, S11 said:

I could enjoy reading English because I didn't have to translate. However, I need to improve my English grammar and vocabulary, so I wanted more grammar explanation and vocabulary tests in class.

Student S14 concurred, for a similar reason:

I like extensive reading better…. But when I found the word I didn't know, I almost never used a dictionary, so my vocabulary has not improved.

These students admitted that they enjoyed extensive reading, but at the same time, they did not feel that enjoyable reading counted as language learning.

In effect, ER seemed to be too pleasurable, too much fun to be study. Elley (1991, p. 403) mentioned this point, though in reference to elementary school administrators and parents, not students. In several of the studies he reviewed, Elley notes that skeptics viewed ER as students "merely enjoying themselves, rather than learning." Yu (1997, p. 3), also noted that students regarded ER "an optional extra, a luxury" rather than an effective means to foster a "reading culture." Certainly, the notion that learning requires diligence and conscientious effort is deeply rooted in traditional attitudes toward education. Evidently, some students seem to have internalized a conception of English study that equates learning not with using the language, but with explicit and earnest attention to its constituent parts.

Proximity

Another notable finding involved the way students thought about themselves vis-à-vis English. There was a perceptible shift in terms of proximity, as many students stated they felt

closer to English because of ER. For example, Guriko said, "I'm not afraid of reading English now" and Tahoya remarked that English had "become familiar" through ER. In part, the shift in proximity was related to the way students understood English, as the language became less a subject of study that required correctness and attention to detail, and more a tool of communication. In particular, students began to relate to English through the enjoyment and excitement of narrative. Many students made a point to contrast ER with the traditional grammar-translation approach, which nearly everyone (92%) reported using to some extent in high school. With ER, "I've been able to feel the enjoyment of English," Jofumi noted, adding that "I feel that I've drawn closer to English." Similarly, Banichi commented that when he "studied for the entrance exams," he "hated" English, while now he "enjoys reading more." Junoya said, "My way of reading changed because I don't translate English into Japanese anymore," and Moyuki insisted that while he had learned English in the traditional intensive style for years, he now thinks it "useless" because extensive English is "more fun."

The shift in proximity generated a sense of ownership and connection, as students began to appropriate English as something they related to personally. Manami said, for example, "I learned to read English as English. Now I seem to understand how people using English feel." Similarly, Haseo commented, "I don't translate English into Japanese. Rather, I now think in English as I read. I feel absorbed in English as it is." In other words, student identities began to change through ER. Students developed, in Hakui's words, "English brains" as they began to think in English and see themselves as readers to whom English belonged. Rather than locating English externally, as a subject of study for exams with right and wrong answers, English became rooted in the imagination and personal sense of value.

Part of the shift in identity was generated by the growing ability to engage with extended narrative. That is, the reading was tied to creative involvement. Students read successfully because they were able to construct the story, which worked in turn to attract them as readers. This narrative engagement was evident in the adventure and excitement of the plot, the appeal of the characters, and the relevance of underlying themes. For many, however, affective engagement was neither immediate nor intrinsic to the story; rather, it was an active process of construction that developed over time, only with persistence and concerted effort. For example, Goto, one of the most skeptical students in class commented:

> *The book itself is relatively easy. If it were Japanese, even elementary school kids would understand. But I'm not good at English (especially reading) and doing the homework*

every week was really difficult. The only thing that saved me was the interesting story. You don't often find this kind of atmosphere in Japanese novels, and I like it. What's more, I can empathize with the heroine, who is weak and unattractive.

For Goto, it was the connection he was able to construct with the story that allowed him to deal with the workload and the challenge of getting through the reading assignments. More specifically, Goto's ability to get involved with the story provided the stimulus to keep reading.

The connection that other students built with the narrative was also strong, though the focus differed. For example, Saiko pointed to Molly, the protagonist's changes of heart and stated, "I was encouraged by Molly." Makiko commented on the vicarious excitement of imagined delinquency: "I experienced committing a robbery with Molly and Rocky. Though robbing a bank is bad, I enjoyed the thrill." Another student, Etaya, talked about the excitement of travel to America: "This book taught me something important about my life. I also learned about American culture. I want to go to America." And Mami remarked that, like Molly, she too "wanted to use hypnotism!"

However, a number of students found the reading difficult at first, which made the story hard to relate to. Sansuke, for example, noted how the level of difficulty changed as he began to become involved in the story and make connections with the protagonist:

Until I caught the story, I thought it was difficult to read and there are many vocabulary words I didn't know. But as I became used to reading, I began to like Molly, who noticed that hypnotizing is a dangerous power while she lived a gorgeous life.

Similarly, Manami pointed to her initial skepticism of the main character, because of what she saw as the heroine's selfishness:

The first part of this book is so disgusting. I disliked Molly's character and the idea that she could make anyone do what she wanted by hypnotism. I thought again and again that I wanted to give up reading, but the story turned good. Molly changed, and I like her now. This book is so enjoyable.

Until Manami thought the story had a moral point to which she could relate, her basic dislike worked as an impediment and held her back to the point where she was ready to give up.

What is significant here is that these comments express connections with literature that experienced readers make. Students were able to go beyond reading the text as an assignment and make contact with the narrative, developing an emotional investment in the story. In addition, students were able to read beyond their competence, and develop increased proficiency because they were drawn into the story by the narrative. However, the emotional involvement did not come automatically. Understanding was a constructive process which took place over time within the supportive structure of the class as a whole.

Discussion

In this paper, I have reported on a local investigation of using a modified ER approach in a first-year university reading class. The evidence suggests that in reading the two novels extensively, most students felt that their reading fluency (speed and comprehension) increased. Further, interest in English and attitudes toward reading also improved for many, with some students expressing a noticeably stronger sense of proximity to the language. A small but significant number of students felt concerned that reading extensively did not develop their vocabulary, because attention was focused on meaning, not explicitly learning new words. The diagnostic tests suggest, however, that while academic vocabulary may have indeed dropped slightly, vocabulary at 3K level (that is, non-specialized words that appeared in the novels) may actually have increased.

Research in ER often talks about the inherent motivation of enjoyable texts (e.g., Stoller, 2015), but it seems that it may be important, not simply to trust the intrinsic power of ER, but to "sell" students on the idea, explicitly reminding them of the value of reading for meaning and the particular challenges that the approach entails. Given the experience many first-year students have with English in high school, it would also seem to be helpful to address unstated assumptions about how the language is acquired and remind students that not only does learning and enjoyment often go together, but that an important aspect of second language proficiency is acquired in the tacit focus on meaning and authentic content.

At the same time, evidence from this study seems to suggest that it would probably be a good idea to supplement ER with additional vocabulary study, even if assigned outside class, a position that Nation (2013) advocates as a matter of course. Browne, Culligan, and Phillips (2013) have updated the academic word list and placed it online with flash-card exercises that provide excellent individual practice, with the potential to not simply stem possible attrition, but actually develop academic vocabulary. While there is a strong argument to be

made that ER is more about fluency, narrative interpretation, and response, a supplementary practice component would likely prove effective in addressing student concerns about losing vocabulary.

Some critics might contend that popular fiction has no place in the university EFL classroom, which demands a more serious approach to content. A focus on academic content involving larger social issues is needed, the argument goes, to prepare for upper-level EMI courses and, after graduation, a professional career in a globalizing world. Certainly this is the ideal, which is possible for students already proficient in English, but intermediate learners who lack fluency and tend to be skeptical about reading in general need an alternative pedagogy that strengthens language skills. Findings from this study suggest that a modified ER approach using popular fiction may in fact prove effective in strengthening fluency and positive attitudes toward both reading and English itself. In sum, learning to read with speed and enjoyment in the EFL classroom can play an important role in preparing first-year intermediate students to deal successfully with upcoming challenges, including English medium instruction.

Acknowledgements

References

Bamford, J. & Day, R.R. (1997). Extensive reading: What is it? Why bother? *The Language Teacher, 21,* 6-8/12.

Barfield, A. (2000). The promise and practice of extensive reading: An interview with George Jacobs and Willy Renandya. *Literacy Across Cultures, 3*(2), 25-30.

Bell, T. (2001). Extensive reading: Speed and comprehension. *The Reading Matrix, 1* (1).

Brown, H. (2014). Contextual factors driving the growth of undergraduate English-medium instruction programmes at universities in Japan. *The Asian Journal of Applied Linguistics, 1*(1), 50-63.

Browne, C., Culligan, B. & Phillips, J. (2013). *The new academic word list.* Retrieved from http://www.newgeneralservicelist.org.

Burns, A. (2009). *Doing action research in English language teaching: A guide for practitioners.* New York: Routledge.

Byng, G. (2002). *Molly Moon's incredible book of hypnotism.* London: Macmillan.

Carty, P., & Susser, B. (2015). Global education and classroom teaching: From CBI to EMI. In P. Clements, A. Krause, & H. Brown (Eds.), *JALT 2014 Conference Proceedings.* Tokyo: JALT.

Casanave, C. P. (1993). Reading and responding: Using journals. In R.R. Day (Ed.), *New ways in teaching reading* (pp. 149-156). Alexandria, VA: TESOL.

Chow, P. H., & Chou, C. T. (2000). Evaluating sustained silent reading in reading classes. *The Internet TESL Journal, 6*(11).

Day, R. R., & Bamford, J. (1998). *Extensive reading in the second language classroom.* Cambridge: Cambridge University Press.

Dupuy, B., Tse, L., & Cook, T. (1996). Bringing books into the classroom: First steps in turning college-level ESL students into readers. *TESOL Journal, 5,* 10-15.

Elley, W. B. (1991). Acquiring literacy in a second language: The effect of book-based programs. *Language Learning, 41,* 375-411.

Hill, D. R. (1997). Setting up an extensive reading programme: Practical tips. *The Language Teacher, 21*(5), 17-20.

Hill, M., & Van Horn, L. (1997). Extensive reading through book clubs: How book clubs have changed lives. In G. M. Jacobs, C. Davis, & W. A. Renandya (Eds.), *Successful strategies for extensive reading* (pp. 98-108) Singapore: SEAMEO Regional Language Centre.

Hino, N. (1988). *Yakudoku:* Japan's dominant tradition in foreign language learning. *JALT Journal, 10*(1), 45-55.

Jacobs, G. M., & Hannah, D. (2009). Combining cooperative learning with reading aloud by teachers. *International Journal of English Studies, 4*(1), 97-118.

Jacobs, G., & Gallo, P. (2002). Reading alone together: Enhancing extensive reading via student-student cooperation in second-language instruction. *Reading Online, 5*(6).

Kanno, Y. (2003). *Negotiating bilingual and bicultural identities: Japanese returnees betwixt two worlds.* NY: Routledge.

Kirkpatrick, A. (2016). The learning and teaching of English as an international language in Asia-Pacific universities: Issues and challenges. In C.H.C. Ng, R. Fox, & M. Nakano (Eds.), *Reforming learning and teaching in Asia-Pacific universities* (pp. 233-247). Singapore: Springer.

Krashen, S.D. (2000). What does it take to acquire language? *ESL Magazine, 3*(3), 22-23.

Krashen, S.D. (1985). *The input hypothesis: Issues and implications*. NY: Longman

Lees, D. & Althomsons, S. (2015). Approaches to extensive reading in the Japanese education system. *Kwansei Gakuin Language Education Research Center Report, 18*, 73-82.

Lyster, R. (2007). *Learning and teaching languages through content: A counterbalanced approach*. Amsterdam: John Benjamins.

MEXT. (2012). *Higher education in Japan*. Tokyo.

Moll, L.C. (1990). Introduction. In L.C. Moll (Ed.), *Vygotsky and education*. New York: Cambridge University Press

Nation, P. (1997). The language learning benefits of extensive reading. *The Language Teacher, 21*, 13-16.

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

Nation, P. (2013). *What should every EFL teacher know?* Compass Publishing.

Nation, P. (2015). Principles guiding vocabulary learning through extensive reading. *Reading in a Foreign Language, 27*(1), 136.

Nation, P., & Wang, K. (1999). Graded readers and vocabulary. *Reading in a Foreign Language, 12* (2), 355-380.

Nishino, T., & Watanabe, M. (2008). Communication oriented policies versus classroom realities in Japan. *TESOL Quarterly, 42*(1), 133-138.

Noguchi, J. (2015). Secondary English education in Japan. In B. Spolsky & K. Sung (Eds.), *Secondary school English education in Asia: From policy to practice* (pp. 33-46). Routledge.

Patton, M. (2015). *Qualitative research & evaluation methods* (4th ed.). LA: Sage.

Prowse, P. (2002). Top ten principles for teaching extensive reading: A response. *Reading in a Foreign Language, 14*(2).

Renandya, W. A., Rajan, B. R. S., & Jacobs, G. M. (1999). Extensive reading with adult learners of English as a second language. *RELC Journal, 30*, 39-61.

Robb, T. N., & Susser, B. (1989). Extensive reading vs. skills building in an EFL context. *Reading in a Foreign Language, 5*, 239-251.

Rosszell, H. R. (2010). Class readers: The learner's perspective. *Global Japanese Studies Review of Meiji University, 2*(1) 51-76.

Sachar, L. (1998). *Holes*. NY: Farrar, Straus and Giroux.

Spack, R. (1997). The acquisition of academic literacy in a second language: A longitudinal case study. *Written Communication, 14* (1), 3-62.

Stoller, F. L. (2015). Viewing extensive reading from different vantage points. *Reading in a Foreign Language, 27*(1), 152.

Susser, B., & Robb, T. N. (1990). EFL extensive reading instruction: research and procedure. *JALT Journal, 12*(2), 161-185.

Swain, M. (1993). The output hypothesis: Just speaking and writing aren't enough. *The Canadian Modern Language Review, 50*, 158-164.

Swain, M. (2000). The output hypothesis and beyond: Mediating acquisition through collaborative dialogue. In J. P. Lantolf (Ed.), *Sociocultural theory and second language learning* (pp. 97–114). Oxford, England: Oxford University Press.

Thomas, D.R. (2006). A general inductive approach for analyzing qualitative evaluation data. *American Journal of Evaluation, 27*(2), 237-246.

Waring, R., & Takaki, M. (2003). At what rate do learners learn and retain new vocabulary from reading a graded reader? *Reading in a Foreign Language, 1*5(2).

Wood, D., Bruner, J. & Ross, G. (1976). The role of tutoring in problem solving. *Journal of Child Psychology and Psychiatry, 17*, 89-100.

Yu, V. W-S. (1997). Encouraging students to read more in an extensive reading programme. In G. M. Jacobs, C. Davis, & W. A. Renandya (Eds.). *Successful strategies for extensive reading* (pp. 1-9). Singapore: SEAMEO Regional Language Centre.

Student Perception on Giving Peer Feedback in EFL Writing and its Implications

Orita Yoshikawa, Kyoko

Abstract

The qualitative case study reported in this paper explored Japanese senior high school students' perceptions on *giving* peer feedback in an EFL writing class. The study unveiled two contrasting findings. The survey data revealed that the students perceived *giving* peer feedback positively and recognized the benefits. The students identified the process of *giving* feedback as an experience of reading non-standard, imperfect English essay actively. The benefits they saw in *giving* feedback were related to the improvements in their own English proficiency: the reviewing of their own English usage, the application into their own writing of what they learned from *giving* peer feedback, and the exposure to the wide variety of English their classmates used. On the other hand, the observational and interview data unveiled the students' irritation and impatience in *giving* feedback on essays written by their peers in imperfect, non-standard English. It depicted the intolerance toward students' own non-standard English. The study stresses the importance of *giving* peer feedback in EFL writing sessions with regards to 1) the students recognizing its benefits, and 2) the students needing to develop tolerance and perseverance in dealing with their own non-standard English.

Introduction

As the communicative approach gained its momentum over the grammar-translation method in the 1990s, the Japanese EFL classrooms began to witness an increased amount of pair and group work. The use of peer feedback was also not ignored. A large amount of research on peer feedback was conducted, varying in its purpose, nature, and results. Some studies emphasized the effectiveness of peer feedback (Keh, 1990; Hedgcock&Lefkowitz, 1992; Tsui&Ng, 2000; Nakanishi, 2006) while others criticized its ineffectiveness

(Conner&Asenavage, 1994; Zhang, 1995).

Despite the differences in their results, most of the research on peer feedback focused on the end product, the *given* peer feedback. Some examined the student feedback *given* on paper in writing and dissected it using different approaches. Others studied the students' reactions to the *given* feedback, while some others studied how much the *given* student feedback was reflected in their second essays. However, when the focus is shifted from the *given* feedback to the process of *giving* feedback, the educational and social benefits unique to peer feedback came into the picture. This is precisely what this research paper is focused on, the students' perceptions on *giving* peer feedback in an EFL writing session, rather than the end product, that is, the *given* peer feedback.

Literature Review

Studies on peer feedback can be categorized into the four types outlined below: 1) students' preferences, 2) comparisons of teacher feedback and student feedback, 3) the constraints of peer feedback, and 4) the benefits of peer feedback.

1) Students' preference of feedback type

There are studies that focused on learners' preference in terms of peer and/or teacher feedback. Zhang (1995) conducted research involving 81 university students in an ESL setting. Her study attempted to examine the changes in the students' preferences as regards teacher and peer feedback when social support was provided in the classroom. She concluded that 93.8% of the students preferred teacher feedback to peer feedback.

Tsui & Ng (2000) observed 27 high school students in an ESL setting to determine the effectiveness of the peer feedback. They attempted to clarify whether secondary L2 learners benefited from peer comments or not. They found that the students preferred teacher feedback, and that the teacher feedback was incorporated a lot more than the peer comments in the revision of their writing. In addition, they stated that the reasons for the preference were based on the teachers' experiences and the directness of the feedback. Regarding the peer feedback itself, they found that 88% of the students stated that *given* peer feedback was more beneficial than *receiving* it.

Moreover, Chaudron (1984) examined 23 university students in an ESL setting and found that nearly all the students preferred teacher feedback over peer feedback. He stated that some students did see benefits of receiving peer feedback, but they doubted the validity of

the peer feedback.

Regarding the validity of the feedback and its relation to the students' preference of the types of feedback, Nakanishi (2006) concluded in her study that the validity and the preference of feedback varied according to the language proficiency level of the reviewers (pp.103-109). She had 13 teachers, 25 university students of higher-English proficiency, and 25 university students of lower-English proficiency review one common composition. She found that the teacher feedback outweighed the student feedback in quantity and in quality. Similarly, the students with higher-English proficiency gave much more direct and coded feedback than those with lower-English proficiency. Her results aligned with the students' preference for the types of feedback they received. In other words, the students preferred teacher feedback to that of the high English proficiency students, while the high English proficiency student feedback was preferred to that of the lower English proficiency students. Likewise, Mendoca & Johnson (1994) stated that the corrections made in the peer feedback tended to arouse doubts on the side of the receivers, especially when their fields of study differed, resulting in an increase in the students' preference for teacher feedback over peer feedback. Interestingly, Sumida (2010) examined two-year college students' perceptions of EFL writing instruction through a survey analysis, concluding that the learners regarded receiving feedback from teachers and peers to be a lot more effective than from teachers alone. His study showed that the students acknowledged the fundamental differences of the two and considered both to be significant. Similar results were reported by Tsui and Ng (2000) and Wakabayashi (2008).

2) Comparisons of teacher feedback and peer feedback

Teacher feedback and peer feedback are fundamentally different. The strength of peer feedback can be found in its social, cognitive and affective aspects all of which cannot be fully supplemented by the teacher feedback. Conversely, the strength of teacher feedback can be identified in its methodological and linguistic aspects (Rollinson, 2005, p.24; Liu&Hansen, 2002, p.8). Teacher feedback and peer feedback complement one another to provide the most effective feedback. Thus, for example, when peer feedback is compared to teacher feedback in its accuracy over grammatical corrections, the effectiveness of teacher feedback naturally outweighs the other. However, despite such fundamental differences, there have been a large number of studies concerned with a comparison of the two.

There is research examining how much influence peer feedback had on the final draft of

students' compositions when compared to the teacher feedback. Connor and Asenavage (1994) reported on eight university students in an ESL environment, showing that approximately 35% of the teacher feedback was reflected in the final drafts, while it was only 5% for the peer feedback (pp.266-267). They pointed out that the majority of the teacher feedback appearing in the students' final draft was of direct corrections.

In a study comparing the effects of different editing methods on EFL writing quality at high school level, Shizuka (1993) found that the quality of the final draft improved the most when the direct corrections were made by the teachers. Nakanishi (2006) compared the features of the teacher feedback and peer feedback in her study involving Japanese university students and reported that teachers made far more "grammatical" corrections than the peers (p.106). Approximately 12% of the teacher feedback focused on the language, while it was only 4% for the peer feedback.

Similarly, Hedgcock and Lefkowitz (1992) compared the final drafts of 30 FSL (French as a Second Language) students that received either teacher or peer feedback. They found that while the final drafts written after the peer feedback tended to improve in their content, structure and vocabulary, those with the teacher feedback exhibited improvements on their language use the most. Shizuka (1993) however, went so far as to question the effectiveness of the direct corrections offered by the teachers. He stated that direct corrections in the teacher feedback tended to appear as they were instructed in the final drafts, but suspected that in the long run, such direct corrections may prevent learners from developing self-feedback skills to revise and improve their own writing (p.20).

In the studies comparing the teacher feedback and peer feedback, the scores of the students' first and revised compositions were often examined. In most cases, the score of the revised composition was better than the first one. Simply put, the "revision" improves the score regardless of the type of feedback the students receive. Nakanishi (2006) conducted a study involving four types of feedback (self-feedback, peer-feedback, teacher-feedback, and peer-&-teacher feedback) and compared the scores of the revised composition. She concluded that the improved scores of the second draft were due to the effect of "revision" rather than the types of feedback the writers received (pp.171-172). Similarly, Yakame (2005) studied the revised composition of students who received peer feedback and concluded that their scores improved because the students simply added sentences to increase the volume of their compositions (pp.108-109).

Furthermore, Sumida (2010), Wakayabashi (2008), and Kambara (2011) suspected the

writing topic selection resulted in the score improvement. They all suggested that the score of the revised compositions increased depending on the topic. Wakabayashi (2008) cited one of the limitations of her study was the failure to adjust the composition conditions for the first and the revised ones (p.106). Thus, regardless of what type of feedback the students received, the score for the second composition increased in general.

3) *Constraints of peer feedback*

Outlined below are some problematic aspects of peer feedback including time constraints, cultural dispositions, and both the students' and teacher's characteristics and beliefs.

Time constraint has been the major obstacle in implementing peer feedback (Rollinson, 2005, pp.25-26). Whether it is given orally or written, the process of giving peer feedback itself is a time consuming activity. In addition, most learners are not accustomed to reviewing papers. Pre-training on how to give feedback may be needed, which naturally demands more time (Liu&Hansen, 2002, pp.151-155; Nakanishi, 2008, p.46). Moreover, usually in secondary education, the EFL curricula have been set in advance, often leaving no room for peer feedback sessions. Yu (2013) stated that Chinese EFL teachers were occupied with aligning their content and schedule around the common term examinations that it was nearly impossible to incorporate peer feedback (p.76).

Collectivism, often a symbol of East Asian cultures, may be a cultural disposition which impedes the effective implementation of peer feedback. In their study, Carson and Nelson (1998) suggested such collectivism to be the reason behind Chinese students' tendency to withhold critical comments. They stated that Chinese students were reluctant to critique their peers in order to maintain group harmony within the class and to avoid exerting authority over their peers. With similar notions in mind, Nakanishi (2006) conducted a study involving 48 Japanese college students. In order to adjust such cultural dispositions, she asked the students to give two kinds of peer feedback, one of which was named and the other anonymous. Her results showed that the anonymous evaluation contained lower scores and more comments on points to be improved. She also stated that the students commented more specifically and directly when their feedback was anonymous. On the other hand, in the feedback including their names, the students gave higher points and commented more on the good points. Their wording was much more polite. Her study showed that conducting peer feedback anonymously allowed the Japanese students to critique their peers more effectively (p.124).

Student characteristics and beliefs often interfere with the implementation of peer feedback. Rollinson (2005) stated that many students might need to be persuaded to value peer feedback (p.26). Some of them might not accept the notion that their peers were qualified to read and critique their writing in place of a teacher. They might feel that only a better writer, in most cases a native speaker or the teacher, was qualified to judge or comment on their work. This is more apparent in a foreign language setting than that of L2 because the exposure to the target language is extremely limited and the only model accepted is the teacher. At the same time, Masaki (1997) stated that some students remained hesitant in critiquing their peers' writing due to their lack of confidence in their English ability.

Teacher characteristics and beliefs, although not often pointed out, may hinder the implementation of peer feedback. Rollinson (2005) stated that teachers often questioned the value of peer feedback and wondered how they could reconcile such a time consuming activity with course or examination constraints. He also pointed out that some teachers had difficulty handing over a significant degree of responsibility to the students. They feared being unable to supervise each and every peer feedback taking place. Moreover, "the teacher may find it difficult not to interfere by providing feedback in addition to that of the student readers, which might well reduce the students' motivation and commitment to their own responding" (Rollinson, 2005, p.26).

Indeed, Yu (2013) conducted her qualitative research with 26 Chinese EFL teachers and found that there was a gap between their beliefs and practice regarding the peer feedback. Although most teachers acknowledged the significance of peer feedback, they did not put it into practice due to various constraints, including their own beliefs. For example, some teachers believed that their students did not possess enough English proficiency to make peer feedback meaningful. According to her study, nearly all the teachers who incorporated peer feedback in their class gave teacher feedback after peer feedback. This not only increased the workload of the teacher, but it clearly illustrated how the teachers thought peer feedback was not sufficient. Her study concluded that EFL teachers lacked the awareness of the value and potential of peer feedback and thus needed to be persuaded.

Moreover, a teacher-centered approach and large class size prevent teachers from implementing peer feedback in their classroom. Zaman and Azad (2012) found that although the Bangladeshi teachers had a very positive attitude towards peer feedback in their study of teachers' perceptions of peer feedback, none of the teachers attempted to incorporate it into their classroom due to the large class sizes (p.150).

4) Benefits of peer feedback

The advantages of peer feedback have often been summarized as follows: Learners receiving more feedback than from the teacher's alone, learners receiving comments from other learners' perspectives, learners gaining audience awareness, learners focusing on ideas for better revision, learners being able to improve their self-revision skills, learners being able to enhance critical reading and critical thinking skills, and learners building a sense of community in the classroom (Chaudron, 1984; Hedgcock & Lefkowitz, 1992; Liu & Hansen, 2002).

One of the most influential benefits is depicted in the Cone of Experience (Dale, 1954, p.43). Dale stated that students retained more in the active participation than through passive participation. Moreover, in defining active participation, he concluded that the students learn more by teaching one another. This is precisely what peer feedback is about. By having students become involved as evaluators, the learning benefits of such students are incomparably better than more passive approaches. Making corrections on peer's essays are typical examples of such benefits. In the process of *giving* peer feedback, the students consult their friends and teacher; they take the initiative in making corrections; and they extend their language retention period through such hands-on experiences.

As is self-evident, the amount of feedback the learners receive increases when peer feedback is added to teacher feedback. This benefit has been reported by Sumida (2010) in his study of Japanese college students. The students acknowledged that receiving feedback from their peers and teacher was more beneficial than the teacher's alone. Unlike the top-down structure of teacher feedback, peer feedback allows the learners to receive comments from the learners' perspectives. It encourages and motivates them in improving their written work.

Liu and Hansen (2002) stated that the benefits of peer feedback are found in its social aspect (see Table 1). They stated that peer feedback enhanced the communicative power of learners, allowed them to receive authentic feedback, and established collegial ties and friendship. Learners may also gain audience awareness through peer feedback (p.8). Keh (1990) asserted that using peer feedback allowed learners to gain a greater sense of audience when their text was reviewed by several readers, instead of just by one teacher (pp.299-303).

Audience awareness enables the learners to conduct self-feedback and self-assessment at every stage of the writing process. In other words, they are able to revise and edit their work as it progresses from a reader's point of view. Rinnert & Kobayashi (2001) also noted

that peer feedback required leaners to become more aware of their own writing style, and simultaneously, forced them to guess where readers might have problems reading their writing (p.203).

Table 1 Benefits and Constraints in Using Peer Response (Liu & Hansen, 2002; p.8)

	BENEFIT	CONSTRAINT
Cognitive	1. Exercise thinking 2. Take active role in learning 3. Engage in exploratory talk 4. Build critical skills 5. Demonstrate and reinforce knowledge 6. Build audience awareness	1. Uncertainty concerning peers' comments 2. Lack of learner investment
Social	1. Enhance communicative power 2. Receive authentic feedback 3. Gain confidence and reduce apprehension 4. Establish collegial ties and friendship 5. Influence learners' affective state	1. Discomfort and uneasiness 2. Lack of security in negotiating meaning 3. Commentary may be over critical
Linguistic	1. Enhance metalinguistic knowledge 2. Explore linguistic knowledge 3. Gain additional language skill practice 4. Enhance participation and improve discourse 5. Find right words to express ideas	1. Too much focus on surface structure 2. Lack of L2 formal schemata 3. Difficulty in understanding foreign accent
Practical	1. Applicable across students proficiency levels 2. Flexible across different stages in the writing process 3. Time-efficient in some cases 4. Reinforces process writing	1. Time constraints 2. Counter-productive feedback 3. Lack of student preparation

To conclude this review of current literature regarding peer feedback, we can see that most of the research introduced here focused on the *given* feedback. Comparisons between teacher feedback and student feedback were made, often resulting in the teacher feedback outweighing the other in quality and quantity. Such studies have failed to focus on the process of *giving* peer feedback. They disregarded how the students felt as they gave their feedback and what they thought of the whole process. The practical research case study reported in this paper aims to reveal the students' perception of *giving* peer feedback and find out its implications.

Practical Research

Methodology

(1) Participants

The participants in this research were 73 first-year senior high school students. The academic level of this school against the national standard was considered to be high with nearly all the graduates advancing to a four-year university. The students were divided into two levels according to their English proficiency (intermediate and advanced), and all the participants of this study belonged to the intermediate level. All the participants took four 50-minute EFL classes per week taught by the researcher.

(2) Procedure

First, the participants wrote English compositions on a given topic: *my turning point in life*. The compositions were collected in class by the instructor. Later, the instructor photocopied each composition onto a peer feedback sheet with the writers' names hidden (see Appendix A). Second, these photocopied sheets were handed out to be evaluated by the participants. They gave feedback directly onto the peer feedback sheet in class using red pens. Third, the English compositions with the peer feedback were each returned to the writers, who then revised their compositions and submitted their second drafts in class.

(3) Instruments

As described above, four instruments were used: peer evaluation sheet, questionnaire, interview notes, and observation notes.

  i) Peer evaluation sheet

   The peer evaluation sheet consisted of four parts. First, the participants twice summarized the compositions they read – once in 50 Japanese characters or less, and once in English. They then made changes and comments directly onto the composition referring to a proofreader's guide. Next, they evaluated the composition according to five criteria - content, organization, vocabulary, language use, and mechanics – using a scale of four. Finally, they gave comments and advice to the writer in Japanese (see Appendix A).

  ii) Questionnaire

   The questionnaire was given to the students when they submitted their second draft. Nearly 10 minutes were allotted for the participants to answer two questions.

    a) Do you think *giving* peer feedback on EFL composition is beneficial to you?

    b) What is your reason for answering YES or NO to the above question?

iii) Interview notes

Interviews were conducted with the five participants who answered that they felt that *giving* peer feedback on EFL composition was *not* beneficial to them. All the interviews were conducted separately under very casual settings. The interview time varied from five to ten minutes and was all conducted in Japanese.

iv) Observation notes

The students were observed under two circumstances: when giving peer feedback, and when writing up their second draft. The observer was the instructor. Notes were made after class based on the recollection of the observer.

Analysis

The responses to the questions on the questionnaire were analyzed according to the qualitative data analysis scheme – data reduction, data display, and conclusion drawing and verification (Miles & Huberman, 1994, pp.55-66).

Results

(1) Questionnaire

Question #1　Do you think *giving* peer feedback is beneficial to you?

Of the 73 participants, 67 answered that *giving* feedback was beneficial to them. Five answered that it was not beneficial to them; and one answered "neither".

Question #2　What is your reason for answering YES or NO to Question #1?

For those who answered *yes*, the improvement in their overall English skills appeared the most as the benefit of *giving* peer feedback. In order of perceived benefit, they were 1) English composition skill, 2) grammatical skill, 3) reading skill, and 4) vocabulary skill (see Appendix B).

First, regarding the English composition skill, the participants replied that *giving* peer feedback enabled them to learn from peer's writing and mistakes, compare their own work to those of peers, and review their own writing. They further mentioned that they could imitate peers' writing in their second draft.

Second, the participants felt that *giving* peer feedback strengthened their English grammar. In order to confirm if the grammatical corrections they made on the peer feedback sheet were correct or not, they had to consult dictionaries, grammar books, and other classmates. Moreover, they mentioned that they became more careful so as not to make the similar

grammatical mistakes as their peers.

Third, as for the reading skill, the majority of responses reflected how their proofreading of the compositions was carried out attentively and carefully. It appeared that reading imperfect compositions written in "raw" English without any set of answers required much attention on the part of the reviewers.

Fourth, regarding vocabulary skills, the participants mentioned checking the spellings, witnessing peers using vocabulary they had never used before, and actually using a set of vocabulary they had learned from *giving* peer feedback onto their second draft.

Aside from the general English skills, other responses by those who answered *yes* included learning from comparing one's writing with those of peers, being able to realize one's own mistakes, being motivated by the idea of peers reading one's work, and being able to provide detailed feedback.

First, regarding learning from the peer comparison, the participants noted that they learned where they stood among their peers in regards to their English proficiency. Some responses pointed out the surprise of discovering how much their own and peers' writing differed in quantity, quality, and topic.

Second, as for being able to notice their mistakes, the participants discovered some patterns in mistakes made by their peers; and those patterns were taken into account when they wrote the second draft. They also learned how misinterpretations occurred in EFL writing. The majority of them pointed out that *giving* peer feedback made them conduct self-checks, or self-feedback, while writing.

Third, some participants thought that *giving* peer feedback motivated them in writing English compositions. They said that it was inspirational that somebody actually gave feedback to their writing, just as they were giving feedback. The last point was being able to provide individualized, detailed feedback. Unlike the teacher feedback, the participants felt that they were giving more individualized feedback with, for instance, grammatical explanations and examples.

For the five who answered that *giving* peer feedback was not beneficial, the main reason given was their own lack of English proficiency and therefore ability to make meaningful judgments. They further pointed out that even if they could locate a mistake, they did not possess enough knowledge to make necessary correction or give advice.

(2) Interview

The five participants who answered *no* to the questionnaire were then interviewed casually in order for them to clarify their reflections. Of the five, two had low-intermediate level of English proficiency, while three were high-intermediate.

All of them reiterated that they lacked enough knowledge of English to "correct" their peers' compositions. The low-intermediate participants stated that they could not locate mistakes and problems within the composition. They said that their time was spent merely trying to figure out what was said in English. Thus, they felt extremely useless in the process of giving peer feedback. Those with high-intermediate English proficiency felt that none of the students in class possessed enough English knowledge to give feedback to another student. One felt that it was a waste of time to give feedback when neither the reader nor the writer knew the "right" English.

Regarding the peer feedback itself, one participant of low-intermediate proficiency stated that he feared he might be "contaminated" by non-standard, incorrect English. Since he lacked a basic knowledge of English, he worried that he might accept the grammatical mistakes made by peers as correct, and end up learning them in such a way. One participant with high-intermediate proficiency insisted that the teacher feedback was the only feedback he wanted. He saw no significance in peer feedback.

As for the method of incorporating peer feedback into the EFL writing classroom, one participant said that it was difficult to give and receive peer feedback only on paper. He said that it was meaningless without an oral explanation and that he could not make corrections or advise well enough on paper. By the same token, he could not understand the corrections and advice given to him on his own paper and wished to ask the peer reviewer directly, face to face, for more details.

(3) Observation

The participants were observed under two circumstances: first, during the procedure in which the participants gave peer feedback to anonymous writers, and second, during the process when the writers began working on their second draft, after receiving the anonymous peer feedback of their first draft.

During the process of *giving* peer feedback, the majority of the participants worked individually following the peer feedback guidelines at first. Once they completed the whole process, they began to consult those sitting nearby for advice on certain points they could

not give feedback on. Meanwhile, some of the participants, regardless of their English proficiency, reacted openly to peers' essays. Some were surprised at the level of English they displayed. Some were irritated by being unable to decipher an essay written in non-standard, grammatically-flawed English. The positive and negative comments by the participants while *giving* peer feedback were in such terms as (researcher's translation): *"Wow!" "This guy knows how to use English." "This writer should be in the advanced class." "What does he want to say?" "How can you read it, Sensei? It's incomprehensible!" "Native speakers would never say it this way, right?"* (see Appendix C for remarks in Japanese). As the class period approached its end, many participants voiced the need for more time to give peer feedback.

When the anonymous peer feedback sheets were returned to the students, some students reacted emotionally, verbalizing their feelings. Some criticized the anonymous peer reviewers for pointing out the mistakes and attempted to prove that their peer reviewers had wrongfully pointed out the mistakes. It seemed that the participants were fighting back at the unseen anonymous reviewers. Some students appeared to be threatened by the given peer feedback and seemed desperate in trying to justify the mistakes pointed out in the feedback sheets. One participant nearly begged his instructor for an approval that his paper was sufficient and that the peer reviewer made unnecessary revisions. The participants were making full use of dictionaries and English textbooks, and consulting the teacher and classmates in writing up their second draft. Some of the comments expressed out loud by the participants when reading the anonymous peer feedback sheets were: *"Why can't you get it?" "You should be able to read it and understand it." "I checked the dictionary. It can't be wrong!" "What I wrote is right, isn't it?"* (see Appendix C for remarks in Japanese).

Discussion

After analyzing the above results using qualitative data analysis scheme (Miles and Huberman, 1994), the study raises the following three main characteristics, further discussed below: 1) the benefits and constraints students saw in *giving* peer feedback, 2) the students' definitions of *giving* peer feedback, and 3) the students' reactions to anonymous peer feedback.

(1) The benefits and constraints students saw in *giving* peer feedback

The majority of the participants (67/73) answered that giving peer feedback was beneficial to them. This is similar to the finding by Tsui & Ng (2000) in which 88% of the participants

thought *giving* peer feedback was more beneficial than *receiving* it (p.158). On the other hand, less than one-tenth of the participants (5/73) answered that giving peer feedback was not beneficial to them. The reasons given by the participants aligned with the benefits and constraints of peer feedback summarized by Liu and Hansen (2002, p.8) (see Table 1). The findings are significant in that they revealed that the students recognized and acknowledged the benefits and the constraints of *giving* peer feedback.

First, in all the responses noting benefits, cognitive, linguistic, and practical aspects are evident. Cognitively, some participants mentioned of gaining audience awareness. For example, one wrote, *"through the process, I can see what kind of writing is easy to comprehend, and on the other hand, what kind of writing causes confusions for the readers."* Others touched upon taking active roles in learning: *"I had to use all my resources in giving peer feedback, and I think the process contributes to improving my grammatical skills in return."* Linguistically, the students found benefits in improving their English skills. Some referred to gaining additional language skills. For instance, one wrote, *"I discovered some new English expressions."* Another wrote that he noticed a certain tendency for grammatical mistakes while giving peer feedback, and hoped not to repeat similar mistakes in his own writing: *"[Through the process of giving peer feedback,] I get to see what kind of grammatical mistakes are made often, and it helps me when I write."* As for practicality, one of the students wrote: *"Regardless of the language proficiency of the peers, we can all learn something through giving peer feedback. From those who are good at English, we learn how to construct the composition. From those who are not good at English, we can learn the solutions of improving not-so-good compositions."*

Second, the responses illustrating the constraints of peer feedback could all be witnessed in cognitive, social, linguistic, and practical aspects. Cognitively, two participants of high-intermediate English proficiency expressed uncertainty concerning peers' comments: *"I believe none of us has enough knowledge of English skill to give peer feedback. I doubt the appropriateness of the feedback I give, and of that I receive."* Socially, some participants with lower-intermediate English proficiency expressed discomfort and uneasiness. One wrote, *"I hate giving peer feedback because I don't have any confidence in my command of English and I can neither point out the mistakes nor correct them."* Linguistically, one participant focused too much on surface. He said that he was hesitant in marking the paper directly because he wasn't sure what he wrote would be right. Another student said in the interview, *"I don't want to pick up some incorrect English knowledge through reading peers' writing. I*

don't want to be 'contaminated' by non-standard English." As for a practical aspect, a majority of the participants pointed out the time constraints. They stated that they needed much more time in giving peer feedback: *"I couldn't finish giving feedback. One class period is not enough."*

(2) The students' definitions of *giving* peer feedback

The results of the questionnaire revealed how the participants defined the process of *giving* peer feedback (see Appendix D). Two definitions emerged from the results: 1) pointing out the mistakes and 2) reading imperfect English. The significance of these definitions is that it made it apparent that the students were intolerant toward non-standard English writing, and that a realistic depiction of present-day English language was essential in an EFL writing class.

First, more than one third of the participants (27/73) defined the process as *"pointing out the mistakes"*. For example, one wrote, *"I can improve my grammatical knowledge because I read to find the mistakes."* Another wrote, *"[The process of giving peer feedback] is like the error-recognition questions on the entrance exam."* Of these participants, one-third (9/27) additionally answered *"making corrections"* as part of pointing out the mistakes. Interestingly, three of them (3/9) were the participants who responded that giving peer feedback was not beneficial. One of the responses reads as follows: *"Our English proficiency is not so high. So, the process of making us point out the mistakes accurately and expecting us to correct them is too time consuming, and it is even doubtful that the corrections are right."* Rather than defining the process merely as "pointing out the mistakes," these participants placed greater expectations on the function of peer feedback as simultaneously offering corrections. On the other hand, those who responded positively said that they could strengthen their grammar as they needed to check to make sure the corrections they were making were correct: *"By pointing out the mistakes and making corrections, the process is inscribed in our memories. So, we don't repeat the same mistakes. Through it, we are reviewing our own grammar knowledge unintentionally, but aggressively."*

Second, nearly one-third of the participants (23/73) defined the process as *"reading other people's writing."* Under this definition, the most commonly expressed feature was *"reading imperfect English"* followed by *"reading ambiguous English,"* and *"reading various kinds of English."* One wrote, *"I learned a lot from giving peer feedback because reading composition written in imperfect English was a new situation for me."* Another wrote, *"I am reading*

and evaluating a composition with no correct answers." Others wrote that they had to read carefully and attentively because compositions were written with *"raw English"* unlike those in English textbooks. This definition indirectly confirmed the lack of student exposure to varieties of non-standard English writing. It showed that they were exposed only to texts written in Standard English, and thus reading non-standard English compositions left a strong impact on the participants, which was closely reflected in their definition of the process of giving peer feedback.

(3) The students' reactions to anonymous peer feedback

The students' reaction to anonymous peer feedback unveiled their intolerance towards non-standard English writing, while showing the effectiveness of anonymity in overcoming certain cultural dispositions, especially collectivism.

Some of the expressions heard in the classroom during the process of giving anonymous peer feedback highlighted the non-standard English. However, after receiving anonymous peer feedback, comments expressed frustration at the feedback and looked for support as their own non-standard English came under attack.

The students tried to justify their writing by validating it with the use of dictionary, consulting their instructor, their peers and counter-attacking the anonymous reviewers. What all this illustrated was that the students, although non-native speakers of English and still at a relatively early stage of learning English, lacked tolerance and/or immunity towards non-standard English writing.

The effectiveness of anonymity over the culturally-disposed barrier of collectivism also stood out in the observational study. The participants were able to maintain group harmony within the class and were not forced to take authority over their classmates. They could critique the composition objectively without worrying about any after-effects of their reviews. This allowed the peer feedback to be implemented effectively, fulfilling its purpose. Nevertheless, it is true that anonymity did bring about some over-critical reviews, impolite wording, and insufficient responses. Despite this, the anonymity in this study can be said to have reduced the cultural barrier against the effective implementation of peer feedback in EFL writing classrooms.

Conclusion of Practical Research

Some limitations of this study could be found in its inapplicability to other situations, the students' tendency to answer in the favor of their teacher/researcher, and the limited amount

of interview data as a qualitative research. A longer term study would be beneficial so as to see how the students' attitudes developed over time as they improved their English ability, and became more accustomed to the peer feedback method.

However, consideration of the results of this practical research pointed out a strong need for incorporating peer feedback in EFL writing sessions for two apparent reasons. The first one is the students' recognition of the benefits of *giving* peer feedback as reflected in the questionnaire and supported by the observational data. The second is the need for a realistic representation of English in EFL classrooms. The students' lack of tolerance was shown in how they defined the process of giving peer feedback, and how they reacted in giving and receiving peer feedback on their non-standard English writing. This study made it clear that incorporation of *giving* peer feedback in EFL writing sessions is crucial because students do indeed acknowledge the benefits of peer feedback. There also seems to be a need for the students to develop tolerance and immunity towards non-standard English writing as the majority of English speakers in our present-day world are non-native speakers whose English may be far from standard.

Conclusions and Implications

Based on the findings from this study, there are two implications for teachers incorporating peer feedback into their EFL writing lessons: 1) the need to acknowledge students' recognition of the benefits of *giving* peer feedback, and 2) the need to develop students' perseverance and willingness to communicate even when faced with non-standard varieties of English.

The constraints of peer feedback, as mentioned in the literature review, weigh so heavily on the teachers that it seems nearly impossible to conduct peer feedback in class regularly. However, as shown in the study, the students are well aware of the benefits of *giving* feedback to their peers. They find *giving* feedback strengthens their own English skills. Furthermore, the students need to be exposed to non-standard varieties of English to develop tolerance and acceptance toward their own non-standard English, and to nurture perseverance and willingness to communicate even when faced with such non-standard English. This is because non-standard English is prevalent in our present-day society when the English language is used as a lingua franca and when the majority of its speakers are non-native speakers. The students' rejection of non-standard English is indeed a rejection of their own non-standard, developmental-stage English. It is crucial that the English language is presented realistically

in accordance with the present-day society, and that our students possess the means and the will to communicate with others even when faced with non-standard varieties of English.

In conclusion, this paper examined the students' perceptions on *giving* peer feedback. Through the practical research, it was found that the students recognized the benefits of *giving* peer feedback, while being intolerant toward essays written in non-standard English. Based on the findings, the study concludes that the incorporation of peer feedback is essential in EFL writing classes. Teachers need to acknowledge that their students themselves recognize the benefits of *giving* peer feedback, and that their students need to obtain the means and willingness to communicate even when encountering non-standard English.

References

Carson, J.G., & Nelson, G.L. (1998). ESL students' perceptions of effectiveness in peer response groups. *Journal of Second Language Writing,* 7(2), 113-131.

Chaudron, C. (1984). The effects of feedback on students' composition revisions. *RELC Journal, 15*, 1-16.

Connor, U., & Asenavage, K. (1994). Peer response groups in ESL writing classes: How much impact on revision? *Journal of Second Language Writing, 3*(3), 257-276.

Dale, E. (1954). *Audio-Visual Methods in Teaching*. New York: Holt, Rinehart and Winston.

Hedgcock, J. & Lefkowitz, N. (1992). Collaborative oral/aural revision in foreign language writing instruction. *Journal of Second Language Writing,1*(3), 255-276.

Kambara, K. (2011). Writing katsudo niokeru peer response to kyoshi feedback no koka [Effects of teacher feedback and peer response in an writing activity]. *Eiken Bulletine, 23*, 127-139.

Keh, C.L. (1990). Feedback in the writing process: a model and methods for implementation. *ELT Journal, 44*(4), 294-304.

Liu, J. & Hansen, J.G. (2002). *Peer response*. Ann Arbor: The University of Michigan Press.

Masaki, M. (1997). Peer feedback in a Japanese EFL writing classroom. *JACET Kansaishibu Daiyonji (1995-1999) Kenkyu Project Writing Shido Kenkyukai Kiyo, 3*, 11-18.

Mendoca, C.O. & Johnson, K.E. (1994). Peer review negotiations: Revision activities in ESL writing instruction. *TESOL Quarterly, 28* (4), 745-769.

Miles, M.B. & Huberman, A.M. (1994). *Qualitative data analysis*. Thousand Oaks, CA: Sage Publications.

Nakanishi, C. (2006). *A teaching approach to Japanese college students' EFL writing*. Tokyo: Keio University Press.

Nakanishi, C. (2008). The effects of trained self-feedback on revision. *JACET Journal*, 47, 35-49.

Rinnert, C. & Kobayashi, H. (2001). Differing perceptions of EFL writing among readers in Japan. *The Modern Language Journal,* 85(2), 189-209.

Rollinson, P. (2005). Using peer feedback in the ESL writing class. *ELT Journal*, 59(1), 23-30.

Shizuka, T. (1993). Effects of different editing methods on EFL writing quality at high school level. *JACET Bulletin, 24*, 139-158.

Sumida, A. (2010). College students' perception of EFL process writing instruction: A survey analysis. *Niigata Seiryo Daigaku Tankidaigakubu Kenkyu Houkoku*, 40, 157-164.

Tsui, A.B.M., & Ng, M. (2000). Do secondary L2 writers benefit from peer comments? *Journal of Second Language Writing, 9*(2), 147-170.

Wakabayashi, R. (2008). The effect of peer feedback on EFL writing: Focusing on Japanese university students. *OnCUE Journal, 2*(2), 92-110.

Yakame, H. (2005). The role of peer feedback in the EFL writing classroom. *Annual Review of English Language Education in Japan, 16*, 101-110.

Yu, S. (2013). EFL teachers' beliefs and practices regarding peer feedback in L2 writing classrooms. *Polyglossia, 24*, 74-79.

Zaman, M., & Azad, A.K. (2012). Feedback in EFL writing at tertiary level: Teachers' and learners' perceptions. *ASA University Review, 6*(1), 139-156.

Zhang, S. (1995). Reexamining the affective advantage of peer feedback in the ESL writing class. *Journal of Second Language Writing, 4*(3), 209-222.

Appendices

Appendix A   Sample peer feedback sheet

1. Summarize the essay in Japanese.　50文字前後に要約

(grid of empty cells for writing)

2. Summarize the essay in English.　英文 1 文に要約

3. Give feedback directly on the essay.　（直接書き込む、赤ペンで）
※下線を引くだけや「？」を書き込むだけは避け、具体的なフィードバックを！
校正記号例：

削除	~~the~~	校正前に戻す	~~the~~ ⓞⓚ	大 / 小文字	大 小	スペース	␣
挿入	⌃	段落変更	¶	文法ミス	Ⓖ	スペリング	ⓢⓟ

4. Evaluate the essay by circling the appropriate rating.　※「評価」というよりも「印象」

内容	1 (very poor)	2 (poor to fair)	3 (average to good)	4 (good to excellent)
構成	1 (very poor)	2 (poor to fair)	3 (average to good)	4 (good to excellent)
語彙	1 (very poor)	2 (poor to fair)	3 (average to good)	4 (good to excellent)
文法	1 (very poor)	2 (poor to fair)	3 (average to good)	4 (good to excellent)
形式	1 (very poor)	2 (poor to fair)	3 (average to good)	4 (good to excellent)

5. Write a comment addressed to the writer.　日本語・英語可

Appendix B   Students' perceived benefit of *giving* peer feedback

	Code	No. of Responses
English-skill related	English composition skill	15
	Grammatical skill	14
	Reading skill	12
	Vocabulary skill	5
Non English-skill related	Comparing one's writing with those of peers	10
	Being able to realize one's own mistakes	8
	Being motivated by the idea of peers reading one's work	2
	Being able to provide detailed feedback	1

Appendix C   Student remarks in Japanese

Observation 1: **Student remark when *giving* anonymous peer feedback**	Observation 2: **Student remark when *receiving* anonymous peer feedback**
「こいつスゲーなぁ」 「上級だよ！」 「何が言いたいの？」 「ありえない！」 「先生、これよく読めますね。大変だ。」 「ネイティブはこう言いませんよね。」	「なんでわからない？」 「フツウ、わかるだろう！」 「辞書にそうあった。ちゃんと調べた。」 「間違えてないですよね？」 「俺、あってますよね？」

Appendix D   Students' definitions of *giving* peer feedback

Code	Definition: Giving feedback means…	No. of Responses	Responses to Q1		
			Yes	No	Neither
Pointing out the mistakes	to look for mistakes	18	17	0	1
	to make corrections of the mistakes	9	6	3	0
Reading other people's writing	to read imperfect English	14	14	0	0
	to read ambiguous English	8	8	0	0
	to read various kinds of English	3	3	0	0
No definition	to give feedback (redundant response)	4	4	0	0
	nothing mentioned	17	15	2	0

Motivation of Japanese Elementary School EFL Learners Illuminated through the Test of English as a Foreign Language (TOEFL®)

Ohgi Ito

ABSTRACT

This article reports on the practice of an English Language Teaching (ELT) experiment at an elementary school in Japan, employing an international standardized test as a stimulator in a private school context. The introduction of assessment tasks and the implementation of the Test of English as a Foreign Language (TOEFL) Primary towards 11 to 12-year-old students revealed a strong learning motivation of young learners and their enthusiasm for the practical use of the foreign language. Although the discussion has been continuing about whether or not English language should be taught at an earlier age, the field of Teaching English to Young Learners (TEYL) is increasingly expanding, and the government policy has set the directions of elementary school English to be started at a lower grade. This article illustrates the impact of assessment opportunities and highlights the motivation of Japanese EFL (English as a Foreign Language) learners. The power of social values, including family influence, is observed and the experimental classes exhibit the strong desires of Japanese young learners to communicate in English however limited their abilities may be. The results could potentially contribute to the development of TEYL in Japan and suggest the need for further longitudinal research into motivational transition from elementary to secondary school English education.

1. Introduction

This study was undertaken with the aim to shed light on Japanese EFL learners' motivation in undergoing assessment and their views on learning English in an elementary school context. In Japan's education scenario, there has been a discussion about whether or not ELT should be implemented at an earlier age. Since the announcement of the *Action Plan*

to Cultivate "Japanese with English Abilities" by Japan's Ministry of Education, Culture, Sports, Science and Technology (MEXT 2003), the field of TEYL has occupied the attention of Japanese people. Traditionally in Japan, ELT had relied heavily on the grammar-translation methods, and it was not until the mid-1980s that people started to call for the need for 'English for communication' and the method of Communicative Language Teaching (CLT) gained ascendancy in the 1990s. Those changes were brought on by the growth of linguistic influence in international relations, media, travelling and communications (Crystal 2003:86-122) and the social context has influenced the direction of ELT in Japan. The field of TEYL has been no exception to this trend, and many express their belief that the 'earlier is better' in child second language (L2) acquisition. The number of young learners has been growing, and the government most recently proclaimed the English education reform plan (MEXT 2014) and officially announced that English language classes would be taught from Year 3 (age 8-9) in all state schools from 2020.

Meanwhile, private schools have a different background with a long history of ELT, and the institution in this article is a private elementary school in Tokyo with children from Year 1 (age 6-7) to Year 6 (age 11-12), who take English classes once or twice a week as part of their regular school subjects. From a teacher-researcher's point of view, ELT has brought positive outcomes and strong learning motivation can be observed, although various factors, e.g. individual interest, aptitude, intellectual curiosity and influence from family or peers, seem to be deeply connected with their motivation. Interpreting this highly complex EFL setting can illuminate what plays an important role in stimulating elementary school learners and in what aspects the objective measurement has potential for TEYL. The experiment first commenced with the idea that the impact of an assessment could motivate the young EFL learners for further development as well as provide the teachers with meaningful factors about learning motivation.

For this experiment, a case study method is employed in order to portray what it is like to be in a particular EFL situation and 'to catch the close up reality' (Cohen *et al.* 2007:253). Empirical research is conducted on 18 Year 6 students with research tools such as observation and questionnaires, and the data collected from the TOEFL Primary attempt to illuminate the enthusiasm of Japanese elementary school EFL learners for acquisition of English as a foreign language.

2. Context

For this research experiment, it is essential to give full descriptions of background context, particularly of TEYL in Japan, and to throw light on the layers of context. Crystal (2003:59) identifies two aspects for the present-day world status of English as 'the expansion of British colonial power, which peaked towards the end of the nineteenth century, and the emergence of the United States as the leading economic power of the twentieth century', both of which can be evidenced in the history of ELT in Japan. From the mid-1980s through the 1990s, the method of CLT started to gain ascendancy, and the primary goal of ELT became 'fostering communication skills' for global stages (MEXT 2003). Some critics put in question the Japanese perception towards English language learning that reflects their belief in 'communicative English for intercultural communication' (Kubota 2002; Sakui 2004), and the changes to the trend for CLT have apparently brought confusion to schools; Sakui (2004:155) describes metaphorically that CLT in Japan is somewhat like 'wearing two pairs of shoes', taking on 'a parallel situation' where the teachers keep the grammar-translation approach for university examinations, while teaching 'communicative English' as required by the government. These historical and social backgrounds of ELT are deeply entrenched in Japan, and the field of TEYL began to flourish in the 1990s with great expectations from adults about young learners' potential for 'communicative' competence.

2.1 Teaching English to Young Learners (TEYL) in Japan

As indicated, the social context has influenced the direction of foreign language education in Japan, including TEYL, and the number of young EFL learners has been growing in spite of the fact that Japanese children have extremely limited exposure to foreign languages. There has been a belief that the 'earlier is better' in child L2 acquisition, and the government decided to introduce English language in all state schools from Year 5 (age 10-11) in 2011 and from Year 3 (age 8-9) in 2020 (MEXT 2014). The decision inevitably affected the whole ELT in the school context and raised issues from both extremes. Critics such as Otsu (2004) and Torigai (2006) argue that Japanese children are not ready to learn a foreign language and that early learning of a foreign language might disturb the acquisition of their first language (L1). Others fear that TEYL might take away the 'identity' as Japanese and suggest that ELT should be combined with other subjects such as social studies or intercultural understanding at school (MEXT 2006). Content and Language Integrated Learning has been advancing in Europe; however, this approach cannot be easily adopted into the environment where foreign language

exposure is limited and comprehensible input is sparse. Among these negative perspectives, Tomita (2004:18-20) advocates that linguistic input is not required for young EFL learners and that their 'awareness' to unknown, new things around the world should be the goal for TEYL. It has been a concerning issue what can be done to facilitate the foreign language learning of young learners in a 'low-immersion' context like Japan.

Under these controversial circumstances, TEYL has become a reality in Japanese schools and the focus is now moving towards what age would be the best to introduce English language as a school subject. In private schools, on the other hand, ELT has long been implemented from younger ages, since these schools can select their own educational policies with minimal control from the government. In the early years of CLT, the effects of TEYL activities such as chants, games and short plays were highlighted and researched mainly; however, the situation has changed, and several studies have revealed linguistic and emotional effects of ELT, such as Nishida (2013:93-109), who investigates motivational differences in age and gender among Japanese school children. At present, the link between elementary and secondary English education has become a new topic for discussion (Higuchi *et al*. 2008:58-69). Therefore, TEYL is still in the middle of pursuing methods and materials, and empirical studies are rather thin on the ground, remaining unexplored as yet.

2.2 English Language Teaching (ELT) in this institution

The institution in which this experiment was conducted is a private elementary school in Tokyo, and since the foundation in 1874, ELT has been implemented with original materials and methods. Learning English is a part of the school curriculum from Year 1 (age 6-7) and the class size in upper grades is relatively small for a Japanese school, dividing a class of 36 students into three groups (12 students each for English classes) from Year 4 (age 9-10) onto Year 6 (age 11-12). There is a unique assessment for the upper grades, offering the students speaking and writing sessions, and the whole course of study has been positively accepted, although educational research into learning motivation has been scarcely administered.

The Year 4 English generally carries an enthusiastic class atmosphere, but the dynamics start to change as the students grow physically and mentally; competitiveness between 'fast' and 'slow' learners arises, and parental influence on achievement becomes remarkable. Some parents provide children extra classes outside school, and others express concerns about children's achievement in the reports. During Year 4 and Year 5, the main focus of teaching is to get the students familiar with English sounds through chants and simple expressions,

not so much with vocabulary or structure. The Year 6 English classes introduce syntax with new vocabulary by degrees so that the students can prepare for secondary school English; in Japanese schools, English language formally becomes a required subject from the secondary education. Still, the Year 6 students are vigorous and energetic in learning English, thanks to their curiosity about different languages and cultures, and the milieu of a small class size enables them to speak out without hesitation. From Year 4 onward, the students learn English twice a week (40 minutes each), which is twice as much as state school students learn, and the proficiency level of the Year 6 students is assumed to be high, compared to other average Japanese school children. It should be noted, however, that their language abilities are varied and individual differences in motivation and competence are immense. In addition, parents play a crucial role in motivation with some, for example, believing that earlier learning would result in better achievement of 'communicative competence'. The learning environment greatly depends on each student's background, particularly in the private school context as in this article.

2.3 The Test of English as a Foreign Language (TOEFL) and TOEFL Primary for EFL learners in Japan

The TOEFL is a standardized test to measure the English language abilities of non-native speakers who wish to study at English-speaking universities or institutions, and it has been taken and accepted globally for many years. The test is designed and administered by the Educational Testing Service (ETS) in the United States and the official score report is received as a certificate of applicants' English abilities by major academic and professional institutions around the world. In 2005, the internet-based test (iBT) was introduced and gradually replaced the paper-based test and the computer-based test. The iBT format consists of four sections, measuring four language skills, i.e. Reading, Listening, Speaking and Writing sections, and it requires comprehension and integrated communicative skills as well as academic knowledge for higher education. A significant feature of iBT is the quality and quantity of its content as the four-hour test is conducted by means of a personal computer with a headset and a microphone. Integrated skills are measured, for instance, in the Writing section, the test takers first read a passage, then listen to a related lecture, and finally summarize the whole content, typing within a limited time frame and word amount. In the Speaking section, various topics are indicated and some require the test takers to read a short passage, then to listen to a conversation, and finally to express their opinions or suggestions

on what has been previously presented.

According to *Test and Data Summary for TOEFL iBT® Tests* (ETS 2015), the score means (average score) of each country is published as a guide every year, and the data is primarily for comparing the performances of individual test takers in the same country, not for comparing countries by rank. However, the results provoke concerns in non-native English-speaking countries, including Japan, because the iBT format appears to be difficult and the low score might become a barrier for EFL learners to study abroad. With such a background, the TOEFL Junior for learners from age 11 and the TOEFL Primary from age 8 were introduced to meet the needs of young EFL learners around the world. Both tests intend to measure English language skills, guide to the next step of learning, and help both teachers and young learners to identify their strengths and weaknesses. The scores are mapped to the Common European Framework of Reference (CEFR) levels released by the Council of Europe (2001), and both tests enable the young learners to get accustomed to the test format and receive an appropriate measurement of their linguistic abilities. These junior versions of the TOEFL have been introduced in Japan, although the test opportunities are limited and the contents are not yet well-known.

The experiment in this article started with this background, reflecting the voices of former students of the researcher who intended to study abroad, but struggled to achieve the TOEFL scores high enough to be accepted to the overseas universities in which they wished to enroll. The experiment, therefore, has a mission to highlight learners' motivation as well as to enhance EFL learning by inclusion of an extracurricular assessment opportunity at school in Japan.

3. Literature Review

3.1 Research methodology

When research involves children who cannot be regarded as being on equal terms with adult subjects, special attention should be paid to respect their freedom and self-determination. Thus, it is necessary to take into consideration 'the issues of sampling, reliability and validity at the very outset' (Cohen *et al.* 2007:49), and 'sensitivities' of ethical issues should be addressed. Another critical phase is that teachers have an extremely strong influence on students' decisions and actions in class, hence the participants should not feel coerced, nor suffer any disadvantages, retaining anonymity and confidentiality of their identities to protect privacy. For this particular EFL experiment in an elementary school

context, a case study approach is effective because it attempts to examine 'an instance in action' and to provide 'a unique example of real people in real situations, enabling readers to understand ideas more clearly than simply by presenting them with abstract theories or principles' (Cohen *et al.* 2007:253).

According to Nisbet and Watt (1984), case studies have strengths in that 'the results are more easily understood by a wide audience', as they are frequently written in 'everyday, non-professional language' and that 'they can catch unique features that may otherwise be lost in larger scale data'. The findings might hold the key to understanding the situation and provide insights into other similar contexts. Weaknesses, on the other hand, are that 'the results may not be generalisable' and 'they are not easily open to cross-checking, hence they may be selective, biased, personal and subjective' (Nisbet and Watt 1984:78-79). The outcome of a theory-seeking or theory-testing case study might lead to a 'fuzzy generalization' (Bassey 1999:30); therefore, it should be understood that a case study may lack rigor with little basis for scientific generalization. Nevertheless, Simons (1996:225) welcomes 'the paradox between the study of the singularity and the search for generalization' and helps researchers to understand the characteristics of case studies.

3.2 Teaching English to Young Learners (TEYL)

It was widely believed until quite recently that L1 acquisition was largely complete by the age of five, and it has been hypothesized that 'children learn a L2 better than adults' (Cameron 2001:12-13). Nevertheless, literacy skills of children are still in the early stages of development under school age, and some studies provide evidence that there is no such 'cut-off point' for L2 acquisition as suggested by the Critical Period Hypothesis (CPH; first proposed by Penfield and Roberts in 1959), the idea that children can learn a L2 more effectively before puberty because their brains are still able to use the mechanisms that have assisted their L1 acquisition. Lightbown and Spada (1999:16) present some evidence for and against the CPH and they point out that if the goal of the L2 learning is 'native-like proficiency', the learning benefits from an early start, but if the goal is 'communicative ability' in a foreign language, the benefits are much less clear. In EFL teaching, the process of child language acquisition is highly suggestive; however, attention should be paid to different needs, motivations, and contexts of diverse learners. O'Grady (2005:164-175) admits that 'there is still no real solution to the mystery of child's language acquisition' and the substantial indication is that the mystery cannot be explained by 'imitation' nor 'teaching'. It is a popular idea that children

learn languages by imitating adults in immersion contexts, but the imitation cannot account for the speed at which children acquire vocabulary and language rules. According to Eyres (2007:26), children are in fact not very good at imitating sentences containing unfamiliar words and structures, and they typically repeat only what they can already say. Children often produce sentences (e.g. *allgone milk, come car*) which they have never heard anybody actually say before. Their language learning cannot be explained in one simple 'straight line' of progress.

Cameron (2001:14-15) asserts that different languages have different ways of carrying meaning, and that the particular ways in which a language encodes meaning act as 'cues' to interpreting the meaning of what is said. When faced with the new language, children try to understand it in terms of 'salient cues of their L1 and also pay particular attention to items of L2 vocabulary that they are familiar with' (Schmidt 1990:141). Generally, children are said to be more enthusiastic and lively as learners than adults, seem less embarrassed with speaking in the L2, and their 'lack of inhibition' seems to help them acquire a more native-like accent. Meanwhile, they lose interest more quickly and are less able to keep themselves motivated on one task than adults. Cameron (2001:xi-xii) claims that 'Theorising the teaching of young learners has an important role to play in complexifying oversimplifications about working with children and thereby increasing the quality of foreign language education'. When teaching young learners between the age of five and twelve, significance lies in the fact that they are still in the middle of development in their L1 and individual differences are remarkable; formal L1 literacy skills are still in the early stages of growth, discourse skills are continuing to develop. As a result, there is a burden on the teachers to provide exposure to the L2 with effective classroom activities. Young EFL learners attempt to make sense of the world in terms of their limited knowledge and experience, but they definitely need 'skilled' help in noticing and attending to various aspects of the English language.

The implications for TEYL that can be drawn at this point are threefold: firstly, if language learning is neither 'imitating' nor 'teaching', effective methods and programs (other than the L1 acquisition approaches alone) can be explored to facilitate the L2 learning. Young EFL learners in 'low-immersion' contexts like Japan have very few opportunities to 'imitate' or 'produce' the L2 in their everyday life, meaning that their learning approaches should be distinguished from the 'L2 immersion context'. Secondly, teachers play an important role to support young EFL learners by focusing their attention on useful sources of information, however limited their 'cues' are in the target language. Lastly, the process of L1 acquisition

and young learners' characteristics should be fully considered in TEYL, and the socio-cultural context needs to be examined as it exercises an invisible but enormous influence on learning motivation.

3.3 Assessing young EFL learners

The nature of young EFL learners might generate issues of assessment in need of attention: factors such as age, gender, the influence of L1, background context, the linguistic development and diverse learning motivation, all need to be taken into consideration. Cameron (2001:215) suggests that it seems 'reasonable to require an assessment to *serve* teaching, by providing feedback on pupils' learning that would make the next teaching even more effective, in a positive, upwards direction' [her italics]. Despite that, the reality is often contrary; 'assessment seems to *drive* teaching by forcing teachers to teach what is going to be assessed' and some of the classroom activities are restricted to test preparation in extreme cases. Thus, the assessment for young EFL learners is expected to 'identify the needs of young learners, to determine what level, if any, of proficiency they have in the target language, to diagnose their strengths and areas in need of improvement, and to keep track of their progress in acquiring the language' (Alderson and Bachman 2006:x).

At the same time, vulnerability might become apparent when the impact of assessment on young EFL learners is to be studied. Children are sensitive to praise, criticism and approval, and their self-esteem is influenced by their experiences among friends and teachers at school. Therefore, they need experiences that will help them to succeed and maintain enthusiasm and creativity; a lack of positive experiences might result in loss of motivation. The assessment has a power to change their learning both positively and negatively, and McKay (2006:18-19) asserts that effective assessment provides valuable information to educators, parents and students themselves, and the valid assessments are 'those that measure what they are supposed to measure'. The TOEFL Primary in this article can be classified as a 'summative assessment' to measure language skills of young EFL learners as well as an 'assessment to encourage and motivate' them. By utilising objective 'scales and scores for each language skill' rather than 'pass-fail results', the TOEFL offers the Japanese learners an opportunity to understand their English language abilities as well as to see the progress that they have made up to the present.

3.4 EFL learning motivation

In regard to learning motivation, Dörnyei and Ottó (1998:65) attempt to give a comprehensible definition stating that 'motivation can be defined as the dynamically changing cumulative arousal in a person that initiates, directs, coordinates, amplifies, terminates, and evaluates the cognitive and motor processes whereby initial wishes and desires are selected, prioritised, operationalized and (successfully or unsuccessfully) acted out.' Dörnyei (2001:46) also points out that the 'motivation to learn' is 'an intricate, multifaceted construct' on its own, but if the target is the mastery of the L2, 'the picture becomes even more complex'. He indicates two aspects of L2 motivation; one is that the L2 is a 'learnable' school subject and the results of its studies on academic achievement are of direct relevance to motivation. The other is that language is 'socially and culturally bound, and serves as the primary channel of social organisation in society'. The mastery of the L2 is not merely an educational issue but is also connected with social events that require incorporation of a wide range of factors in respective cultures. Gardner (1979:193-194) reveals the social nature of L2 acquisition and claims that foreign language learning in the school context should not be viewed as 'an educational phenomenon' in much the same light as any other school subject, but as 'a central social psychological phenomenon'. For school-age young learners, the 'school' literally represents a social arena, and the effects of the '*microcontext* (i.e. the immediate learning environment)' as noted by Dörnyei (2001:32-34 [his italics]) need to be examined thoroughly in motivation research.

The question remains, however, whether young learners' motivation is driven by an intrinsic interest in the field of subject (e.g. intellectual curiosity and preferences of challenges), or by an extrinsic orientation motivated by desire, for instance, getting good grades or skills in order to win teachers' approval or meeting external demands for higher education. As Dörnyei questions the '*researchability*' (2001:183) of motivation, what makes motivation research a formidable task is that it is not subject to direct observation but must be inferred from 'some indirect indicator, such as the individual's *self-report accounts, overt behaviours or physiological responses* (e.g. change of blood pressure)' [his italics]. What is recommended as a principal step for motivation research is to keep the level of this inherently subjective nature of motivation to the minimum, especially when analyzing motivational factors of young EFL learners.

4. Methodology

4.1 Research methods

Having taken the layers of context into consideration and reviewed relevant literature, the experiment in this article commenced with the question whether an assessment opportunity would stimulate the motivation of Japanese elementary school EFL learners and what motivational factors could be observed through the preparation for the TOEFL Primary. It was not the researcher's intention to make assertions on the learners' achievement from the scores or to draw attention to a simple interpretation of the small-scale data. The experiment was planned for 18 Year 6 students, due to the limited facility of the computer-assisted language learning (CALL) system needed for the Speaking test; prior to the actual test day, two preliminary classes were scheduled. The announcement was made in a letter to all 144 Year 6 students and 41 students showed an interest and applied to participate. With these circumstances, the participants were decided by lottery and all attended two preliminary classes to experience the mock tests and get accustomed to the test format of the TOEFL Primary. In addition to the recordings of the interactions between the learners and the teachers, a semi-structured questionnaire and post-interview were conducted after the test. A supplementary free comment space was included to gather various forms of 'self-report accounts' and 'overt behaviours' (Dörnyei 2001:185), and respective research tools are described below.

4.2 Research tools

Preliminary classes prior to the TOEFL Primary

The participants in this experiment carried diverse English learning backgrounds and their fluency levels varied, but the researcher had taught them previously or was teaching them in regular classes, so each learner's English levels were acknowledged. The relationship between the participants and the teacher-researcher was intimate and friendly, which made it easy to draw their true feelings and honest opinions. The preliminary classes were held on consecutive weekends, each for three hours, in order for the participants to understand the test content and prepare for the unfamiliar format. Two teachers implemented the lectures, one for reading and listening practice and the other for speaking, using the textbook "*Preparation Book for the TOEFL Primary® Step 1* (Kumon Publishing 2015)". The participants practiced answering a variety of questions and were given an explanation of the vocabulary and comprehension tasks. After this, they took the mock tests of each

section. Throughout this process, two observers were present from the TOEFL testing organization (Global Communication & Testing Co. Ltd.), who supported the test session and audio-recorded the lectures. Target factors to be investigated were categorized accordingly in relation to the post-test questionnaire questions as shown in Table 1 (see also Appendices 1 and 2), and the teacher-researcher, while engaging in interactions with the participants, took observation notes in written format, utilizing subsequent video analysis of classroom exercises.

Table 1. Target factors to be investigated through the preliminary classes prior to the test

Questions in the questionnaire	*Target factors to be investigated*
1. Did you know about the TOEFL Primary before this experiment?	− **Recognition of the standardized English test** − **Personal interest in assessment opportunities**
2. Why did you take part in this experiment and decide to take the test?	− **Specific motivational factors** − **Influence of family/peers/school milieu** − **Attitude towards learning English**
3. How were the TOEFL Primary tests? Which were difficult? 4. What did you think about the Speaking test?	− **Fondness of the assessment/test content** − **Confidence or fear for speaking English** − **Factors related with characteristics of young EFL learners**
5. *Agreement rating:* The TOEFL Primary preliminary classes were good. 6. *Agreement rating:* This opportunity to take the TOEFL Primary was good.	− **Fondness of linguistic activities** − **Attitude towards assessment opportunities** − **Learning motivation towards English**
7. *Agreement rating:* I would take the TOEFL Primary again if planned. 8. *Agreement rating:* Further TOEFL test opportunities in affiliated schools would be good. 9. Do you wish to study abroad (either short or long term) in the future?	− **Future images of learning English** − **Perception towards learning English** − **Motivational factors related with personal wishes**

Observation of the actual assessment and post-interview

Observation is a strong tool for interpretive research and offers researchers an opportunity to gather 'live' data from naturally occurring social situations (Cohen *et al.* 2007:384), although it is also beset by issues of validity and reliability. For this case study, the actual test

session was observed and recorded in written format by the teacher-researcher in accordance with the target factors shown previously, and for structured observation, four motivational domains clarified by Morrison (1993:80) were used, i.e. physical setting, human setting, interactional setting and program setting. For the Speaking test, two groups were formed to maintain the personal space while speaking to the microphone; one group contained 8 students, taking the Speaking test for the first 30 minutes, while the other had 10 students, who took the paper-based Reading and Listening tests for the first 60 minutes. The Listening test was played on an audio-CD player and took approximately 30 minutes to finish with the directions in English. To reinforce individual accounts, the participants' live and instant comments were extracted verbally immediately after the test and written into the form of notes. Post-interview was administered one month later when the test results arrived. Interviewing was beneficial 'to understand the world of children through their own eyes rather than the lens of the adult' (Cohen *et al.* 2007:374), and it revealed how the participants perceived the assessment. The duration was five minutes, and each participant was asked the following questions: 1) How did you like the TOEFL? 2) Were the preliminary classes and the test session beneficial? 3) Would you like to develop your English abilities?

A semi-structured questionnaire to 18 participants

Subsequent to the actual test, a semi-structured questionnaire (Appendix 1, translated by the researcher from Japanese into English) was undertaken with the 18 Year 6 students. The questionnaire is a useful tool 'for collecting data from a large number of respondents' and 'the information sought is not so complex' (Hinds 2000:42-3). The primary objective was to grasp the participants' interest in the assessment and their attitude towards English language learning. A Likert scale (named after its deviser in 1932, Rensis Likert) with five point responses was employed, as it was easy for the young learners to indicate their opinions by simply circling a mark which should represent their ideas most. A free comment space was included to invite 'an honest, personal comment from respondents' (Cohen *et al.* 2007:330), which might contain information that otherwise might not be caught in standardized questionnaires. Although 18 was not a sufficient number to yield a general tendency of young learners, the results contributed to illuminating certain motivational factors such as 'a sense of achievement', 'parental influence' and 'future images of EFL learning'. This experiment was not for generating hypotheses, classifying or categorizing; rather, it aimed to illuminate, interpret and describe what was observed in the practice of TEYL, informing 'on-going'

teaching and learning in a specific context in Japan.

5. Results and Analysis

5.1 Results of the preliminary classes

During the preliminary classes, all the participants showed strong interest and curiosity in the content of the test, asking enthusiastically about the meaning of each text and sharing the new knowledge with their peer group. As the TOEFL Primary Step 1 is specifically aimed at beginner-level young learners, both the Reading and the Listening tests carry a number of illustrations to match the correct vocabulary. The participants found them amusing and answered the questions with confidence. In latter parts of the Reading test, there are riddles and short passages with about 80-100 words in order to examine the test takers' knowledge of vocabulary and syntax as well as to test their comprehension speed in English. Contrary to the researcher's anticipation, the participants vigorously answered the questions and sought more detailed explanations in Japanese. While reading the passages beyond their lexical level, they made inferences with some verbal assistance and exhibited high motivation to answer the questions. This suggested that an assessment opportunity could give young learners much more than mere linguistic knowledge; a learning motivation driven by young EFL learners' pure interest in new things, not just to score good points on the test.

Similarly, the Listening test contains illustrated conversations in the beginning, but the later sections include several dialogues and recorded messages. The participants practiced oral reading of some scripts with approximately 80-100 words in the preliminary classes, but they struggled with the high speed of spoken English as well as the deliberate choices of vocabulary. Moreover, in the Speaking test, practical usage of the English language is required; for instance, the test takers

Figure 1. Preliminary classes with the CALL system for the Speaking test (December 2015)

look at several pictures and describe the situations in progress within 15 seconds, or in other topics, they watch an animated cartoon story and are required to explain what has happened within 30 seconds. The preliminary classes offered the participants an opportunity to exercise this format with the CALL system within a required time (Figure 1). The results displayed some remarkable behaviors, for example, some students uttered whatever words that flashed

into their minds, or others showed puzzlement with the gap between L1 and L2, being unable to describe the situations in English. Nevertheless, the participants enjoyed the activities for the most part, indicating a positive attitude towards learning English.

5.2 Results of the assessment session

During the actual TOEFL session, several factors were identified in terms of motivational domains, following Morrison's categorization (1993:80) as shown in Table 2.

Table 2. Motivational domains observed in the assessment, Morrison (1993)

Physical setting	− Physical environment of the test room and the layout of the desks created a serious, quiet milieu for the Reading/Listening tests. − The distance between the participants was sufficient for the test. − The CALL room was assisted with the teachers and the testing staff, which seemingly added a tense ambience.
Human setting	− Observable actions and behaviors were witnessed such as tone of voice, facial expressions and postures before and during the test session. − Most participants showed fear on their faces in the beginning. − Characteristics of individual participants were observed, e.g. puzzled faces when listening to the instructions in English.
Interactional setting	− Both formal and informal interactions were observed in the tests. − When each group finished with the Speaking tests, the participants moved and some interactions occurred among the two groups. − Some showed confidence in the Reading test, while others looked troubled, especially in and after the Speaking test.
Program setting	− Preliminary classes seemed effective as the participants had learned beforehand what to expect in the actual test. − Some participants cited they would study more speaking and listening.

In the Reading test, most participants finished the section earlier than the time limit, and the score results showed that 8 students out of 18 obtained 109 points (the best score for the Primary Step 1 test), and 15 students (83%) reached A2 level of CEFR, scoring more than 107 points (see Appendix 3).

In the Listening test, on the contrary, the content was perceived as challenging, and

many of the participants indicated confusion in the latter section. As was hypothesized, the Listening test was difficult for the beginner-level participants, and the test duration of half an hour appeared too long for the young learners to remain focused. As a result, the scores varied greatly, and 6 students reached 109 points (the best score). Despite that, all of the participants reached A2 level of CEFR listening ability, scoring more than 105 points.

The most challenging section was the Speaking test, and the participants were overwhelmed by the time limit. During the session, some keyword vocabulary was presented verbally on individual computer screens, but the test format seemed far beyond their previous experiences with spoken English, with the result that the Speaking test was regarded the most difficult. In terms of CEFR speaking ability levels (maximum point score of 27), 6 students (30%) achieved A2 level, of whom 4 obtained 18 points or more. Another 11 students (61%) were graded at A1 level.

The actual session of the TOEFL Primary revealed that the Japanese elementary school EFL learners lacked the experience of listening to spoken English as well as speaking the language in authentic settings.

5.3 Results of the semi-structured questionnaire and the post-interview

The results of the nine closed-questions were analyzed with some charts from gross responses (Appendix 2); Questions 1 and 2 explored the participants' knowledge about the TOEFL and the reason why they took part in the test. Question 3 searched for the learners' perception of the test in regard to three language skills, i.e. reading, listening and speaking. Question 4 focused particularly on the Speaking test, investigating what aspects seemed easy or difficult to the participants and whether they became motivated or demotivated by the test format. Questions 5 to 8 related to the young learners' attitudes towards EFL learning and how the extracurricular assessment opportunity was regarded. Question 9 looked into the learners' interest in studying abroad, and the last space invited free comments, which were translated from Japanese into English by the researcher and listed in relation to two broad categories of motivational factors.

Question 1 was the premise for learning motivation, and three quarters, 14 out of 18 participants, did not know about the TOEFL before the experiment. Question 2 was directly connected with the individual motivations (Figure 2), and the top three answers were: a pure interest by the young learners in a new experience, preparation for secondary school English, and a simple motivation to acquire foreign language. It was shown that for young EFL

learners, family influence plays a central role; more than 40% responded that their parents advised them to take the test. Those participants wishing to study abroad in the future were strongly motivated and considered the assessment as a good opportunity for acquiring practical skills.

Figure 2. (Question 2: Why did you take part in this experiment and decide to take the test?)
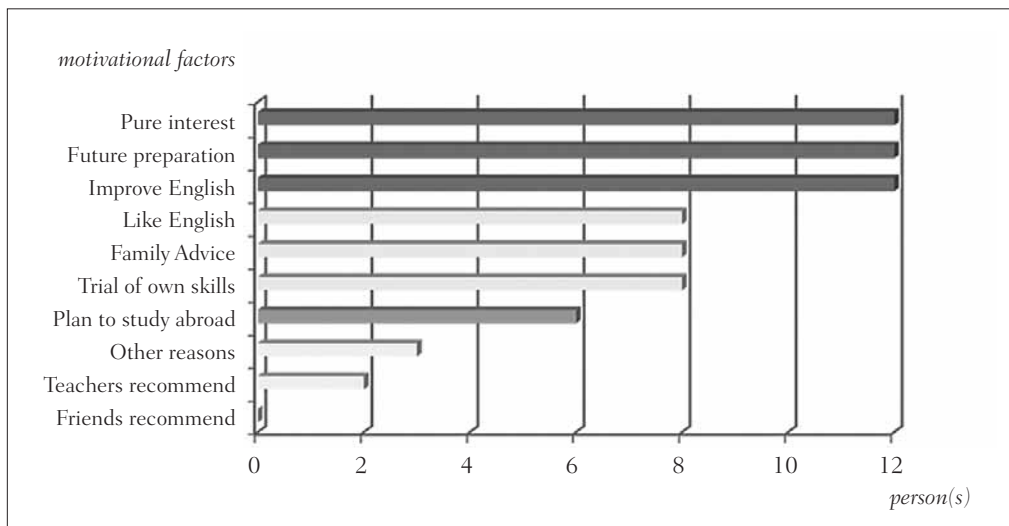
Question 3 intended to find the participants' impression on each test (Figure 3), and the overall view was that the test was difficult, particularly the Speaking test. Roughly 60% thought the Reading test was hard, and more than 80% considered the Listening test difficult. Over 88% believed the Speaking test was 'beyond their abilities', according to their verbal accounts after the session. The results reflect that the Japanese young EFL learners are not accustomed to the actual usage of English language, and the test format is far too unfamiliar. Among them, the Reading test was perceived positively and was actually finished within the time limit. This implies that some young learners already possess a certain level of literacy and can read English texts to some extent.

Figure 3.（Question 3: How were the TOEFL Primary tests? Which were difficult?）
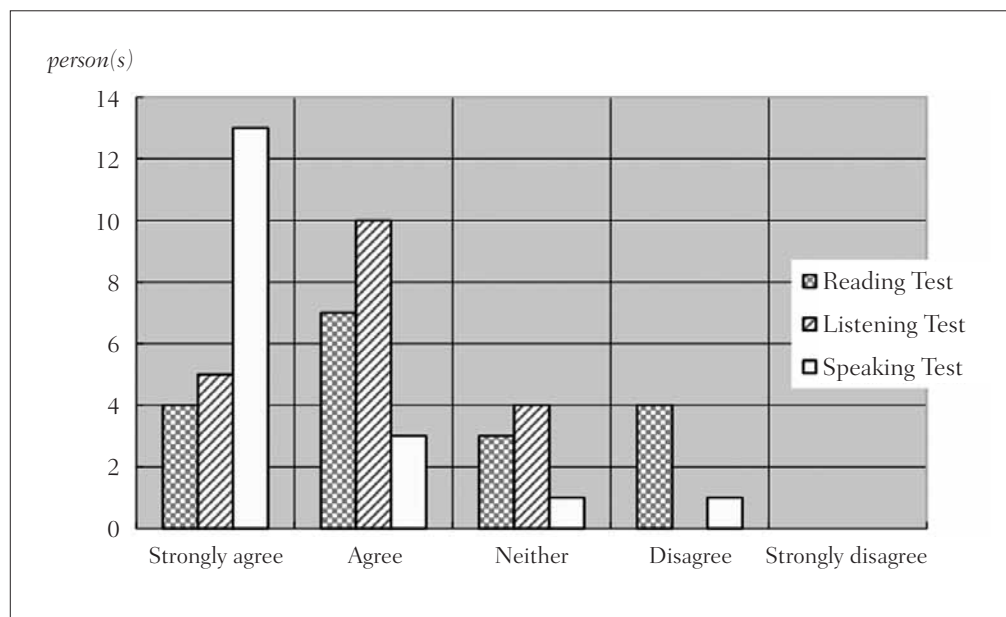
Question 4 specified what factors of the Speaking test were considered difficult, as the test format of using a microphone was unfamiliar to the Japanese EFL learners; nearly 90% agreed that it was their first experience to speak English to a machine. The participants also found the test content demanding and struggled to express their ideas within a limited time frame, although the instructions in English were understood properly. In regard to learners' motivation, more than 70% showed an intention to take the test again, proving their enthusiasm and a sense of competitiveness. The results showed the strong motivation of promising young EFL learners who are willing to advance their learning.

Questions 5 and 6 investigated the post-assessment impressions, and the results were 100% positive; the participants appreciated the preliminary classes and the assessment opportunity, verifying that the inclusion of extracurricular test opportunities is meaningful and could stimulate young learners' motivation. Questions 7 and 8 shed light on the participants' desire to take another assessment in the future, and the majority (77%) displayed eagerness to try again, although some negative responses were observed at the same time.

In addition, Question 9 explored learning motivation connected with the participants' eagerness to study abroad in the future (Figure 4), and nearly 80% agreed, half expressing

a strong wish. This implies that Japanese elementary school EFL learners feel it necessary to live in an English-speaking environment to become fluent users of English language. A powerful motivation was witnessed.

Figure 4. (Question 9: Do you wish to study abroad (either short or long term) in the future?)

■ Strongly agree	8
■ Agree	6
■ Either	2
■ Disagree	0
■ Strongly disagree	2
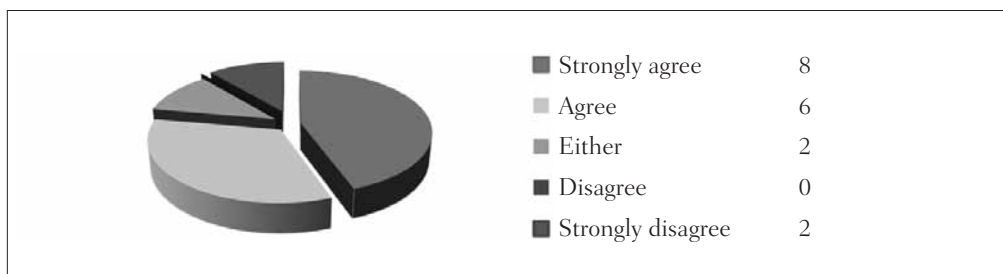
Lastly, the free comment space collected individual voices and diverse perspectives on English language learning, which highlighted the young learners' subtle yet enthusiastic attitudes towards EFL learning. A supplementary question, "What English abilities would you like to acquire?" was added so that the participants could write their thoughts without losing focus. The responses could be divided into two broad categories for clarification, 'linguistic knowledge' such as vocabulary and syntax that the young EFL learners hope to acquire, and 'communicative experience' such as speaking and listening skills that the Japanese learners hardly have a chance to gain outside the classroom (Table 3).

The results of the post-interview reflected how the individual participants perceived the assessment and the responses were diverse, for example, "The preliminary classes were enjoyable, but I never expected the Speaking test to be that difficult." or "It was exciting and I want to improve my English." A large number of the comments pointed out the importance of English language learning for the purpose of school, business, travel etc. in the future. In addition, positive images of English were seen from their verbal comments such as 'It's cool to speak English' or 'I feel great when using English'.

Table 3. Voices from the free comment space:

Linguistic Knowledge (vocabulary/syntax factors)	
I want to learn more words and phrases.	I want to get better reading and listening skills.
I want to read faster.	I want to learn grammar.
I want to get more vocabulary, acquire grammar as well as composition skills.	I want to use English more accurately.
Communicative Experience (sound/communication factos)	
I want to understand English spoken in the movies and TV dramas.	I want to get English abilities accepted and understood abroad.
I want to be able to speak English without being too nervous.	I want to speak English fluently and travel around the world by myself.
English would be very helpful when talking to foreign people for the first time.	I want to work as a bridge between Japan and foreign countries.
I want to speak to foreign people with fluency.	I want to get better in speaking.
I want to speak with foreign people.	I want to be able to sing songs in English.
I want to be good at English so that I can understand what people are talking about.	I think grammar is important, but I want to get speaking proficiency too.
I want to know more convertational expressions and vocabulary out of textbooks.	I want to listen to English more and get better listening and speaking skills.
I want to explain the directions to foreign visitors instantly when they are in trouble.	I want to work on the global stage in the future, and English is necessary to get there.

5.4 Analysis of young learners' motivation observed through the experiment

The observation and the questionnaire in this case study exposed to view various but fundamental motivations of Japanese EFL learners, and the experiment illuminated the elementary school learners' attitudes towards EFL learning. The free comment space particularly succeeded in gaining motivational factors in detail, reinforcing each factor to be narrowed down for further research. Some responses might be 'impressionistic' (Cohen *et al.* 2007:330); however, the comments mirror the reality of the present TEYL in Japan, especially in the elementary school context, and the significance of the data collected can be

summarised as follows:

1) The preliminary classes made it apparent that the participants perceived the assessment opportunity positively and considered it as a chance to test their language abilities. Parental influence is extremely powerful for shaping these positive attitudes and images.

2) The challenge of assessment means more than a mere measurement of language competence to young learners. The participants in this experiment found pleasure in gaining new knowledge and their motivation was driven by their pure interest and intellectual curiosity. The assessment can function as a stimulator and the elementary school students can obtain foreign language experiences with appropriate support from the teachers.

3) The questionnaire revealed positive images of EFL learning among the young learners despite the reality that the English language has no major role in Japanese society. An important implication for TEYL is that the social context plays a crucial role in motivating young EFL learners.

4) The participants were highly motivated and some possessed advanced language abilities for their age as the test results showed (Appendix 3). They appreciated occasional assessment opportunities, especially those who were wishing to study abroad. The issue to be discussed further is the link between elementary and secondary school English education; the experiment has opened a question to the present ELT curriculum, in which no linguistic skill-based subdivisions have been constructed between the two.

This was the first attempt for all the participants at the TOEFL Primary, but the results showed that the Japanese elementary school EFL learners possessed diverse motivations, and they found enjoyment in EFL learning. The experiment has illuminated the enthusiasm of young learners and their motivational factors for English language learning. As was observed all through the experiment, socio-cultural influence appears to be enormous in creating learner perception, indicating the necessity of setting appropriate contexts for young EFL learners.

6. Discussion: the impact of assessment opportunity and motivational transition

In this case study, motivation of Japanese elementary school EFL learners is highlighted through the course of study for the TOEFL, and the results imply that the inclusion of assessment might function effectively in TEYL. The motivational factors are diverse and the data indicates both intrinsic and extrinsic motivations, the former mostly from the learners' pure interest, curiosity or competitiveness, and the latter from family influence, desires for appraisal or positive images of EFL learning engraved in Japanese social values. Occasional assessment opportunities like the TOEFL could stimulate young learners' enthusiasm, and high expectation towards global status can be observed. However, the question remains whether the young learners can maintain high motivation through puberty on to adolescence in the present EFL environment in Japan. Each young EFL learner differs in motivation and likewise in aptitudes for foreign language learning. As this experiment has elucidated, different motivations are originated in each learner's 'value mechanisms' or 'appraisal system' (Schumann 1997:2). A constructive outcome is the fact that the positive young learners naturally find pleasure in learning linguistic knowledge and acquiring communicative experience through the assessment.

On the other hand, as Nishida (2013:98) recognizes, 'motivation tends to decrease as students get older' and 'no objective measure of linguistic ability is available' in Japanese elementary school contexts. As learners grow older, linguistic achievement is required, and many EFL learners tend to weaken in motivation in Japan. Yet, the results from this experiment have contrasted young learners' willingness to obtain real-life communicative competence, and the participants possessed a strong desire for further learning. The problematic features of standardized tests like the TOEFL would be that the settings are always 'virtual', not authentic, which has often led to the question as to whether this would facilitate genuine communication skills. As the learners get older, the issue of 'authenticity' can affect their motivation, because foreign language learning tends to become a mere 'school subject'. The differentiations in learning motivation might call the need for 'individualization' in EFL teaching, particularly in school context, and a dramatic improvement in the link between elementary and secondary school English education is greatly needed. Along with needs analysis, a variety of teaching methods and techniques should become essential for elementary school EFL learners, and it is suggested that the young EFL learners in Japan would benefit from more effective, individualized English teaching in a smaller group and in an appropriately streamed form of learning for better communicative practices.

7. Conclusion

The experiment in this article has shed light on the motivational factors of Year 6 students in a private elementary school context in Japan and portrayed diverse motivational factors of EFL learning. Through the preliminary classes (Figure 5) and the actual TOEFL Primary session, it was discovered how the assessment was experienced by the Japanese elementary school students, and the impact of the assessment opportunity was closely examined. Despite the argument, the reality is that elementary school education adopts English

Figure 5. Preliminary classes for the Reading and the Listening tests (December 2015)

language as a regular subject, and Japanese people consider it necessary for the future careers of the younger generation. The dominant power of social context including family influence has become evident, and the strong desires of young learners to gain communicative skills were exhibited in this study. The assessment opportunity could function as a stimulator for young EFL learners and as Cameron (2001:218) advocates, a tactfully practiced 'learning-centred perspective' should be assumed in TEYL.

The remarkable feature of the Japanese EFL context is that the positive image of acquiring English language is deeply connected with young learners' motivation, and a social 'value' conveys an immense effect on foreign language learning. The various factors cannot simply fit in one theory or generalization; however, the inclusion of assessment can offer important opportunities for EFL learners to see their linguistic abilities by an international standard. It is, therefore, recommended to create these opportunities regularly at school in order to help both learners and teachers to measure language skills and identify the next step of learning and teaching.

This experiment has also called for the need to explore further motivational transition from elementary to secondary school English education longitudinally. It is to be hoped that the voices of the young EFL learners in this experiment can contribute to the betterment of the field of TEYL in Japan, in some part enabling methods of effective learning and teaching of English language in school contexts to continue to be enlightened by further research.

Bibliography

Alderson, J. C. and L. F. Bachman. (2006) Series editors' preface. In McKay, P. *Assessing Young Language Learners*. Cambridge: Cambridge University Press.

Bassey, M. (1999) *Case Study Research in Educational Settings*. Berkshire: Open University Press.

Cameron, L. (2001) *Teaching Languages to Young Learners*. Cambridge: Cambridge University Press.

Cohen, L., Manion, L. and K. Morrison (2007) *Research Methods in Education* [sixth edition]. Oxson: Routledge.

Council of Europe. (2001) *Common European Framework of Reference for Language Learning, Teaching, Assessment* (CEFR). Cambridge: Cambridge University Press.

Crystal, D. (2003) *English as a Global Language* [second edition]. Cambridge: Cambridge University Press.

Dörnyei, Z. (2001) *Teaching and Researching Motivation*. Harlow: Pearson Education.

Dörnyei, Z. and I. Ottó. (1998) Motivation in action: A process model of L2 motivation. *Working Papers in Applied Linguistics*. 4:pp.43-69. London: Thames Valley University.

Educational Testing Service (ETS). (2016) *Test and Data Summary for TOEFL iBT® Tests 2015*. Princeton: ETS.
[on-line] https://www.ets.org/s/toefl/pdf/94227_unlweb.pdf

Eyres, I. (2007) *English for Primary and Early Years. Developing Subject Knowledge* [second edition]. London: The Open University.

Gardner, R. C. (1979) Social psychological aspects of second language acquisition. In Giles, H. and R. St. Clair. (eds.) *Language and social psychology.* pp.193-220. Oxford: Blackwell.

Higuchi, T., T. Kunikata, Y. Omura, Y. Tanabe, E. Shimo, E. Izumi, T. Kinugasa, T. Kagata and Y. Hakozaki. (2008) A follow-up survey of primary English learners and implications for primary and secondary English education. [translated from Japanese by the researcher] *The English Teachers' Magazine* [special edition]. October 2008, pp.58-69. Tokyo: Taishukan.

Hinds, D. (2000) Research instruments. In Wilkinson, D. (ed.) *The Researcher's Toolkit – The complete guide to practitioner research*. Oxon: Routledge.

Kubota, R. (2002) The impact of globalization on language teaching in Japan. In Block, D. and D. Cameron. (eds.) *Globalization and Language Teaching*. Oxon: Routledge.

Kumon Publishing. (2015) *Preparation Book for the TOEFL Primary® Step 1*; original copyright by Ji, T. Lim. (2014) First published in Korea. Japanese translation supervised by Global Communication & Testing. Tokyo: Kumon Publishing Co.Ltd.

Lightbown, P. and N. Spada. (1999) *How Languages are Learned* [second edition. Oxford: Oxford University Press.

McKay, P. (2006) *Assessing Young Language Learners*. Cambridge: Cambridge University Press.

Ministry of Education, Culture, Sports, Science and Technology (MEXT) - Japan. (2003) *Regarding the Establishment of an Action Plan to Cultivate "Japanese with English Abilities"*.
[on-line] http://www.mext.go.jp/english/topics/03072801.htm

Ministry of Education, Culture, Sports, Science and Technology (MEXT) - Japan. (2006) The official record of the proceedings of the International Education Division, Elementary and Secondary Education Bureau. The original document is available only in Japanese.
[on-line] http://www.mext.go.jp/b_menu/shingi/chukyo/chukyo3/015/siryo/06080203/004.htm

Ministry of Education, Culture, Sports, Science and Technology (MEXT) - Japan. (2014) *Report on the Future Improvement and Enhancement of English Education (Outline): Five Recommendations on the English Education Reform Plan Responding to the Rapid Globalization.*
[on-line] http://www.mext.go.jp/english/topics/1356541.htm

Morrison, K. R. B. (1993) *Planning and Accomplishing School-Centred Evaluation.* Dereham: Peter Francis.

Nisbet, J. and J. Watt. (1984) Case study. In Bell, J., T. Bush, A. Fox, J. Goodey and S. Goulding (eds.) *Conducting Small-Scale Investigations in Educational Management*, pp.72-92. London: Harper & Row.

Nishida, R. (2013) A Comprehensive Summary of Empirical Studies of Motivation Among Japanese Elementary School EFL Learners. In Apple, M. T., Da Silva, D. Fellner, T. and Authors of Individual Chapters. (eds.) *Language Learning Motivation in Japan*, pp.93-109. Bristol: Multilingual Matters.

O'Grady, W. (2005) *How Children Learn Language*. Cambridge: Cambridge University Press.

Otsu, Y. (2004) Objections to English language teaching in state primary schools in Japan. [translated from Japanese by the researcher] *The English Teachers' Magazine*. Vol.53/2 May 2004, pp.8-11. Tokyo: Taishukan.

Sakui, K. (2004) Wearing two pairs of shoes: language teaching in Japan. *ELT Journal*. Vol.58/2 April 2004, pp.155-63. Oxford: Oxford University Press.

Schmidt, R. (1990) The role of consciousness in second language learning. *Applied Linguistics*. II, 2, pp.129-58. Oxford: Oxford University Press.

Schumann, J. H. (1997) *The Neurobiology of Affect in Language*. Malden, MA: Blackwell.

Simons, H. (1996) The paradox of case study. *Cambridge Journal of Education*. 26(2), pp. 225-40. Cambridge: Cambridge University Press.

Tomita, Y. (2004) Heading to authentic foreign language activities as a part of intercultural understanding. [translated from Japanese by the researcher] *The English Teachers' Magazine*. Vol.53/2 May 2004, pp18-20. Tokyo: Taishukan.

Torigai, K. (2006) *Critical Condition of Elementary School English Teaching.* [translated from Japanese by the researcher] Tokyo: Bungei-Shunju

Wilkinson, D. (2000) Analysing data. In Wilkinson, D. (ed.) *The Researcher's Toolkit – The complete guide to practitioner research*. Oxon: Routledge.

Appendices

Appendix 1: A semi-structured questionnaire (translated into Japanese by the researcher)

慶應義塾大学外国語教育研究センター
Keio Research Center for Foreign Language Education
http://www.flang.keio.ac.jp/

一貫校から始める TOEFL Tests 講座

【Experimental Class "Starting TOEFL Preparation": Questionnaire】

【1】 Did you know about TOEFL Primary before this experiment?

【2】 Why did you take part in this experiment and decide to take the test? (multiple answers)

- ☐ I was interested in the test.
- ☐ My family advised me.
- ☐ My friends recommended.
- ☐ My teacher recommended.
- ☐ The test will be necessary in future.

- ☐ I like learning English.
- ☐ I thought I could improve my English.
- ☐ I want to study abroad in future.
- ☐ I want to test my proficiency
- ☐ Other reasons （　　　　　　　　　）

【3】 How were the TOEFL Primary tests? Which were difficult?

Strongly Agree — Agree — Neither — Disagree — Strongly Disagree

① Reading was difficult.

② Listening was difficult.

③ Speaking was difficult.

【4】 What did you think about the Speaking test?

① Speaking to the microphone was difficult.

② I couldn't express myself in English well.

③ Instructions in English was difficult.

④ I want to try the Speaking test again.

Strongly Agree — Agree — Neither — Disagree — Strongly Disagree

【5】 The TOEFL Primary preliminary classes were good.

【6】 This opportunity to take the TOEFL Primary was good.

【7】 I would take the TOEFL Primary again if planned.

【8】 Further TOEFL test opportunities in affiliated schools would be good.

【9】 Do you wish to study abroad (either short or long term) in the future?

【Free Comments】 (What English abilities would you like to acquire in the future?)

Appendix 2: Gross results of responses for the semi-structured questionnaire (with original Japanese version, the numbers indicate the responses)

慶應義塾大学外国語教育研究センター
Keio Research Center for Foreign Language Education
http://www.flang.keio.ac.jp/

一貫校から始める TOEFL® Tests 講座

【2015年12月実験授業（幼稚舎・受講者数 18 名）】 アンケート結果

【1】以前から TOEFL Primary® を知っていましたか？　　　　はい　　4
　　　　　　　　　　　　　　　　　　　　　　　　　　　　いいえ　14

【2】今回 TOEFL®Test の集中講座とテストを受験することにした理由は何ですか？（複数回答可）
　　□ 興味があったから　　　　12　　　　□ 英語が好きだから　　　　　　　　8
　　□ 家族に勧められたから　　8　　　　□ 英語の力を伸ばせると思ったから　12
　　□ 友達に誘われたから　　　0　　　　□ 将来留学を考えているから　　　　6
　　□ 先生に勧められたから　　2　　　　□ 力試しをしたかったから　　　　　8
　　□ 今後必要になりそうだから　12　　□ その他（ 将来 TOEFL ibt を受ける練習を兼ねて
　　　　　　　　　　　　　　　　　　　　英語と関わる仕事につきたい　愉しそう、とりあえずやる
　　　　　　　　　　　　　　　　　　　　自分の力がどれくらいかるかわかる ）

【3】TOEFL Primary® Step 1 の集中講座とテストを受験してどう思いましたか？

強くそう思う／少しそう思う／ちょうど良い／あまりそう思わない／まったくそう思わない

① Reading が難しかった　　　4　7　3　4　0
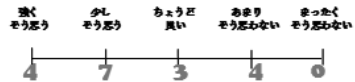
② Listening が難しかった　　5　10　4　0　0
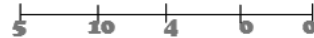
③ Speaking が難しかった　　13　3　1　1　0
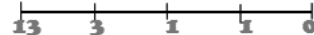
【4】Speaking テストについてどう思いましたか？

① PC 相手にマイクで話すのが難しかった　　10　6　0　1　1

② 思ったよりも英語が口から出てこなかった　14　3　0　1　0
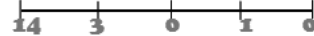
③ 指示がすべて英語で分かりづらかった　　8　5　4　1　0

④ チャンスがあったらもう一度挑戦したい　　9　5　1　1　2
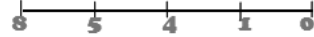
強くそう思う／少しそう思う／どちらでもない／あまりそう思わない／まったくそう思わない
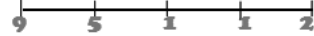
【5】今回 TOEFL Primary® Step 1 の集中講座を受講して
　　良かったですか？　　　11　7　0　0　0
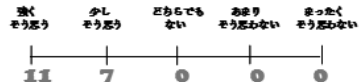
【6】今回 TOEFL Primary® Step 1 のテストを受験して
　　良かったですか？　　　13　5　0　0　0
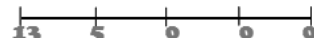
【7】次回 Primary® Step 2 の集中講座とテストの機会が
　　あれば受けたいですか？　　6　8　2　1　1
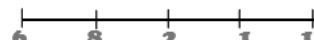
【8】TOEFL®受験の機会が今後も一貫校であると良いと思いますか？　12　3　2　0　1
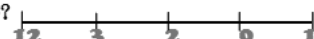
【9】将来、留学（短期・長期どちらでも）をしたいと思いますか？　8　6　2　0　2
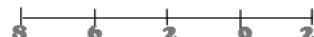
Appendix 3: [Upper part] Gross results of the open-ended question
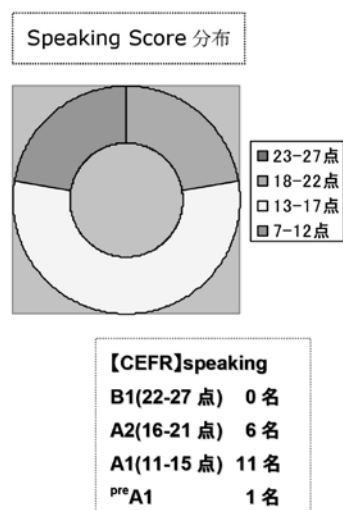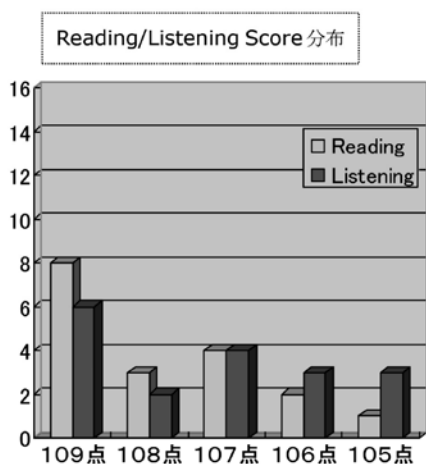
[Lower part] Score results of the participants; the TOEFL Primary Step 1

★その他、自由記入欄（今後どのような英語の力を身につけたいですか？）

Vocabulary・Syntax の要素	
単語をもっといろいろ覚えたい	Reading や Listening の力もがんばりたい
スラスラと読めるようになりたい	もっと文法の力をつけたい
語彙力・英作文、文法を学びたい	
Sound・Communicative の要素	
ドラマとか映画を観て英語がわかりたい	海外で通用する英語力
リラックスして話せるようになりたい	外国の人とペラペラしゃべれるようになる
初対面の人ともしゃべれると役に立つ	将来、日本と外国の架け橋になりたい
ペラペラになり海外旅行に1人で行きたい	特に Speaking 力を伸ばしたい
道で聞かれてすぐに言葉が出るようになる	歌を歌いたい
通りすがりの人の話を理解したいから	文法も大事だが Speaking を身につけたい
日常会話や教科書にない言葉も知りたい	いろんなものを聞いて聞く・話す力つける
外人相手でも気楽に緊張せず話せるように	

【幼稚舎 2015 テスト結果】

Reading score 109 … 8 人	Listening score 109 … 6 人	Speaking score 23-27 … 0 人
108 … 3 人	108 … 2 人	18-22 … 4 人
107 … 4 人	107 … 4 人	13-17 … 10 人
106 … 2 人	106 … 3 人	7-12 … 4 人
105 … 1 人	105 … 3 人	0-6 … 0 人
8 割が A2	全員 A2	3 割が A2

Reading/Listening Score 分布

Speaking Score 分布

【CEFR】speaking
B1(22-27 点) 0 名
A2(16-21 点) 6 名
A1(11-15 点) 11 名
preA1 1 名

外国語学習法の非言語的認知機能への影響：
日本語を母語とするドイツ語初習者を
対象とした実験的検討

矢田部　清　美

皆　川　泰　代

Abstract

The present study examines changes in young adults' non-verbal cognitive abilities induced by foreign language learning, with a view to investigating how a particular instruction method may affect students' general executive function as well as language proficiency. Seventeen Japanese university students were divided into two groups, on the basis of their scores on tests evaluating their non-verbal cognitive abilities and knowledge of German. Participants in the first group were taught German using the inductive method and participants in the second group were taught the language using the deductive method. Both groups were taught the same content with the same material; the only difference between the two was in the type of instruction. There was no significant difference in the level of language proficiency attained between the two groups. However, there was a weak but significant effect of instruction method on participants' ability to perform a non-verbal cognitive task. We also found a significant difference in reaction time between the two groups in a German and a non-verbal cognitive task. These findings suggest that it might be possible to enhance certain types of one's non-verbal cognitive abilities through language learning, although more research is needed to reach a firm conclusion.

１．　はじめに

　本論文では、外国語学習法、特に外国語を異なる学習様式で学習すると、学習者の外国語の熟達度だけでなく、非言語的認知機能に変化に違いが生じうるのかという問題について論ずる。まず、一般的に、外国語学習の結果として非言語的認知機能に変化が生じるということがありうるのかを、従来のバイリンガルと外国語学習の研究を参照して論じ、続いて本研究の課

題設定を論じる。さらに、異なる学習様式での外国語学習研究の必要性を述べ、日本語を母語とするドイツ語初習者を対象としたパイロット実験を説明する。最後に実験結果を踏まえ、異なる学習様式の外国語学習による学習者の認知機能への考えうる影響と今後の外国語教育への展望を述べる。

1.1　外国語学習経験による認知機能の変化
1.1.1　バイリンガルであることによって生じる認知機能の変化

　最近、国際化が進展するにつれ、複数の言語を同時に習得する、または母語以外の外国語を学習する等により、日常的に複数の言語を使用する機会が増えている。このように、ある個人が少なくとも二つの言語を話すこと、またはその使用者をバイリンガル（bilingual）と呼ぶことがある（American Speech-Language-Hearing Association, 2004）。本節では、バイリンガルであることが認知処理全般に影響を及ぼし得るのか、さらに本研究の課題設定について論じる。

　1960年代までバイリンガルは一つの言語を話すモノリンガル（monolingual）の言語レベルに到達することが出来ず、それゆえ多くの認知処理課題においてモノリンガルより劣っていると考えられていた（Tunmer, Pratt, Herriman, & Bowey, 1984）。しかし1960年代に入ると、過去の研究には調査対象のグループ間の社会経済レベルや年齢が統制されていない、調査問題に文化背景が異なると答えにくい問題が含まれている等の深刻な方法論上の欠点があったことが指摘されはじめた（Adesope, Lavin, Thompson, & Ungerleider, 2010; Bialystok, Craik, & Luk, 2012; Peal & Lambert, 1962; Tunmer et al., 1984）。そのため、こうした点が統制されたそれ以降の研究により、バイリンガルはモノリンガルより多くの認知処理課題において優れた点を有するという報告が多数を占めるようになった。

　バイリンガルが優位性を持つような認知処理には何種類かのものがあると言われている。ひとつは言語知識を構造化し説明するメタ言語的気づき（metalinguistic awareness）である。バイリンガルのメタ言語的気づきの成長を促すためには一定の閾値を超える第一言語と第二言語の両言語の能力が必要であることが知られている（Cummins, 1979; Ricciardelli, 1992）。また、なぜバイリンガルがメタ言語的気づきに優れているかについては複数の言語を使用することで言語の違いなどの言語知識の抽象化が発達するからではないかと示唆されている。もうひとつは、非言語課題を含めた多くの認知処理課題に解答する際に共通して必要となる実行機能（executive function）（Miyake et al., 2000）である。これは、注意や記憶などの限られた資源を使いつつ情報更新や目標設定を行い思考や行動を制御する高次の認知機能である。バイリンガルの実行機能の成長を促すためには、二言語の熟達度ではなくバイリンガル教育の年数が影響すると言われている（Bialystok & Barac, 2012）。しかしバイリンガルの実行機能の優位性にバイリンガル教育におけるどのような過程が影響したのかは不明であった。

　さらに後者のバイリンガルの実行機能の優位性についてはまだ議論も多い。たとえば、言語の獲得時期・熟達度・バイリンガル言語の類似性などを統制すると、優位性があるという研究結果は再現しない、とする報告もある（Paap, Johnson, & Sawi, 2015）。しかしそうした報告においても特定のバイリンガル経験に限れば実行機能に正の影響を及ぼすことが否定できないとしている。

　そこで、本研究では新しい言語を学ぶときに特定の学習過程を設定し、学習過程が学習者の実行機能に影響を与える可能性を検討する。その一例として、ある推論法を用いた学習方法が非言語的認知処理機能、特に一連の情報から課題目標に適切な情報のみを維持し効果的に問題解決する実行機能に影響を及ぼすという仮説をたて、外国語学習における教示法の認知処理への影響を検討する。

1.1.2　外国語学習前後における認知機能の変化

　前節で論じた言語機能に限らずその一般的な認知機能（general cognitive function）の差異をバイリンガルとモノリンガルという群間で比較する共時的研究とは別に、より直接的な方法として、ある外国語に熟達していない被験者群の外国語学習前後で一般的な認知機能の変化を測定する通時的研究がある。通時的研究は、外国語学習経験内容及びその開始・終了の統制、個人要因、熟達を確認しやすく、外国語学習経験が認知機能の変化を生むという因果関係を明らかにするために有効な方法のひとつであるが、バイリンガルとモノリンガルのある時点での対比といった共時的研究に比べ、その数はまだ少ない。

　これまで Bialystok や Bak らのグループが、ある外国語に熟達していない被験者群の外国語学習前後で、言語に限らない注意や抑制機能などの認知機能の評価をしている（Bak, Long, Vega-Mendoza, & Sorace, 2016; Bialystok & Barac, 2012; Janus, Lee, Moreno, & Bialystok, 2016; Sullivan, Janus, Moreno, Astheimer, & Bialystok, 2014）。その結果、ひとつの研究を除く多くの研究で統制群より顕著な認知機能の向上が認められた。また、顕著な認知機能の変化がみられなかった研究でも、非言語課題中の脳波計測をしたところバイリンガルの脳波パタンに類似した結果が得られたという。こうした結果が多く集まれば外国語学習経験による認知機能の変化は存在するという仮説を強く支持するものとなる。

　しかし、これまでの外国語学習実験では、外国語学習方法は様々で、学習群と比較する統制群が心理学を同時期に学ぶ群や音楽を学ぶ群、外国語学習以外の色々なコースで学ぶ群などであった。そのため、どのような外国語学習方法が認知機能の変化を生みやすいかといった知見を得ることは難しかった。このことは、前節でみたように、特定の言語学習過程が学習者の実行機能に影響を与えるかを問う本研究では大変重要な問題であるといえる。

1．2　外国語学習方法における学習様式

　言語学者 Pinker によれば言語学習とは語彙と規則の学習から成る（Pinker, 1999）。古くから言語の学習法について議論があり、特に言語規則をどのように教授するかをめぐっては多くの研究がある（Ellis, 2011; Robinson, 2008）。一般的に規則の学習は多く推論に基づいてなされることが多い。推論は与えられた命題（proposition）から新しい命題を導く方法を示すもので、知的活動を行う上で欠かせないものである。そのため多くの論理学者や心理学者は人間の行う推論と認知処理の関係に興味を持ってきた（Isaac, Szymanik, & Verbrugge, 2014; Johnson-Laird, Khemlani, & Goodwin, 2015）。そして最近の研究では、異なる推論の方法では、認知処理も異なる可能性を見出した（Goel & Dolan, 2004; Heit & Rotello, 2010）。

　推論の方法、つまり推論規則は二つに大別され、一方を演繹規則（演繹法；rules of deduction）と呼び、もう一方を帰納規則（帰納法；rules of induction）と呼ぶ。演繹規則によれば、真なる言語的発話（linguistic utterance）から他の真なる言語的発話が導かれ、帰納規則によれば、真なる言語的発話から、それが真であるかどうかはわかっていないがその確率がわかっているような仮に真である経験的言明が導かれる（Reichenbach, 1947）。一般に、演繹規則は既知の法則（premise）を個々の事象に適用し、帰納規則は個々の事象からの経験的観測により仮の法則を導き出すことができる。

　これを外国語における言語規則の学習方法に当てはめると、演繹的学習法では最初に言語規則が明示的に与えられ学習者は規則を個々の例に適用し、帰納的学習法では個々の例のみ与えられ学習者はこれらの経験的観察から暗示的に与えられた言語規則を学習する。演繹的学習法は言語規則が明示的に与えられる文法訳読法（grammar-translation method）に代表され、帰納的学習法は言語規則が個々の例のうちに暗示的に与えられる点で直接法（direct method）や AL 法（audiolingual method）に深く関連している（Corder & Allen, 1975）。学習法が明示的か暗示的かという点に特に着目し明示的学習法（explicit learning）や暗示的学習法（implicit learning）といった分け方をする場合もある。多くの研究によって、そのような外国語学習法の対照的な違いは着目され（Gollin, 1998）、その効果、つまり学習者の熟達度の向上ばかりが検討されてきたが、その点でもどちらの学習法の効果が高いかについて、意見は分かれている（Erlam, 2003）。

　本研究では二つの外国語学習法、即ち演繹的学習法と帰納的学習法による違いを実験的方法によって検討する。特に、これまでの外国語学習法の対照研究がすべて言語の熟達度の変化のみに焦点を当てていたのに対し、外国語学習に伴って生じる実行機能の変化に着目する。すなわち、非言語的課題においても、学習法によって学習前後の成績変化の度合いに違いが生じるのかどうかについて検討する。

２．外国語学習実験の方法

２.１　参加者

　本研究の参加者は、大学内のインターネットやポスターによる広報をみて応募した、神経学的既往歴がなく、右利き・日本語のみ母国語・外国で３カ月以上暮らしたことがない健康な17名の大学生である。１名は実験開始後に参加を取りやめた。参加者のドイツ語学習歴は最長で６カ月を超えずその平均は1.06ヶ月（標準誤差 0.44ヶ月）だった。参加者は（未成年は保護者共に）研究説明を受け、同意の下に研究参加を開始した。全ての研究参加に謝礼が支払われた。本研究は慶應義塾大学文学部研究倫理委員会により承認された（番号13006）。

２.２　手法

２.２.１　学習前後の評価方法

　学習前に参加者はドイツ語の熟達度をはかる二種類の課題と非言語的認知機能をはかる二種類の課題、合計四種類の課題に取り組んだ。ドイツ語の熟達度をはかる課題では、ドイツ語技能検定（German Diploma in Japan）４級・５級の2014年春と2015年秋の問題から20問と、自作のドイツ語文法性判断課題36問に回答した。また非言語的認知機能をはかる課題では、非言語的知能検査であるレーヴン漸進的マトリックス検査（Raven's Progressive Matrices; RPM検査）（Raven, 1941）の奇数問題30問と、自作の図形を用いた推論課題18問に回答した。

　学習後にも参加者は学習前と同様のドイツ語の熟達度をはかる二種類の課題と非言語的認知機能をはかる二種類の課題、合計四種類の課題に取り組んだ。ドイツ語の熟達度をはかる課題では、ドイツ語技能検定４級・５級の2014年秋と2015年春の問題から20問と、自作の別のドイツ語文法性判断課題36問に回答した。また非言語的認知機能をはかる課題では、非言語的知能検査である RPM 検査の偶数問題30問と、自作の別の図形を用いた推論課題18問に回答した。

　ドイツ語の熟達度をはかる課題の一方であるドイツ語技能検定は年度などによって難易度が異なるため、学習前後の課題間の難易度も異なることが予想された。しかし、学習前と学習後それぞれの群が同じ課題に回答することにより、前と後の時点の学習群間の差をはかること及び学習前後変化の両群間の差をはかることができることからもう一方の自作課題と併用することとした。また、非言語的認知機能をはかる課題の一方である RPM 検査は無作為に偶数問題と奇数問題を抽出したため、学習前後の課題間の難易度が異なることが予想された。しかし、同様に前と後の時点の学習群間の差をはかること及び学習前後の変化の両群の差をはかることができることから併用することとした。

　もう一方のドイツ語の熟達度をはかる課題であるドイツ語文法性判断課題は、統語範疇と語彙を同程度のものとすることで学習前後の課題間の難易度をおおよそ等しくするように努めた。各問題のドイツ語は文もしくは句の形で出題され、参加者はドイツ語文法表現が正しいか

正しくないかを二択で答えた。また、もう一方の非言語的認知機能をはかる課題である推論課題は、過去の推論課題研究（Crone et al., 2009）を改良し大きさや模様といった別の様式を用いながら推論に要求される計算の複雑性を統制することで学習前後の課題間の難易度をおおよそ等しくするように努めた。

　自作の推論課題についてさらに詳細を説明することとする。この課題は図形の形に着目しその並びにしたがって空欄となっている図形を推論するものである。各問題は上段の３行３列の図形の並びと下段の１行４列の図形の並びから成り、常に上段の３行３列目の右下の図形が点線の四角図形で空欄となっている。参加者は上段の３行３列目の右下の図形以外の図形の並びから、３行３列目の右下の図形が何であるべきか、最も適した図形を下段の四つの図形から一つ選んで答えることが求められた。また学習者の認知的特性をより明らかにするために、推論課題の計算の複雑性の水準を３つに分けた。

　図１は計算の複雑性の水準が低い問題の代表例である。上段の３行３列目の右下の図形以外の図形の並びは同じものが並んでおり、したがって３行３列目の右下の図形に最も適した図形は同じ図形である下段の左から２番目を選んで答えるのが正解である。図２は計算の複雑性の水準が中の問題の代表例である。上段の３行３列目の右下の図形以外の図形の並びは列ごとに同じものが並んでおり、したがって３行３列目の右下の図形に最も適した図形は上段の３列目の図形と同じ図形である下段の左から２番目を選んで答えるのが正解である。図３は計算の複雑性の水準が高の問題の代表例である。上段の３行３列目の右下の図形以外の図形は入れ子になっており、その並びは外側の図形は行ごとに、内側の図形は列ごとに同じものが並んでおり、したがって３行３列目の右下の図形に最も適した図形は外側が３行目と同じで内側が３列目と同じ図形である下段の左から３番目を選んで答えるのが正解である。課題の慣れを防ぐ目的で、形に着目する以外に図形の大きさや模様に着目する課題を用意した点が先行研究より優れた点であるが、紙面の都合上、形のみ例示した。
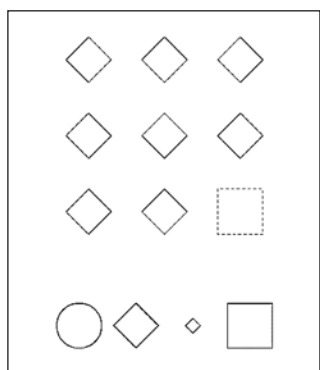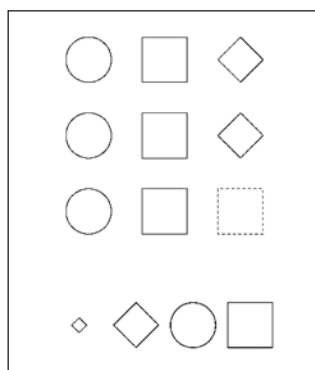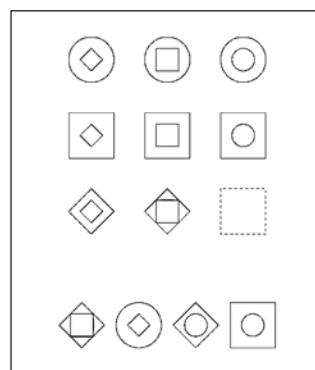
図１　計算の複雑性が低い問題	図２　計算の複雑性が中程度の問題	図３　計算の複雑性が高い問題

　各課題の成績は次のように求めた。ドイツ語技能検定は20問中の正答数を求めた。ドイツ語文法性判断課題における二択の弁別性の指標には、信号検出理論に基づいたディー・プライム（dprime; d'）指標を用いた。この弁別指標は d' = z（正答率（hit））− z（フォールス・アラーム率（false alarm））によって計算した（Macmillan & Creelman, 1991）。バイリンガルについての先行研究より青年期の認知機能をはかる問題では認知的負荷の高い問題でのみ統制群との差が出やすい（Bialystok, 2006; Costa, Hernandez, Costa-Faidella, & Sebastian-Galles, 2009）との知見があり、RPM 検査は後半の問題ほど難易度が上がる設計になっているので、30問を5分割し最後の6問の正答数を求めた。推論課題では要求される計算の複雑性が高い課題の正答数を求めた。いずれの課題も一問ごとにコンピュータの画面に問題を表示しながら、一問ごとの反応時間を計測し、反応時間の平均は正答の反応時間から計算した。参加者はいずれも自分のペースで問題を解いたが、ドイツ語技能検定と RPM 検査のときには無制限時間内でキーボード上の数字キーを押し、ドイツ語文法性判断課題と推論課題のときにはそれぞれ3.5秒の制限時間内と10秒の制限時間内でキーパッド上のキーを押して答えた。

2.2.2　学習群と学習内容

　参加者は、学習前に回答した四課題の成績・反応時間によって、非言語的知能とドイツ語の熟達度におおよそ差のない演繹的学習群8名と帰納的学習群9名の2群の学習群に分けられた。演繹的学習群8名の平均年齢は20.28歳（標準誤差 11.72カ月）、帰納的学習群9名の平均年齢は19.04歳（標準誤差 5.24カ月）で年齢についても両群は統計的な差異は認められなかった（t（15）＝1.21, *p＞0.24*）。

　両群ともに同じ教材で同じ内容を学習した。ただ一つ両群で異なる点は、演繹的学習か帰納的学習かといった学習様式の差だった。演繹的学習群は最初から言語規則を明示的に与えられ、それから例に言語規則を当てはめながら学習した。一方、帰納的学習群は最初に個々の例のみ与えられ、これらの経験的観察から暗示的に与えられた言語規則を推測しつつ学習した。

　学習に用いられる例は情報技術の発展と共にコーパス等から入手が容易になった言語データを言語教育に用いるデータ駆動型学習（data-driven learning; DDL）（Johns, 1991）に倣い、研究室内で用意した学習者のドイツ語レベルに統制したおよそ1700見出し語コーパスに基づいて作成され、クイズ形式などで学習者に示された。各群ともに3ヶ月間9回の60分授業と宿題を通じてドイツ語の名詞・冠詞・形容詞・動詞・助動詞・接続詞・前置詞を用いた初級文法を大量の例と共に学習した。DDL には学習者が言語データを自由に探索し学習に利用する拡散型（divergent）学習と、予め決められた学習目標に向けて導かれつつ言語データを学習に利用する収束型（convergent）学習があるが、今回は学習者が初習者であり、必ずしも全ての学習者がコーパス利用を得意としていなかったため、後者の方法を選択した。即ち、授業では

ドイツ語の例はコンピュータもしくはハンドアウトによって示され、宿題では参加者に貸与された iPod touches（Apple Inc.）という携帯型端末機器を通じて簡便なクイズ形式として例が示された。さらに授業内では学習の際の正誤をフィードバックとして与え、宿題の中では個々の問題に正誤のフィードバックと20問ごとの成績が表示された。学習者は次回の授業までの課外のすきま時間を生かして言語データを通じて言語規則を学習した。DDL は帰納的学習法に利用されることが多いが、演繹的学習法に利用されることも可能である。また、DDL の特徴として言語規則即ち文法から語彙情報が切り離されにくく、統合的に学習できる面がある。

2.2.3　検定方法

　以下のように統計的仮説の検定を行った。学習前後のドイツ語技能検定の正答数及び反応時間、ドイツ語文法性判断課題の弁別指標及び反応時間、RPM 検査の反応時間、計算の複雑性が高い推論課題の反応時間を、学習様式別で比較する場合には、学習前後を参加者内要因、学習様式を参加者間要因とする二要因分散分析を行った。ドイツ語文法性判断課題弁別指標の学習前後の差分、RPM 検査正答数の学習前後の差分を学習様式別で比較する場合には t 検定を行った。RPM 検査正答数、計算の複雑性が高い推論課題の正答数の学習前後差を比較する場合には Wilcoxon の符号付順位検定を行った。RPM 検査の学習前正答数および学習後正答数、計算の複雑性が高い推論課題の学習前正答数および学習後正答数、計算の複雑性が高い推論課題正答数の学習前後の差分、RPM 検査反応時間の学習前後の差分を、学習様式別で比較する場合には Mann-Whitney 検定を行った。

3．外国語学習実験の結果
3.1　外国語の熟達度の変化

　前章で述べたように、ドイツ語の熟達度の変化を測るために、学習前後で、ドイツ語技能検定の問題から20問と、ドイツ語文法性判断課題36問が用意された。

　ドイツ語技能検定の学習前後では二群をあわせた参加者全体の正答数の平均値の向上がみられ、学習前後の主効果は統計的に有意な傾向があった（$F_{(1,15)} = 3.83$, $p < 0.07$）。しかし、両群間の主効果もなく（$F_{(1,15)} = 0.03$, $p > 0.8$）、学習前後と群間の交互作用がなかったため（$F_{(1,15)} = 0.007$, $p > 0.9$）、両群の学習方法によるドイツ語の熟達度及びその変化には差がみられなかった。

　図4に各群のドイツ語文法性判断課題における弁別指標の学習前後変化を示す。実線で表されているのが帰納的学習群で、破線で表されているのが演繹的学習群である。図4にみられるように、ドイツ語文法性判断課題の学習前後では二群をあわせた参加者全体の弁別指標の平均値の向上がみられ、学習前後の主効果は統計的に有意であった（$F_{(1,15)} = 22.22$, $p < 0.001$）。

しかし、両群間の主効果もなく（F（1,15）＝0.07, *p＞0.7*）、学習前後と群間の交互作用がなかったため（F（1,15）＝0.01, *p＞0.4*）、両群の学習方法によるドイツ語の熟達度及びその変化には差がみられなかった。図5は各群の学習後の弁別指標から学習前の弁別指標を引いた結果である。図5にみられるように、両群の平均値の散らばりは重なっており、効果の差がみられないことがわかる（t（15）＝0.78, *p＞0.4*）。

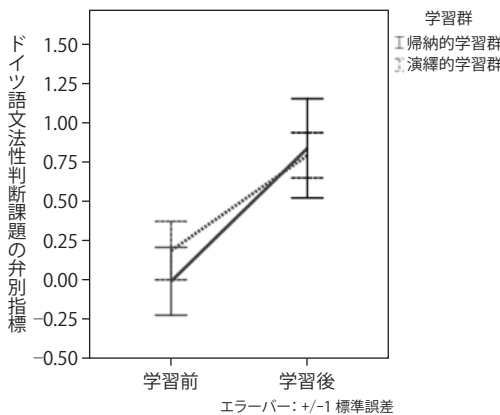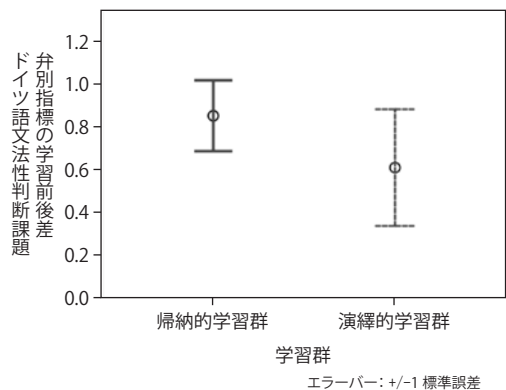**図4　文法性判断課題弁別指標の推移**　　**図5　文法性判断課題弁別指標の学習前後差**

　反応時間はドイツ語技能検定の二群をあわせた参加者全体で学習前後の変化がみられず（F（1,15）＝0.45, *p＞0.5*）、両群間の差もみられなかった（F（1,15）＝0.24, *p＞0.6*）。しかし、学習前後と群間の交互作用に有意傾向があり（F（1,15）＝3.33, *p＜0.09*）、演繹的学習群では反応時間は学習前で8.7秒程度だったのが学習後には7.5秒に短縮したのに対し、帰納的学習群では学習前で7.4秒程度だったのが学習後には9.8秒に延長したことが観察された。

　ドイツ語文法性判断課題の反応時間では二群をあわせた参加者全体で学習前後に有意に延長した（F（1,15）＝10.02, *p＜0.01*）。反応時間は学習前で1.9秒程度だったのが学習後には2.2秒に延長したことが観察された。しかし、両群間の差もみられず（F（1,15）＝0.34, *p＞0.5*）、学習前後と群間の交互作用もなく（F（1,15）＝0.08, *p＞0.7*）、二群の反応時間に差はみられなかった。

3.2　非言語的認知機能の変化

　前章で述べたように、非言語的認知機能とその変化をはかるために学習前後の RPM 検査最終6問と、計算の複雑性が高い推論課題6問を評価した。

　RPM 検査の二群をあわせた参加者全体の正答数の平均値は学習前後で差がみられなかっ

た（Z＝−1.53, *p＞0.1*）。また両群の正答数の平均値は学習前（U＝34.00, *p＞0.8*）も学習後（U＝36.00, *p＞0.99*）も差がみられず、学習後の正答数から学習前の正答数を引いた成績の向上も両群間で差がみられなかった（t（15）＝0.1, *p＞0.9*）。

　図6に各群の計算の複雑性が高い推論課題における正答数の学習前後変化を示す。実線で表されているのが帰納的学習群で、破線で表されているのが演繹的学習群である。図6にみられるように、計算の複雑性が高い推論課題の二群をあわせた参加者全体の正答数は学習前後で差がみられなかった（Z＝−0.26, *p＞0.7*）。また両群の正答数は学習前（U＝27.00, *p＞0.3*）も学習後（U＝25.5, *p＞0.17*）も差がみられなかった。図7は各群の学習後の計算の複雑性が高い推論課題の正答数から学習前のそれを引いた結果である。図7にみられるように、学習後の正答数から学習前の正答数を引いた成績の向上は帰納的学習群のほうが有意に大きかった（U＝15.5, *p＜0.05*）。
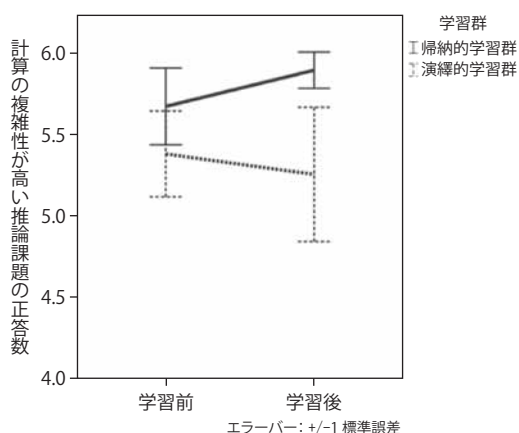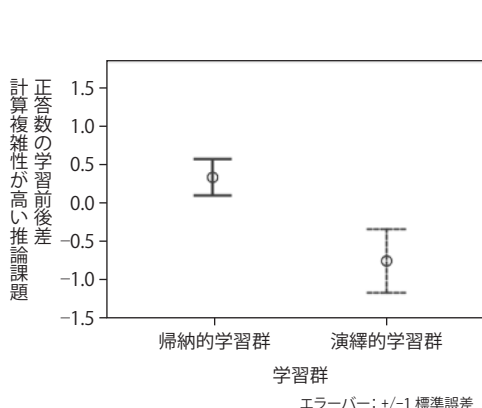
**図6　計算の複雑性が高い推論課題の
　　　正答数の推移**

**図7　計算の複雑性が高い推論課題正答数の
　　　学習前後差**

　図8に各群の RPM 検査最終6問における反応時間の学習前後変化を示す。実線で表されているのが帰納的学習群で、破線で表されているのが演繹的学習群である。RPM 検査最終6問の反応時間は二群をあわせた（正答がなかった参加者1名を除いた）参加者全体で学習前後の変化がみられず（F（1,14）＝1.08, *p＞0.3*）、両群間の差もみられなかった（F（1,14）＝0.13, *p＞0.7*）。しかし、学習前後と群間の交互作用があり（F（1,14）＝4.64, *p＜0.05*）、演繹的学習群では反応時間は学習前で21.6秒程度だったのが学習後には18.5秒に短縮したのに対し、帰納的学習群では学習前で17.2秒程度だったのが学習後には25.9秒に延長したことが観察された。図9は各群の学習後の RPM 検査最終6問の反応時間から学習前のそれを引いた結果である。

図9にみられるように、反応時間は帰納的学習群の方が有意に延長した（U＝14.0, *p＜0.04*）。

　計算の複雑性が高い推論課題の反応時間では二群をあわせた参加者全体で学習前後に差がみられなかった（F(1,15)＝0.56, *p＞0.4*）。また、二群の反応時間にも差はみられなかった。両群間の差もみられず（F(1,15)＝0.26, *p＞0.6*）、学習前後と群間の交互作用もなかった（F(1,15)＝0.08, *p＞0.7*）。
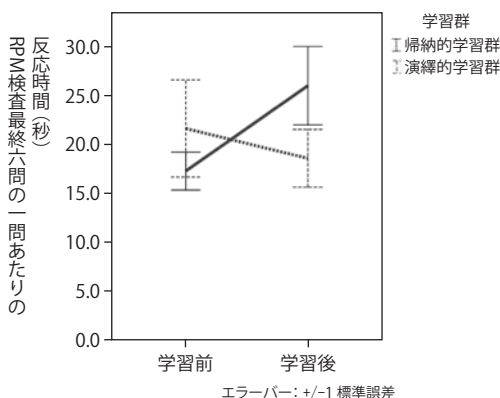

図8　RPM 検査最終六問反応時間の推移
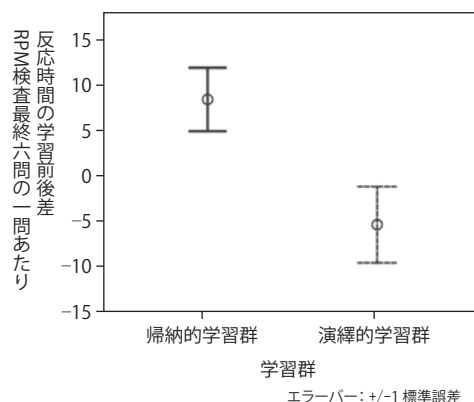
図9　RPM 検査最終六問反応時間の学習前後差

４．議論

　本論文では、外国語の言語規則を異なる学習様式で学習すると、学習者の外国語の熟達度だけでなく、非言語的認知機能にも違いが生じうるのか、バイリンガルや第二言語習得の先行研究を概観するとともに、日本語を母語とするドイツ語初習者を対象としたパイロット実験について検討した。

　実験の結果、外国語の熟達度を測るためのドイツ語技能検定とドイツ語文法性判断課題では学習前後で成績があがる傾向があり、その傾向は両群で差がないことがわかった。一方、学習前後の非言語的認知機能の向上は RPM 検査と推論課題の双方でみられなかったが、学習様式が演繹的か帰納的かによって推論課題の学習前後の変化に一部差がみられた。即ち、認知的負荷が大きいような計算の複雑性が高い推論課題では帰納的学習群の学習前後の伸びが高かった。また、課題の反応時間の分析から、外国語の熟達度を測る課題のときと、非言語的認知機能を測る課題のとき双方で、演繹的学習群は学習前よりも学習後のほうが短時間で解答したのに対し、帰納的学習群は学習よりも学習後のほうが長時間で解答する傾向がみられた。とくに、ドイツ語技能検定と RPM 検査のときにその有意な傾向や有意差がみられた。

　これらの結果から、外国語の熟達度に関しては、学習様式が演繹的か帰納的かに関わらず、両群の参加者はドイツ語の言語規則の学習を通じて、外国語の熟達度の差がなく、同様にドイツ語の成績が向上した可能性が高いといえる。非言語的認知機能に関しては、学習様式が帰納的のときに計算の複雑性が高い推論課題の成績の向上が一部みられた。そして、反応時間に関しては、外国語の熟達度に関する課題と非言語的認知機能を測る課題いずれも、学習を通じて、演繹的学習群は素早く、帰納的学習群はじっくり解答するようになった可能性が高い。これが、ドイツ語技能検定と RPM 検査のときに限られたのは、それ以外の二つの課題は解答するまでの時間に制限を与えたからであると思われる。

　総じて、今回の結果は、外国語の言語規則を異なる学習様式で学習すると、非言語的認知機能にも変化に違いが生じうることを支持する結果といえる。先行研究（Erlam, 2003）で述べられているとおり、異なる学習様式は外国語習得の熟達度には影響を与えるとはいえなかった。しかし、異なる学習様式は非言語的認知課題の成績向上に影響を与えることが示唆された。

　本研究の意義は、新しい言語を学ぶときの学習様式が学習者の実行機能に影響を与える可能性を指摘し、予備的な調査を行なった点にある。バイリンガルが優位性を持つような認知処理には何種類かのものがあることを提唱した先行研究では、メタ言語的気づきには言語の熟達度が、実行機能にはバイリンガル教育の年数が影響を与えることが指摘されている（Bialystok & Barac, 2012）。しかしバイリンガルの実行機能にバイリンガル教育におけるどのような過程が影響したのかは不明であった。またバイリンガルであることが実行機能の成長を促すのではないといった報告のなかでも、将来必要な研究として、ある特定のバイリンガル経験が実行機能に影響する可能性の追求が示唆されていた（Paap et al., 2015）。こうした問いに答えを与えるためには、教育の過程で特定の学習方法を設定し個々の実行機能の変化を計測することが重要であることを本研究は示したといえる。

　一方、本研究は様々な要因で予備的実験であるといえる。まず、参加者の数が少なく、また課題数も少ないため、説明能力が十分とはいえない。かつ、自発的に学習実験に応募した参加者で平均的なドイツ語初習者を代表する参加者群でない可能性もある。今後は効果に関して年齢やどのような認知特性や学習スタイル・経験を持っているのかという参加者のさらに詳細な個人要因の検討も必要と思われる。さらに学習後の評価においてドイツ語課題と非言語課題が同セッションで行われているためドイツ語の授業の体験を想起しながら非言語課題に取り組んだ可能性もあり、課題の施行を別々の日に行うことが必要である。実行機能の向上が長期的に続くのかも確認が必要である。

　今後はこうした点を修正して研究を発展させる必要がある。まず、外国語学習が、新しい言語の知識やスキルの獲得を促すに留まらず、思考や行動様式そのものを変化させること、さらにそれは学習様式によって違った方向性に変化しうるという可能性を追及する必要がある

と考える。この問題は、心理学や神経科学分野における、経験の結果として認知機能や脳の変化が起こり得るのかという問題にも通じる重要な課題である。また、非言語的認知機能を向上させるような外国語学習方法によって得られるであろう学習者の技能が、一般的な学習技能（learning skill）に通ずる汎化可能なものなのかといった点も熟考すべき問題である。つまり外国語学習によって培った学習者の技能が、より一般的な学習技能として教科間で転移（transfer）しうるかといった検討である。このためには教科横断的取り組みも必要になると思われる。

　後者の一般的な学習技能の習得の重要性は多くの教育研究の場で客観的な事実を蓄える際に獲得されるどのように学ぶかを学ぶという方略、方法知（learn how to learn）といった表現で論じられている（Halász & Michel, 2011）。外国語教育研究の場で一般的な学習技能に関連が深いのは、例えば、ヨーロッパ言語共通参照枠（Common European Framework of Reference for Languages）（Council of Europe, 2001）で、コミュニケーション言語能力（communicative language competences）の他に必要で学習者個人にかかわる一般的な能力（general competences）の一種として紹介されている学習能力（savoir-apprendre, ability to learn）である。学習能力は言語学習の動機を高め、学習方略を豊かにするもので、その中にはメタ言語的気づき（metalinguistic awareness）に代表されるような、言語とコミュニケーションに関する意識、一般的な音声意識と技能の他、勉強技能、発見技能が含まれる。また第二章で紹介した DDL がもともと目指している能力にこの方法知は含まれ、確認（identify）、分類（classify）、一般化（generalize）を通じて学習者の自律的な思考様式を促進させる（Johns, 1991）。

　外国語学習では、このような学習能力が学習者の学習の基礎に育つことで、言語間、つまり新しい言語や読解のより容易な学習、使用が可能になると考えられている（de Sousa, Greenop, & Fry, 2010; Kuo & Anderson, 2012）。また複言語を用いることでそれが促進されるという報告（Jessner, 1999）もある。一方、外国語学習によって培った学習者の技能がより一般的な学習技能として教科間で転移可能かどうかの研究は音楽（Moreno, Lee, Janus, & Bialystok, 2015; Patel, 2003）、算数（Vukovic & Lesaux, 2013b）、統計や幾何（Vukovic & Lesaux, 2013a）、科学（Mayer, 2004）についていくつか存在するが、今後より多くの研究によって証明されなければならないといえる。

　また、外国語学習技能がより一般的な学習技能として汎化可能である場合があるならば、それはなぜなのか、についても明らかにされなければならない。例えば音楽と言語は別々の表現型だが認知資源を共有しているからといった説明（Patel, 2003）や、メタ手続き的なより抽象的な言語能力に依存するといった説明（Karmiloff-Smith, 1979）があるが今後の実証が待たれる。

　さらに、どのような方法で学習能力を伸ばしていけばよいかについての研究も必要である。このような能力は、学ぶべき内容、ものごとの構成と使用の原理原則に関する客観的な事実や知識の学習や理解なしには培われず、豊富な知識に根差した上で、それまでの知識に新しい経験を統合（synthesize）することで、相互補完的に伸長するものと考えられている（Perkins & Salomon, 1989）。そのため、学習時には対象の使用に対する理解や感性を伸ばすような内容を使用すべきであり、機械的な暗記学習、空虚な内容を提示し方略のみを明示的に学ばせるような学習、演繹的学習法や帰納的学習法への偏りも避けるべきであろう。そして最適な学習方法は学習者個人により異なりうるという点も考慮すべき点である。

　以上のような問題は外国語学習や教科学習をどのようにすべきかといった教育政策においても重要な課題だと考えられる。

引用文献

Adesope, O. O., Lavin, T., Thompson, T., & Ungerleider, C. (2010). A Systematic Review and Meta-Analysis of the Cognitive Correlates of Bilingualism. *Review of Educational Research, 80*(2), 207-245. doi:10.3102/0034654310368803

American Speech-Language-Hearing Association. (2004). Knowledge and Skills Needed by Speech-Language Pathologists and Audiologists to Provide Culturally and Linguistically Appropriate Services. *Knowledge and Skills*. Retrieved from http://www.asha.org/policy/ks2004-00215.htm

Bak, T. H., Long, M. R., Vega-Mendoza, M., & Sorace, A. (2016). Novelty, Challenge, and Practice: The Impact of Intensive Language Learning on Attentional Functions. *PLoS One, 11*(4), e0153485. doi:10.1371/journal.pone.0153485

Bialystok, E., Craik, F. I. M., & Luk, G. (2012). Bilingualism: Consequences for Mind and Brain. *Trends in Cognitive Sciences, 16*(4), 240-250. doi:10.1016/j.tics.2012.03.001

Bialystok, E. (2006). Effect of bilingualism and computer video game experience on the Simon task. *Can J Exp Psychol, 60*(1), 68-79.

Bialystok, E., & Barac, R. (2012). Emerging bilingualism: Dissociating advantages for metalinguistic awareness and executive control. *Cognition, 122*(1), 67-73. doi:10.1016/j.cognition.2011.08.003

Corder, S. P., & Allen, J. P. B. (1975). *Papers in applied linguistics*. London: Oxford University Press, 1975.

Costa, A., Hernandez, M., Costa-Faidella, J., & Sebastian-Galles, N. (2009). On the bilingual advantage in conflict processing: now you see it, now you don't. *Cognition, 113*(2), 135-149. doi:10.1016/j.cognition.2009.08.001

Council of Europe. (2001). *Common European Framework of Reference for Languages: Learning, Teaching, Assessment*. Cambridge, U.K.: Cambridge University Press.

Crone, E. A., Wendelken, C., Van Leijenhorst, L., Honomichl, R. D., Christoff, K., & Bunge, S. A. (2009). Neurocognitive development of relational reasoning. *Developmental Science, 12*(1), 55-66. doi:10.1111/j.1467-7687.2008.00743.x

Cummins, J. (1979). Linguistic Interdependence and the Educational Development of Bilingual Children. *Review of Educational Research, 49*(2), 222-251. doi:10.3102/00346543049002222

de Sousa, D. S., Greenop, K., & Fry, J. (2010). The effects of phonological awareness of Zulu-speaking children learning to spell in English: A study of cross-language transfer. *British Journal of Educational Psychology, 80*(4), 517-533. doi:10.1348/000709910X496429

Ellis, R. (2011). *Language Teaching Research and Language Pedagogy*. West Sussex, U.K.: Wiley-Blackwell.

Erlam, R. (2003). The Effects of Deductive and Inductive Instruction on the Acquisition of Direct Object Pronouns in French as a Second Language. *The Modern Language Journal, 87*(2), 242-260. doi:10.1111/1540-4781.00188

Goel, V., & Dolan, R. J. (2004). Differential involvement of left prefrontal cortex in inductive and deductive reasoning. *Cognition, 93*(3), B109-121. doi:10.1016/j.cognition.2004.03.001

Gollin, J. (1998). Deductive vs. inductive language learning. *ELT Journal, 52*(1), 88-89. doi:10.1093/elt/52.1.88

Halász, G., & Michel, A. (2011). Key Competences in Europe: interpretation, policy formulation and implementation. *European Journal of Education, 46*(3), 289-306. doi:10.1111/j.1465-3435.2011.01491.x

Heit, E., & Rotello, C. M. (2010). Relations between inductive reasoning and deductive reasoning. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 36*(3), 805-812. doi:10.1037/a0018784

Isaac, A. M. C., Szymanik, J., & Verbrugge, R. (2014). Logic and Complexity in Cognitive Science. In A. Baltag & S. Smets (Eds.), *Johan van Benthem on Logic and Information Dynamics* (pp. 787-824). Cham: Springer International Publishing.

Janus, M., Lee, Y., Moreno, S., & Bialystok, E. (2016). Effects of short-term music and second-language training on executive control. *J Exp Child Psychol, 144*, 84-97. doi:10.1016/j.jecp.2015.11.009

Jessner, U. (1999). Metalinguistic Awareness in Multilinguals: Cognitive Aspects of Third Language Learning. *Language Awareness, 8*(3-4), 201-209. doi:10.1080/09658419908667129

Johns, T. (1991). Should you be persuaded. Two samples of data-driven learning materials. *ELR Journal, 4*, 1-16.

Johnson-Laird, P. N., Khemlani, S. S., & Goodwin, G. P. (2015). Logic, probability, and human reasoning. *Trends in Cognitive Sciences, 19*(4), 201-214. doi:10.1016/j.tics.2015.02.006

Karmiloff-Smith, A. (1979). Micro- and Macrodevelopmental Changes in Language Acquisition and Other Representational Systems. *Cognitive Science, 3*(2), 91-118. doi:10.1207/s15516709cog0302_1

Kuo, L.-J., & Anderson, R. C. (2012). Effects of early bilingualism on learning phonological regularities in a new language. *Journal of Experimental Child Psychology, 111*(3), 455-467. doi:10.1016/j.jecp.2011.08.013

Macmillan, N. A., & Creelman, C. D. (1991). *Detection theory : a user's guide.* Cambridge, U.K.: Cambridge University Press.

Mayer, R. E. (2004). Teaching of Subject Matter. *Annual Review of Psychology, 55*(1), 715-744. doi:10.1146/annurev.psych.55.082602.133124

Miyake, A., Friedman, N. P., Emerson, M. J., Witzki, A. H., Howerter, A., & Wager, T. D. (2000). The unity and diversity of executive functions and their contributions to complex "Frontal Lobe" tasks: a latent variable analysis. *Cogn Psychol, 41*(1), 49-100. doi:10.1006/cogp.1999.0734

Moreno, S., Lee, Y., Janus, M., & Bialystok, E. (2015). Short-Term Second Language and Music Training Induces Lasting Functional Brain Changes in Early Childhood. *Child Development, 86*(2), 394-406. doi:10.1111/cdev.12297

Paap, K. R., Johnson, H. A., & Sawi, O. (2015). Bilingual advantages in executive functioning either do not exist or are restricted to very specific and undetermined circumstances. *Cortex, 69*, 265-278. doi:10.1016/j.cortex.2015.04.014

Patel, A. D. (2003). Language, music, syntax and the brain. *Nat Neurosci, 6*(7), 674-681. doi:10.1038/nn1082

Peal, E., & Lambert, W. E. (1962). The relation of bilingualism to intelligence. *Psychological Monographs: General and Applied, 76*(27), 1-23. doi:10.1037/h0093840

Perkins, D. N., & Salomon, G. (1989). Are Cognitive Skills Context-Bound? *Educational Researcher, 18*(1), 16-25. doi:10.3102/0013189x018001016

Pinker, S. (1999). *Words and rules : the ingredients of language* (1st ed. ed.). New York: Basic Books.

Raven, J. C. (1941). Standardization of progressive matrices, 1938. *British Journal of Medical Psychology, 19*(1), 137-150. doi:10.1111/j.2044-8341.1941.tb00316.x

Reichenbach, H. (1947). *Elements of symbolic logic*. New York: Macmillan Co.

Ricciardelli, L. A. (1992). Bilingualism and cognitive development in relation to threshold theory. *Journal of Psycholinguistic Research, 21*(4), 301-316. doi:10.1007/bf01067515

Sullivan, M. D., Janus, M., Moreno, S., Astheimer, L., & Bialystok, E. (2014). Early stage second-language learning improves executive control: evidence from ERP. *Brain Lang, 139*, 84-98. doi:10.1016/j.bandl.2014.10.004

W. E., Pratt, C., Herriman, M. L., & Bowey, J. (1984). *Metalinguistic Awareness in Children: Theory, Research, and Implications.* Berlin: Springer-Verlag.

Vukovic, R. K., & Lesaux, N. K. (2013a). The language of mathematics: Investigating the ways language counts for children's mathematical development. *Journal of Experimental Child Psychology, 115*(2), 227-244. doi:10.1016/j.jecp.2013.02.002

Vukovic, R. K., & Lesaux, N. K. (2013b). The relationship between linguistic skills and arithmetic knowledge. *Learning and Individual Differences, 23*, 87-91. doi:10.1016/j.lindif.2012.10.007

慶應義塾の中国語教育における高大連携（1）
── 全塾懇談会の理念と既習者の扱い

須　山　哲　治

山　下　一　夫

吉　川　龍　生

一、はじめに

　慶應義塾の一貫教育校・大学で中国語教育を担当する教員の有志が集まり「慶應義塾全塾中国語担当者懇談会（以下、全塾懇談会）」の第1回会合を開催したのは、2012年の３月のことであった。従来、個別に連絡を取ることはあっても、全体で集まることはなかった塾内の各部門の教員が、ここで初めて集合し、塾内の中国語教育の様々な問題について意見交換を行った。

　継続的に会合が開催されるか未定であった第１回会合には、文学部・経済学部・法学部・商学部・理工学部の中国語担当者と、慶應義塾高等学校・慶應義塾女子高等学校・慶應義塾志木高等学校・慶應義塾湘南藤沢中・高等部の「一貫教育校」[i] の授業担当者（女子高等学校は主事も）が参加し、以降の継続的な懇談会開催と、塾内の各部門から最低１人出席することを目指すという方針を確認した。第２回以降では、湘南藤沢キャンパス３学部の中国語担当代表者が参加し、慶應義塾ニューヨーク学院の担当者から送られてきたカリキュラムや履修状況についての情報が紹介され、芝共立キャンパスの薬学部の中国語担当者とも情報共有が図られるようになっていった。また、2016年度から慶應義塾大学外国語教育研究センターの公的な会合として位置づけられるようになった。会場は、参加教員の交通の便を考え、それぞれのキャンパスの「中間地点」ということで、毎回日吉キャンパスが使用されている。第１回以降、各回の詳細は以下の通りである。

　　第１回　2012年３月２日（金）17:00～19:00
　　　　　　日吉キャンパス慶應義塾高等学校 B 棟地下 1F 第３会議室
　　第２回　2012年６月６日（水）18:10～20:30　日吉キャンパス来往舎 2F 大会議室
　　第３回　2013年６月28日（金）18:15～20:30　日吉キャンパス来往舎 2F 大会議室
　　第４回　2014年６月27日（金）18:30～20:30　日吉キャンパス来往舎 2F 大会議室
　　第５回　2015年６月25日（木）18:30～20:30　日吉キャンパス来往舎 2F 大会議室
　　第６回　2016年６月30日（木）18:30～20:30　日吉キャンパス来往舎 2F 中会議室

　全塾懇談会の開催は、各部門の教員によって情報共有の必要性が強く意識されたことが大きな理由となっているが、その中でも特に重要なのが、高校と大学の間の情報交換である。一般の高校では多くの場合、外国語は英語しか学習しないため、大学の第二外国語教育は基本的に未習者を想定して設計されている。しかし一貫教育校では第二外国語の教授が行われており、学生が大学でも同じ言語を履修した場合、学部では未習者とは異なる対応が必要になる。多くの学部では、そうした「既習者」を対象とした特別クラスを設置しているが、クラスの規模を決めるには次年度に何人くらいそうした学生が来るのかを予測しておく必要がある。また、授業設計のために一貫教育校でどのような教科書を使い、どのような授業を行ったのかを知っておく必要もある。さらに高校側も、自分たちが送り出す生徒たちが大学でどのように継続して中国語を学習していくのか、また大学は自分たちにどのような教育を期待しているのかといった問題を知らなければならない。

　近年さかんに「高大連携」が喧伝されているが、高校と大学の連携と言っても、高校の教員が自分たちの生徒を送り出した先の大学の教員と直接話し合いを持つことは、生徒の進路の多様性を考えると一般には困難を伴う。さらに「連携」を行おうにも、第二外国語教育の場合は実施している高校が非常に少ないため、大学側は対象者・対象校を見つけることが難しい。しかし慶應義塾であれば、一貫教育校で第二外国語教育を行っている上に、卒業生の大部分がそのまま大学に進学して来るため、直接的な連携が可能となる。その意味で、全塾懇談会の試みは、「高大連携」の活動の一つと捉えることもできるだろう。

　そこで本稿では、全塾懇談会の試みを通して検討された、慶應義塾の中国語教育における「既習者」問題の現状の一端を報告することで、第二外国語教育における高大連携について考えてみたいと思う。なお、全塾懇談会参加者でも、もちろんそれぞれの教育に対する考え方は異なっており、全体の総意のようなものを形成しているわけではないし、またその必要もない。本稿は全塾懇談会の許可を得た上で執筆者が個人の発想ないし見解を述べるものであり、内容についても全塾懇談会ではなく、執筆者が責任を負うものであることは諒解されたい。

二、一貫教育校と既習者

　慶應義塾の一貫教育校には、初等教育を行う幼稚舎・横浜初等部、前期中等教育を行う普通部・中等部・湘南藤沢中等部、後期中等教育を行う高等学校・女子高等学校・志木高等学校・湘南藤沢高等部・ニューヨーク学院がある。この中で、初等教育や前期中等教育を担当する学校でも、単発的な講座を実施したり、中華圏の学校と交流を行ったりと、中国語に関わる教育が行われていて、慶應義塾全体の中国語教育の中で果たしている役割は決して小さくないと思われるが、ひとまず本稿では検討の範囲としない。

　大学教育と直接接続しているのは、後期中等教育に関わる５つの学校である。いずれも卒業

生は原則的に各学校長の推薦によって全員が慶應義塾大学に進学する。これらの学校で行われている中国語授業の2016年度の授業設置状況は以下の通りである（なお本項での１コマは50分である）。

　（1）高等学校

　　● ２年生：全員が必修選択科目として第二外国語を履修、週２コマ

　　● ３年生：希望者が自由選択科目として第二外国語を履修、週３コマ

　（2）女子高等学校

　　● ２・３年生：全員が必修選択科目として第二外国語を履修

　　➢ ２年生：週２コマ

　　➢ ３年生：週３コマ

　（3）志木高等学校

　　「総合的な学習の時間」で、２年生は設置されている24言語の中から興味のある語種を選んで履修、半年間週２コマ。またそれ以外に、自由参加である１年間週１コマの「語学課外講座」も設置。

　（4）湘南藤沢高等部

　　● ３年生：全員が必修選択科目として第二外国語を履修

　　➢ Ⅰ類（文系科目主体）：週４コマ

　　➢ Ⅱ類（理系科目主体）：週２コマ

　（5）ニューヨーク学院

　　選択科目の「外国語」で、２・３年生共に「中国語Ⅰ」および「中国語Ⅱ」を履修可能。各学年１年間４コマ。なお「外国語」は、中国語以外ではスペイン語のみ設置。

　次に大学の各学部の状況を見てみよう。多くの学部で第二外国語は選択必修科目、すなわち各語種の中から一つ以上を履修し、進級・卒業要件となる科目として設置されている。またそれ以外に、必修ではない選択科目なども多数存在する。前者についての各学部の中国語の設置状況は以下の通りとなる（なお本項での１コマは90分である。また、文学部・経済学部・法学部の３学部には「通信教育課程」も設置されているが、本稿では扱わない。なお、経済学部の通信教育課程には中国語は設置されていない）。

（1）文学部

- 1年：週3コマ
- 2年：週2コマ

「第二外国語選択必修」ではなく、英語・ドイツ語・フランス語・中国語・朝鮮語・スペイン語・ロシア語・イタリア語・日本語（留学生のみ）の中からの「2言語」の選択必修。

（2）経済学部

- 1年：週3コマ6単位
- 2～4年：2～4単位（2単位＝通年週1コマ / 半期集中週2コマ）

2年次に第三外国語として1年次に履修した第二外国語以外の言語を週2コマ4単位履修可。

（3）法学部

- レギュラーコース：1～2年：週2コマ
- インテンシブコース：1～2年：週4コマ、3年：週3コマ。

（4）商学部

- 1～2年：週2コマ
- インテンシブコース：1～2年：週3コマ

「第二外国語選択必修」ではなく、英語・ドイツ語・フランス語・中国語・スペイン語の中からの「2言語」の選択必修。

（5）理工学部

- 1年：週2コマ

（6）総合政策学部・環境情報学部

「第二外国語選択必修」ではなく、「1年生・2年生で外国語科目を一定単位以上履修」（2014年度以降の入学者）という制度。そのうち中国語は、インテンシブコース1期・2期・3期・4期（各半年間週4コマ）、ベーシックコース1期・2期・3期（各半年間週2コマ）。

なおこれ以外の学部だと、医学部では1年生の時に第二外国語の履修が必修となっているが、ドイツ語かフランス語しか選べず、また薬学部や看護医療学部では第二外国語は必修とはなっていない。

上記の必修科目の履修者は、前章でも述べたとおり大多数が未習者であるが、大学入学以前に何らかの形で中国語を学んだ、以下のような「既習者」もいる。

1．一貫教育校で中国語の授業を履修した者
2．一貫教育校以外の高等学校、あるいはそれに相当する国内外の後期中等教育学校で中国

　　語の授業を履修した者

　３．専修学校・市民講座・放送講座などで中国語を学習した者

　４．中国語母語・準母語話者（いわゆる「ネイティブ」「ニア・ネイティブ」）

　この中で、圧倒的多数を占めるのはやはり１である。2016年度で見ると、大学学部新入生約6500名のうち、一貫教育校出身者は1400名弱である。従って、中には高校で選んだのとは異なる言語を大学で学ぶ学生も多少存在するものの[ii]、単純計算すると大学の必修中国語履修者の２割程度を１が占めることになる。これに対して、２〜４は相対的に人数が少ない。そのため、学部の既習者クラスも基本的には一貫教育校出身者を念頭に設計し、２や３の学生については１に合わせて割り振るという方式が採られている（なお４の中国語母語・準母語話者の問題については後述）。

　この既習者クラスの設計に、全塾懇談会で共有された情報が役立てられることが想定されている。具体的には、一貫教育校での履修者数の動向に基づいて、どのくらいの学生が集まるか事前にある程度の予測を立て、クラス規模を調整している。加えて、高校側でどのような教科書を用い、どのようなアプローチで授業を行ったかをもとに、学習内容がこれとうまく接続するよう、既習者クラスの授業内容を調整するのである。

三、大学における既習者クラスの設定

　さて、各学部の必修外国語科目における中国語既習者クラスの設置状況は以下の通りである。

（1）文学部：１年生週３コマの必修では、既習者をレベルによって「中級クラス」と「上級クラス」とに、また２年生週２コマの必修では「上級クラス」と「最上級クラス」とに振り分ける（１年生「上級クラス」と２年生「最上級クラス」は、外国語教育研究センター設置科目、後述）。

（2）経済学部：１年週３コマの必修では、同レベルの既習者がなるべく同じクラスになるように、年度ごとに状況を見て調整を行う。

（3）法学部：１〜２年週２コマのレギュラーコースでは、既習者を「既習者クラス」と外国語教育研究センター設置科目（後述）とに振り分ける。

（4）商学部：既習者クラスは設置していない。既習者については、本人の希望とレベルによって、外国語教育研究センター設置科目などを履修させる。

（5）理工学部：１年週２コマの必修では、既習者を「既習者クラス」に振り分ける。

（6）総合政策学部・環境情報学部：インテンシブコース１期・２期・３期・４期（各半年間週４コマ）、ベーシックコース１期・２期・３期（各半年間週２コマ）は、どちらも学

　　部独自の「検定試験」に合格することでコース途中からの履修を認めており、既習者に
　　はこの制度を利用して自分にあったレベルを選択させる。

　大学の授業で言えば、例えば歴史や漢文のような科目でも、未習者と既習者の間には大きな
相違があるはずだが、そうした授業では既習者クラスのようなものは設定されていない。にも
かかわらず中国語で既習者クラスが設定されているのは、初級・中級・上級と段階的な学習を
設計している以上、初級をすでに終えた学生には中級の授業を提供すべきであるという教学理
念に基づいている。そしてそれを可能としているのは、一貫教育校を擁することで常に一定数
の既習者が入学して来るという、慶應義塾に特有の環境である。また、中国語は進級・卒業要
件に関わる科目となるため、学生たちに対して成績の上での公平性を担保するという問題も関
わっている。
　これら一貫教育校出身者を主体とする既習者は、学習者全体の中ではあくまでも少数派であ
る。しかし、人材育成という観点からは、決しておろそかにしてはならない存在である。一つ
の外国語を、研究や仕事に用いるのに充分な水準まで引き上げるには、３年間〜４年間の時間
が必要だが、未習者の場合、そこまで到達するのに学部の時間のほとんどを使うことになる。
こうした事情は、英語一辺倒を唱える一部の人々にとって、「コストパフォーマンスの悪さ」
として映り[iii]、「第二外国語は結局習得できないのだから、その時間は英語の強化に使うべき
だ」などという主張の根拠とされてしまう。しかし既習者であれば、大学入学の時点で１年〜
２年の学習を経てきているため、未習者よりも在学中の早い段階で高い水準に到達できる可能
性が高い。また、そうした学生が周囲に存在することは、大学入学後に当該言語の学習を始め
た学生たちにとっても大きな刺激となり得る。

四、中国語母語・準母語話者
　前章で触れた４の中国語母語・準母語話者の存在の問題も、全塾懇談会を通じてより明確に
なったことである。従って、このことも全塾懇談会の成果の一つと言えるので、併せてここで
紹介しておきたい。
　まず、中国語母語・準母語話者の存在は、中国語教育において特徴的な問題といえる。もち
ろん、大学で第二外国語として設置されているどの言語であっても、いわゆる「ネイティブ」
の学生は存在する。ただ、日本国内においては近年、中華圏にルーツを持つ人口が増加傾向に
あるため、例えばフランス語やドイツ語などのヨーロッパ諸語に比べると、中国語の母語・準
母語話者は桁外れに多い。
　こうした中国語母語・準母語話者が大学の第二外国語で中国語を履修することを許可しない
という発想は当然あり得る。「楽に単位を取得するためだろう」というわけである。もちろん、

そうした場合もかなり存在することは事実だが、ただ全員が「不純な動機」に基づくというわけでもないようである。中国語母語・準母語話者と言っても、中国語の能力にはかなりの開きがある。その中には中等教育を日本語で受けたため、中国語を話すことはできても読み書きに不安があるから、大学の必修科目で第二外国語を勉強しなければならないのであれば、ぜひそれを利用して自分の中国語を強化したいと、真剣に考えて履修しようとする学生も少なくない。もちろん、そうしたあり方は必修選択外国語科目の理念に反するとして認めていない学部もある。しかし一方で本人のいわばセールスポイントとなる中国語をサポートすることは、教育上決して悪い選択ではないという考え方もあるだろう。

　問題は、そうした場合の受け皿である。この種の学生は、一般の日本人学生では数年かけても到達が難しいような、かなり高度な会話力を有している。しかしその反面、中国語のローマ字発音表記などは不得意なこともあり、高校から中国語学習を始めた学生を前提とする既習者クラスでは対応が難しくなることが多い。

　だからと言って、ローマ字発音表記を未習者クラスで学習させようとしても、うまくいかない。当人にとっては「発音は理解しているが、どう表記するのか解らない」という問題なのに対し、未習者クラスの授業では「中国語の発音を習得するための道具」として使われるので、両者の間には大きな齟齬が生じるからである。また中国語母語・準母語話者と言っても、教科書で教えるような「標準的な発音」の学生は稀で、何らかの方言的要素を有していることが多いが、そうした学生に教員が発音矯正を行うのはあまり意味がない。加えて、さらに事情を知らない未習者がそれを見て、「中国人の発音が教科書と違う」ことに違和感を覚え、授業への信頼をなくしたりすることもあり、教育効果は決して高くない。

　唯一の解決策としては、こうした中国語母語・準母語話者に特化した授業を別に作るということになる。ただ、現状では学生数が少なく、学部単独では設置が難しいため、学部横断組織である慶應義塾大学外国語教育研究センターの特設科目として開講し、各学部が乗り入れるという形にしている。中国語母語・準母語話者に必修選択第二外国語として中国語の履修を認めていない学部も、そうした学生に対しては選択科目として履修することを勧めている。

　なおこの問題は、一見すると「高大連携」とは関係しないように思われるかもしれないが、実はそうではない。全塾懇談会の結果、各一貫教育校にもこのような中国語母語・準母語話者の生徒が存在することが明らかになったからである。彼らは、塾全体における中国語母語・準母語話者の数から見れば、割合としては少数かもしれないが、しかし確実に一定数存在することもまた事実である。

　一貫教育校の中国語母語・準母語話者の生徒の中には、自らの中国語能力に対して、完全な自信を持っていない者も当然ながら存在する。特に、上述のように「日常会話は問題ないが、読み書きは苦手」という者が多い。そうした生徒の多くが、大学入学後に自分の中国語入力の

不足を補ってくれるような中国語の教育が受けられるのか否か、不安を感じている。上述した、大学各学部の中国語母語・準母語話者への対応状況を一貫教育校と情報共有することで、こうした生徒への対応もより十全なものとなり得る。これもまた全塾懇談会の「高大連携」に対する成果の一つと言えよう。

五、おわりに

「高大連携」は、1999年に中央教育審議会が提出した答申（中央教育審議会（1999））を契機とする。答申で触れられている具体的な方法をもとに、「高校生の大学での聴講」、「大学教員の高校での出張授業」、「高校生向けの大学のオープンキャンパス」といった動きが2000年代前半に急速に拡大した[iv]。しかし、高校側は生徒の進学率を高める材料として、また大学側は入試広報活動の一環として利用しているだけという側面もあり、これによって「大学教育へのスムーズな移行」という高大連携の理念が本当に「達成」されたと、手放しで賞賛できるような性質のものではないように思われる[v]。

教育研修事業財団（2011）は、こうした活動の問題点として、（1）高校生は正規の授業以外の時間で参加することになり、様々な点で負担となること、（2）高校教員との個人的な繋がりに基づきボランティアで行われることが多いため、そうした関係を有する一部の大学教員に負担が集中する傾向にあること、（3）高校生の学力向上が主眼となっており、大学生の学力向上を目指すという視点が欠けていること、の3点を指摘している。

注目すべきは（3）であろう。学習者が高校で学んだ内容をスムーズに大学での学習に移行できるカリキュラムを、教育目標の異なる高校と大学が違いを乗り越え、協力して作り上げていく全塾懇談会の試みは、この問題に対する答えの一つだと言えるだろう。もちろん、まだまだ暗中模索といった側面もあるが、今後も着実に続け、発展させて行きたいと考えている。

なお本稿では、字数の関係もあって、全塾懇談会で議論されている既習者クラスの使用教材や、授業内容の設計の問題については詳しく触れることができなかった。こうした問題については、また稿を改めて検討したい。

参考文献

中央教育審議会（1999）「初等中等教育と高等教育との接続の改善について（答申）」、http://www.mext. go.jp/b_menu/shingi/chuuou/toushin/991201.htm、2016年 8 月23日閲覧。

勝野頼彦（2004）『高大連携とは何か—高校教育から見た現状・課題・展望』、学事出版。

堀茂樹（2005）「多言語主義について—英語以外の外国語を（も）教える理由」、平高史也・古石篤子・山本純一（編）『外国語教育のリ・デザイン 慶應 SFC の現場から』、慶應義塾大学出版会、pp.9–27。

教育研修事業財団（2011）「新しい教育の流れとして拡大する高大連携についての調査研究〜高大連携の現状と課題、将来への提言〜【現状報告】」、http://aef.or.jp/_src/sc632/2010.0981408D8291E598A8Cg 81408CBB8FF395F18D90.pdf、2016年 8 月23日閲覧。

日本経済団体連合会（2013）「世界を舞台に活躍できる人づくりのために—グローバル人材の育成に向けたフォローアップ提言—」、http://www.keidanren.or.jp/policy/2013/059_honbun.pdf、2016年 8 月23日閲覧。

境一三・冶山純子・平高史也・丸田千花子・山下一夫・吉川龍生（2015）「慶應義塾の第二外国語教育における高校・大学の連携についての意識調査——フランス語・ドイツ語・スペイン語・中国語の調査結果の概要と分析」、『外国語一貫教育における複言語・複文化能力育成に関する研究（研究課題番号24242018）平成24年度〜平成26年度科学研究費補助金基盤研究（A）研究成果報告書』、pp.16–71。

高大接続システム改革会議（2016）『高大接続システム改革会議「最終報告」』、http://www.mext.go.jp/ component/b_menu/shingi/toushin/__icsFiles/afieldfile/2016/06/02/1369232_01_2.pdf、2016年 8 月20日閲覧。

注

[i] 一貫教育校とは、慶應義塾の中で初等教育・前期中等教育・後期中等教育を担当する学校を指す。

[ii] この問題については、境一三・冶山純子・平高史也・丸田千花子・山下一夫・吉川龍生（2015）を参照。

[iii] 多言語主義を主張している堀（2005）は、英語の汎用性を強調し、第二外国語不要論を唱える人々が、この「英語以外の外国語を学ぶコストパフォーマンスの悪さ」を論拠の一つとして挙げていると述べている。

[iv] 2000年代初頭における高大連携の議論については勝野頼彦（2004）を参照。

[v] なお高大連携については、経団連が2013年に発表した提言に基づき、「グローバル人材育成」のための「高大接続テスト」の導入という議論に進んでいるが、本論ではこの問題には立ち入らない。日本経済団体連合会（2013）、高大接続システム改革会議（2016）などを参照。

网络流行句式"一言不合就～"的意义扩张

嚴　　馥

一．前言

"一言不合"描写双方沟通时话不投机的情况。"一言"指〈一句话〉，"不合"表示〈合不来、不和〉之意。《中国語大辞典》将其列为词条，释义为〈話がうまく合わない〉，并举例如下。

(1)　他们俩说着说着，<u>一言不合</u>就吵了起来[1]。
(2)　<u>一言不合</u>斗起口来。《儿女・39》

[《中国語大辞典》（下），東京：角川書店，1994年，3666页]

"一言不合"的标注是［成］，但该词典的［成］除成语之外，还包括具有成语形式的词语，因此无法完全断定"一言不合"是否真的属于成语。笔者另外调查的成语词典及汉语词典[2]都没有收录。因此，"一言不合"这四个字现今只能看作词与词的组合，尚未归类为成语。

"一言不合"在句子里，除了像例(3)这样单独使用之外，如例(4)、例(5)所示，也常跟副词"就"或"便"一起搭配使用。而且，人们在不能达成共识时，往往会发生冲突，因此"一言不合"后面会出现带负面意义的词语，如例(3)的"成仇"、例(4)的"相打"和例(5)的"拳脚相加"。

(3)　我们现在是要好的，万一将来<u>一言不合</u>，翻转面来，何苦为好成仇，弄到一场没趣？
　　　［清・张春帆，《九尾龟》，上海：上海古籍出版社，1990年，18页][3]
(4)　海明威性喜打猎，富有所谓"男子汉气概"，在酒巴间中饮酒时，<u>一言不合，就要与</u>人相打。［董鼎山，〈酒癖无助灵感〉，《读书》，北京：生活・读书・新知三联书店，1981年，97页][4]
(5)　<u>一言不合</u>便拳脚相加，兵刃相见。［陆文夫，《人之窝》，上海：上海文艺出版社，1995年，261页][5]

最近在网络中经常出现像例(4)那种"一言不合"加上副词"就～"的搭配，只不过，后接的词语并不带任何负面意义，跟典型用法不同，如例(6)所示。句子里的搭配虽然没有语法上的错误，但"跳舞"很难跟言语上的争执扯上关系，大多数的情况下，都是人们带着欢愉的心情进

行的动作，因此这种搭配，在语义上存在着很大的矛盾。

（6）　为什么印度电影里<u>一言不合就跳舞</u>？[6]

语言是活的，并非永远固定不变，"一言不合"跟"就"的搭配因为某种契机产生意义与用法上的变化，而且在网络的世界里，类似例（6）的使用越来越频繁，甚至延伸到电视媒体里，引发了众多的关注。俨然成为一个网络流行新句式。

这些网络流行新句式虽然形成一股势不可当的风潮，但尚未完全固定，因此在教学时，教师势必会把例（6）这种牛头不对马嘴的句子判成病句，要求学习者进行修改。只不过，使用网络的汉语学习者与日俱增，脱离教科书进入真实的汉语世界后，必定会接触到网络上流行的新句式，因此对于其使用和意义的变化也有必要加以探讨。在看似随意发生的意义变化里，是否存在着某种规律？本文从语言学的角度，使用语义扩张的概念来分析网络流行句式"一言不合就～"，对其扩张意义以及扩张意义结构进行考察。

二．多义词的语义扩张

认知语言学认为多义词的每个意义之间绝非独立的，而是互有关联的。卢植（2006:167）指出"词的意思组成一个集合，而该集合的结构是一种放射状网络（racial network），这意味着这样的网络有一个中心点，不同词义从该中心点出发向不同方向辐射出去"。

位于网络中心点的词义被称为原型词义。原型词义是我们在想象这个词的意思时首先想到的那个词义（卢植，2006:186）。松本（2003:141–143）[7] 指出原型词义具备最基本、熟习程度及凸显程度最高等特征。此外，在用法上不带限制，或者相对来说限制较少。

关于在整个放射状网络中，其他的词义跟原型词义之间的连接问题，卢植（2006:167）提到"认知语言学使用四种不同的过程来集中分析和探讨这个普通范畴中的多个成分之间的关联：转喻、隐喻、特化和概括"。换句话说，非典型词义的出现是建立在这四种认知的基础上的。

综上所述，籾山（2002:101）[8] 认为考察多义词时的主要课题有三。（1）决定原型词义、（2）分析各项意义间的关系、（3）厘清语义的放射状网络结构。本文欲探讨的"一言不合就～"的句式是词与词的组合，其意义扩张就如同多义词语义扩张的放大版。因此，下面也逐步按照上述的三个项目来分析其在意义方面的扩张情况。

三．"一言不合就～"的原型意义　＜一句话说得不投合，立刻～＞

"一言不合就～"的原型意义就是以往通用的、收录在词典里的、未经过扩张变化前的意义。笔者认为其原型意义里包含五个意义成分：

〈沟通双方〉〈口头〉〈失败〉〈立刻〉〈负面动作/状态〉

"一言不合"表示〈一句话说得不投合〉[9]，为符合这个词义，首先需要两个沟通者，而且双方是用口头进行沟通。另外还包括双方沟通不顺利之后失败的部分。副词"就"表示〈立刻～〉，也就是双方在一言不合之后立即出现另一种情况。一般来说，沟通失败之后，都不会发生什么好事，因此"一言不合就～"后接的词语所表示的动作或状态会带着负面意义。上述曾提及，大多数的是像例（7）和例（8）那种表示语言或肢体冲突的词语，有时也会出现例（9）和例（10），仅表示某一方的负面情绪等的词语。

(7) 包不同爱和人争辩，却不问亲疏尊卑，一言不合，便争个没了没完。[金庸，《天龙八部》（四），北京：生活·读书·新知三联书店，1994年，1573页][10]

(8) 众人均知南海鳄神是段延庆的死党，但一言不合，便即取了他性命，凶残狠辣，当真是世所罕见，眼看到这般情状，无不惴惴。[金庸，《天龙八部》（五），北京：生活·读书·新知三联书店，1994年，1866页][11]

(9) 对天气与情绪，并没有必要盖棺论定。只是反观现下高温天，为什么一言不合就动怒？[《新民晚报》、2016年08月23日][12]

(10) 问题是聊天的对象不认识，又看不到对方脸部表情，下笔若稍不斟酌，一言不合，误会就发生了。[中央研究院平衡语料库][13]

三. "一言不合就～"的扩张意义

3.1 扩张意义（1）＜思想·观念不一致，立刻～＞

根据网友的说法[14]，"一言不合就～"起源于下面这两个句子。

(11) 一言不合就飙车[15]。

(12) 一言不合就开车[16]。

在这里的重要关键是"飙车"和"开车"。这两个词在网络世界里有了新的解释。百度曾因卖掉网吧的行为受到众多网友的抗议。网友发无意义的贴子、黄图、网盘等去爆吧，用这些行为来表达不满情绪。在爆吧时产生了"老司机（开车）带你上路"这样的说法。"老司机"本指富有经验、相当熟悉路况的驾驶者，"上路"指启程、动身出发到目的地。通过隐喻的手法，"老司机"被比喻为熟悉网络操作的吧友，"上路"成为开始运用各项手段达到爆吧的目的，而"飙车"或是"开车"则被用来表示各种让网吧停止营运的行为，比如在网上发一些种子、网盘等特殊资源等等。在"大家知道钢琴吧被百度卖掉了嘛"中，出现了下面的说法。

（13） 被卖了70万，空降了一枚吧主，现在吧友很不爽在飙车呢[17]。

（14） 以最近"钢琴吧"为例，百度以三十万价格卖给一个钢琴店的老板，钢琴吧的所有吧务在贴吧空降吧主后集体罢工，如今的钢琴吧整个乱套了，不少吧友通过"开车"的方式表示抗议[18]。

老司机	开车	带你	上路
老经验的驾驶者	驾驶车辆	乘客	动身前往目的地
↓	↓	↓	↓
老经验的吧友	发黄图、网盘等行为	经验浅的吧友	达到爆吧目的

[表1] "开车"的隐喻用法形成过程

另外还有一种说法，同样跟爆吧有关。那是起因于一场世界电子游戏竞技大赛。据说某次世界电子游戏竞技大赛的举办方的态度引起 DOTA2 观众不满，观众就跑去爆其百度贴吧，让百度贴吧自动封吧。表示爆吧行为的"开车"的新用法，被套用到"一言不合就～"句式里，慢慢扩张之后就形成现在的网络流行新句式。

这个扩张意义跟原型意义的用法相较后，可以发现〈口头〉及〈负面动作/状态〉这两处发生了变化。"一言不合"的原型意义的用法将沟通形式限定为口头沟通，但是到了扩张意义（1）时，不再局限在口头沟通上。在上述的事件发展中，很难完全排除亲自通过口头进行沟通的网民，然而大多数的吧友应该只通过观察对方的态度和行为就心生不满，并无直接交流。换句话说，在"一言不合就开车"的句子里，沟通双方仍然存在，但沟通方式有所改变，由口语上的交流变成观察。

再者，表示〈负面动作/状态〉的词语性质也发生变化。"开车"是以通过隐喻手法产生的新意义进入"一言不合就～"的句式里，这个新意义带有进行破坏的负面意义，与典型用法的"一言不合"要求的〈负面动作/状态〉契合，因此能顺利进入这个句式，不至于让人感觉突兀。然而，"开车"的原义是〈驾驶机动车〉，并不具备任何价值色彩，属于中性意义的词语。而这个中性意义在后来的意义扩张里被焦点化，因此，在"一言不合就～"的后面，除了原本要求的负面意义的词语之外，一些不带任何价值色彩的词语，甚至通过解释还能成为带有正面价值的词语也都出现了。

比如例（15）的"卖萌"。这个词原为动漫用语，从〈使动漫角色刻意装萌样讨好观众〉，引申出〈故作可爱状，藉以讨好他人〉或是〈故意作秀〉或〈无知〉等贬义。这个词语带有嘲讽或轻蔑的意思，算不上正面意义的词语。但是为了不让场面太难看，年轻女孩常用装可爱这种战术来缓解局面，从某种观点来看，这也可以算是一种处世智慧。跟上述的"开车"的破坏行为相比，"卖萌"的负面意义相对地淡化许多。也就是说，在经过意义扩张后，后接词语不再需要是表示

负面意义的动作或状态。此点在后续介绍的其他的扩张意义里更为明显。

(15)　道理讲不通的时候，卖萌是最好的武器。<u>一言不合就给你卖萌</u>。[19]

综上所述，扩张意义(1)转变成"沟通双方在思想·观念不一致时，立刻进行某种行为"。跟原型意义相比，在沟通方式和最后出现的词语性质上发生了变化。其意义成分如下：

〈沟通双方〉〈不限口头〉〈失败〉〈立刻〉〈不限负面动作/状态〉

3.2　扩张意义(2)＜总是＞

"一言不合"的"言"字，表示句式里存在〈沟通双方〉和〈口头〉上的沟通这两个重要环节。但到了扩张意义(2)里，原本重要的组合成分不再被焦点化。与此同时，诱发"一言不合"的后续动作的契机也改变了。本该是失败的沟通，现在变成某个特定的事态。换句话说，整个句式的意思，从"一碰到话说得不投合的情况，就立刻进行某行为"，转变为"一碰到某个特定情况，就立刻产生某行为"。扩张意义(2)的意义成分变更如下：

〈特定事态〉〈立刻〉〈不限负面动作/状态〉

(16)　现在的人啊，<u>一言不合就秀恩爱</u>，看看这才叫真正的浪漫，爱情的小船说开就开！[20]

"秀"为英语的 show 的音译，表示〈故作～状〉。在例(16)里的"秀恩爱"完全是"现在的人"的单方面行为，无所谓沟通不沟通，当然也不存在口头的沟通方式。既然沟通双方和口头的沟通都消失，自然也不会出现沟通时的不顺利，所以"秀恩爱"这个动作的契机不是沟通上的失败，而只是指一种特定的、经常发生的事态。每个不同的语境都有其不同的触点，根据网络里贴在例(16)下方影片内容可以得知，特定事态是跟另一半同时出现在公众场合的情况。换句话说，现代人习惯一到众人面前，就马上炫耀亲密关系、彰显美好的恋情或幸福的婚姻。扩张意义(2)的"一言不合就～"变成〈一有机会就～、动不动就～、总是〉之意。

此外，扩张意义(2)里的后续动作也不局限于带负面意义的词语。例(16)的"秀恩爱"还带点嘲讽的意味，但例(17)的"发自拍"，就属于中性词语，完全不带任何正面或负面的价值。

(17)　后来一个中国人找我聊天，忽然间有找到同胞的感觉，就加 QQ 聊了起来。……。
　　　　他去过别的国家留学，见识很广，然后认识很多朋友，说真的，在他身上学到很多东
　　　　西，一切都挺好的，本来是值得深交的朋友。结果剧情来个一百八十度大转变，……

　　　他来了句很自然的话，发自拍给我看看呗。我瞬间傻眼了，发自拍干嘛，又不和你相
　　　亲，我长什么样和我们聊天有什么关系吗？　还是说你和一个人交朋友看的是她的外
　　　貌？　……。所以，不要一言不合就要我发自拍，好吗？[21]

　　文章的作者写到在网上与新朋友结识的经过。在这段记述里，存在着两个人进行口头上的沟
通，但纯属情感交流，并非起争执，亦非思想上的对立，自然也没有因沟通不良而失败的局面。
在典型的"一言不合就～"句式的意义成分里存在的〈沟通双方〉、〈不限口头〉、〈失败〉均不复
见。

　　根据前后文可以得知，这位在网上认识的朋友，跟人聊得尽兴之际，就习惯要求看对方的
长相。换句话说，只要一触动了"网聊"这个特定的事态，立刻就会启动"要求发自拍"的行为。
发自拍本身没有什么负面的意思，至此，不带任何价值色彩的词语也能顺利进入"一言不合就～"
的句式里。文章最后之所以产生厌恶或抱怨的语感，完全是因为作者对网友的态度不以为然，跟
词语本身性质无关。

（18）　很多人发自拍的时候，网友们会纷纷地吐槽说，有事好说干嘛发自拍。所以才有一句
　　　话叫做一言不合就发自拍。有些星座的确是这样的，动不动就喜欢发自拍，在群聊里
　　　发自拍，在朋友圈里发自拍，在微博上发自拍，一有机会就发自拍，他们是哪些星座
　　　呢？[22]

　　在例（18）的星座解说里的"一言不合就发自拍"后面，作者还加上"动不动就～"和"一有
机会就～"两个替代词语来解释前述的"一言不合就～"的情况。从此例更可清楚得知，扩张意
义（2）的"一言不合就～"的语义已转变为"总是在某特定事态发生时，立刻产生某种行为"。

3.3　扩张意义（3）＜冷不防地、突然＞

　　到了扩张意义（3），"一言不合就～"的原型意义里的前三个组合成分消失殆尽，与此同时，
〈立刻〉这个部分被大大地焦点化。前者的扩张意义（2）尚保留了诱发后续动作的契机，但在扩
张意义（3）里，做任何动作时的理由已完全不列入考虑。整个句式的意思也因此转变为"没有任
何铺垫就立刻进行某行动"。扩张意义（3）的意义成分如下：

〈立刻〉〈不限负面行为/状态〉

（19）　一言不合就裁员！　科技公司让你走人的 N 种理由[23]。

例（19）是文章标题，文章的内容描述大公司会因为任何看似不可抗力的理由来解雇员工，而且是在无预警的状况下。雇主和雇员之间没有任何沟通，"裁员"也无关乎沟通的成功或失败。整个句式所要传达的只有突然被开除的那种震惊。

此外，在这个"一言不合就～"句式里，后面不仅能出现带有负面意义的词语，也能出现如例（20）里的"唱歌"和"跳舞"，不带任何价值色彩的中性意义词语。

（20）　提到印度电影，不少人想到的都是：一言不合就唱歌跳舞，歌舞桥段来得让人没有一点点防备[24]。

对印度电影略知一二的人都晓得，里面总是少不了唱歌、跳舞的场景。增添歌舞片段的独特表现手法，其实跟印度电影的发展有很密切的关系。在早期的印度，看电影是种奢侈的享受。影片时间过短的话，观众会觉得看得不过瘾。制片者为顺应观众要求，想方设法拉长电影时间，便在情节里加入唱歌和跳舞的场景。有些网友为了调侃那种突如其来的歌舞表现手法，便用"一言不合就～"这个网络流行句式，来描述印度电影里男女主角只要说上几段话，就没铺没垫地开始载歌载舞的情况。在这个句子里完全不存在任何跟沟通有关的情况，所要强调的只有突然开始唱歌跳舞，那种令人不知所措的感觉。

（21）　一言不合就结婚，你真的想好了吗[25]？

"一言不合就～"的后面，甚至还能与正面词语搭配使用。"结婚"象征着男女双方的美好结合，跟原型意义里所提及的"打骂、生气、发生误会"等情况处于两个极端。本文根据这种意义上的相对性，把"结婚"视为带正面意义的词语。由此可知，"一言不合就～"到了扩张意义（3）时，已经能完全违反典型意义的要求，与正面词语一同搭配使用。

四．扩张意义的效果

"一言不合"的扩张意义从"一言不合就开车"或"一言不合就飙车"开始发端。之后，网友可能觉得用起来有意思或者想追随潮流，便大量复制此句式，使其成为新的网络流行句式。主要目的是为了搞笑、调侃，或表示幽默。这些效果能否得到彰显，跟后接的词语性质有极大的关联。由前述内容可知，伴随着意义的扩张，"一言不合就～"后面出现的词语性质，从原型意义要求的负面词语，扩张到不带价值色彩的词语，甚至能出现带有正面意义的词语。而进入此句式后面的词语，越带正面意义的，所形成的反差就越大，网友想要制造出来的那些无厘头的感觉也就越明显。

例（22）和例（23）的"一言不合就～"都不属于原型意义的用法。前者出现"暴雨"、"暴晒"、

"台风"等表示天灾的词语，天灾人祸绝非美事，本身带有负面意义。相较而言，例（23）的"上演男模秀"本身虽不带任何评价色彩，但是能看帅哥是件令人大呼过瘾、感到欢愉的事情，在意义上相对地比较正面。因此，同样套用到"一言不合就～"的句式里，后者的违和感比较显著，在语言表现上所呈现出幽默感也相对比较大。

(22) 一言不合就暴雨，一言不合就暴晒，一言不合就台风，一言不合就上天！ 夏至已过，天气就像娃娃脸。一言不合说变就变，高温暴雨随意切换[26]。

(23) 然而，某些情况下一言不合带来的却是满满的福利，比如正在火热进行中的欧洲杯，一言不合就上演男模秀！[27]

五． 小结

本文从语言学的角度，使用语义扩张的概念考察了网络流行句式"一言不合就～"的扩张意义及其扩张意义结构。其原型意义为＜一句话说得不投合，立刻～＞。从这个原型意义衍生出以下三个扩张意义。

扩张意义（1）＜思想·观念不一致，立刻～＞。

扩张意义（2）＜总是＞

扩张意义（3）＜冷不防地、突然＞

表示原型意义的"一言不合就～"里有五个意义成分：〈沟通双方〉〈口头〉〈失败〉〈立刻〉〈负面动作/状态〉。扩张意义（1）的意义成分转变为〈沟通双方〉〈不限口头〉〈失败〉〈立刻〉〈不限负面动作/状态〉。其与原型意义有两点不同之处。第一，沟通方式从直接的口头交流扩张到间接观察。第二，后面出现的词语性质的限制减少。"开车"或"飙车"通过隐喻产生负面意义后，顺利进入"一言不合就～"的句式里，成为最早的例子。但是"开车"的原意不带负面意义，这个部分被焦点化之后，不带负面意义的词语也能进入"一言不合就～"句式里。这点在其他的扩张意义里也能观察得到。扩张意义（2）的意义成分为〈特定事态〉〈立刻〉〈不限负面动作/状态〉。跟"一言不合"有关的〈沟通双方〉和〈口头〉不再是焦点，〈失败〉也通过隐喻转变成〈特定事态〉。"一言不合就～"本来表示一碰到话说得不投合的情况，就立刻进行某种行为，到了这个扩张意义里，变成一碰到某种特定情况，就立刻进行某行为。扩张意义（3）的意义成分只剩下〈立刻〉〈不限负面行为/状态〉。跟"一言不合"相关的意义成分消失殆尽，取而代之的是〈立刻〉的部分被完全焦点化。

"一言不合就～"的所有意义组成的放射状网络结构，是根据整体和部分的关系架构而成的，也就是说通过焦点的转移，由整个事态逐渐聚焦到中间由副词"就"所表示的部分。〈立刻〉这个意义成分在三个扩张意义里都没发生变化，可视为网络流行句式的核心。以此意义成分为中心点，来观察扩张意义间的关系可以得知，前面表示"一言不合"的词义逐渐淡化，意义成分脱离

焦点成为背景。从"口头沟通"到"非口头沟通",扩张到"无关乎沟通的特定事态",最后完全消失聚焦至〈立刻〉这一点上。与此同时,随着"一言不合"意义成分的淡化,后面能出现的词语性质的限制也逐渐宽松,从"负面意义词语"最后扩大到所有"不带负面词语的词"。此外,与原型意义的意义成分的差距越大,其呈现出来的语言效果也就越显著。

六. 今后课题 ―"一言不合"就消失?

网络流行句式因某种契机突然出现,也有可能突然消失,第一个课题便是观察"一言不合就～"的保质期究竟有多长,是否还会出现新转变。

第二个课题是厘清扩张意义间的模糊界线。"一言不合就～"可以根据上下文随时变换成其他意思。前述的"一言不合就秀恩爱"解释成一天到晚在别人面前秀恩爱,但在"张睿大方认爱!一言不合就秀恩爱,猝不及防就扔狗粮![28]"里,"猝不及防"这四个字让句式转变成〈突然宣布谈恋爱〉之意。句式和后接词语与语境的关系必须再加以厘清。

第三个课题是对负面词语或正面词语,定下更明确的定义。本文使用意义上的相对概念把"结婚"视为正面词,但如何区分负面词语、中性词语和正面词语应该还是需要更明确的基准。

附注

1 汉语例句后面原附带日译。例（1）彼らは話をしていて、うまが合わず、けんかになった。例（2）言葉の行き違いで口論しだした。

2 笔者调查过下列词典，发现均未收录"一言不合"。

［a］《辞海》，辞海编辑委员会编纂，上海：上海辞书出版社，2009。

［b］《成语词典》，常州市教育局编，南京：江苏人民出版社，1981。

［c］《古今汉语成语词典》，《古今汉语成语词典》编写组编，太原：山西人民出版社，1985。

［d］《汉语成语词典》，甘肃师范大学中文系《汉语成语词典》编写组编，上海：上海教育出版社，1978-1983。

［e］《汉语成语词典》，李一华；吕德申编，成都：四川辞书出版社，1985。

［f］《汉语成语大词典》，湖北大学语言研究室编纂，郑州：河南人民出版社，1985。

［g］《汉语成语分类词典》，叶子雄主编，上海：复旦大学出版社，1987。

［h］《汉语新成语词典》，史式；赵培玉编著，西安：陕西人民教育出版社，1986。

［i］《实用成语词典》，常晓帆编，北京：台声出版社，1990。

［j］《现代汉语词典》（第六版），中国社会科学院语言研究所词典编辑室编，北京：商务印书馆，2012。

［k］《中華成語大辭典》，向光忠；李行建；刘松筠主编，长春：吉林文史出版社，1986。

3 例（3）出处：张春帆，《九尾龟》，上海：上海古籍出版社，1990。

4 例（4）出处：董鼎山，〈酒癖无助灵感〉，《读书》，北京：生活·读书·新知三联书店，1981。

5 例（5）出处：陆文夫，《人之窝》，上海：上海文艺出版社，1995。

6 例（6）出处：https://zhuanlan.zhihu.com/p/21596004

7 日语原文：松本（2003:141）「複数の意味のなかで、最も基本的であり、慣習化の程度・認知的際だち（cognitive prominence/cognitive salience）が高いといった特徴を備えたものをプロトタイプ的意味と認定することにする。」（同書：143）「多義語の複数の意味のうち、用法上制約がない、あるいは制約が相対的に少ない意味をプロトタイプ的意味と認定し、（略）。」

8 日语原文：籾山（2002:101）「多義語分析の課題としてすくなくとも以下の３つがある。① プロトタイプ的意味の認定　② 複数の意味の相互関係の明示　③ 複数の意味すべてを統括するモデル・枠組みの解明。」

9 此词义出自汉词网（http://www.hydcd.com/cy/htm5/yy1851.htm）

10 例（7）出处：金庸，《天龙八部》（四），北京：生活·读书·新知三联书店，1994年。

11 例（8）出处：金庸，《天龙八部》（五），北京：生活·读书·新知三联书店，1994年。

12 例（9）出处：张嘉佳. 情绪中暑. 新民晚报. 2016/08/23.

（http://shanghai.xinmin.cn/xmwb/2016/08/23/30357792.html）

13 例（10）出自中央研究院平衡语料库（http://asbc.iis.sinica.edu.tw）。原文为繁体字，经笔者修改。

14 百度知道（https://zhidao.baidu.com）为公众媒体，任何网友都随时可以按照己意对内容进行增删。在可信度方面，的确有值得商榷之处。本文先根据"百度知道"上网友的说法，对扩张意义（1）进行意义成分的分析之后，再与其他扩张意义的意义成分进行比较。相对来说，扩张意义（1）跟原型意义的意义成分的相似度最高，在认知上的变化最小。根据这些实际语言现象，最终将扩张意义（1）视为最早出现的扩张意义。

15　例（11）出处：

　　http://zhidao.baidu.com/link?url=kmfsXA7DI8BTg9GF9d5OOtYz65xQQ0M-c7aJfOVrV2uG3piGv8X
NHjrT1LD7GZwuQ9NmXvhK88gXg7dDtnngyOmhv7bMYkwWAprpPfZXlS

16　例（12）出处：

　　http://zhidao.baidu.com/question/1642054281732843980.html?qbl=relate_question_1&word=%D2%B
B%D1%D4%B2%BB%BA%CF%BE%CD

17　例（13）出处：http://tieba.baidu.com/p/4103200894

18　例（14）出处：https://www.zhihu.com/question/36612887

19　例（15）出处：http://www.gd.xinhuanet.com/newscenter/2016-07/12/c_1119204347.htm

20　例（16）出处：http://www.weibo.com/3490279655/DxqAFzYaX?type=comment

21　例（17）出处：http://www.jianshu.com/p/8045fccbd18c

22　例（18）出处：http://3g.d1xz.net/astro/gexing/art145353.aspx。笔者将原"纷纷的"修改为"纷纷地"。

23　例（19）出处：http://business.sohu.com/20160807/n462985365.shtml

24　例（20）的原文为繁体字，经笔者修改。例句出处：

　　http://www.weibo.com/ttarticle/p/show?id=2309404000797389127598

25　例（21）出处：http://fashion.sohu.com/20160701/n457275249.shtml

26　例（22）出处：http://bj.leju.com/scan/2016-07-04/15236155631982564127711.shtml

27　例（23）出处：http://www.gd.xinhuanet.com/newscenter/2016-07/12/c_1119204347.html

28　例句出处：http://toutiao.com/i6301900495153791489

参考文献

［1］卢植，《认知与语言—认知语言学引论》，上海：上海外语教育出版社，2006。

［2］松本曜，《认知意味论》，东京：大修馆书店，2003。

［3］籾山洋介，《认知意味论のしくみ》，东京：研究社，2002。

［4］大东文化大学中国语大辞典编纂室编，《中国语大辞典（下）》，东京：角川书店，1994。

ドイツの就学前教育における言語支援策

牛　山　さおり

１．　はじめに

　2000年の経済協力開発機構の学習到達度調査（PISA2000）により、ドイツは家庭の社会階層、両親の出身国によって、生徒の学習到達の格差が最も大きい国であることが示された（小玉2008:76-84）。この結果を受け、2001年12月以降、就学前教育における言語支援、科学教育などに比重が置かれるようになった。

　政策が施行されてから15年後の現在、ドイツの就学前教育において、近年何が問題となっているのかを述べ、言語支援策の例として Baden-Württemberg 州（バーデン・ヴュルテンベルク州、以下 BW 州）での取り組みを報告する。そして、外国につながりを持つ就学前児童への日本語教育にも応用可能だと考えられる点を考察する。

２．　ドイツの就学前教育
２.１　旧東西ドイツ地域に見られる相違点

　豊田（2011）によれば、統一前の旧東西ドイツ地域では、就学前教育全般に関する考え方、幼児を取り巻く環境に地域差が見られた。旧西ドイツ地域では市場経済・自由主義体制のもと、保育や教育の基盤は家庭にあり女性が担うものという考え方が根強く、３歳未満児を対象とする保育施設の整備は著しく遅れていた。一方、旧東ドイツ地域では、社会主義的計画経済の政策のもと、乳幼児からの集団保育および就学前教育は、質・量ともに充実していた。統一を機に、旧西ドイツ地域では旧体制を維持し、旧東ドイツ地域では旧西ドイツのモデルに再編成されたが、その過程は決して容易な道のりではなかったと言われている。

２.２　保育施設の種類

　ドイツの保育制度は児童青少年援助法のもと、16州のうち10州で青少年・家族・女性・福祉関係省に、６州は教育を管轄する官庁に所轄されている。連邦政府と連携のもと、州ごとに教育・研究を統括する大臣が存在し、文化大臣会議により、保育制度や教育行政が決定される（斉藤2014）。一般的に、就学前の子供たちが受ける保育サービスには、児童昼間保育[i]と昼間

施設との２種類が存在し、本研究に関する昼間通所施設を以下に挙げる。

・保育園（Kinderkrippe）：０歳から３歳未満児を対象とした施設
・幼稚園（Kindergarten）：３歳から就学前の幼児を対象とした施設
・児童昼間施設（Kindertagesstätte、通称 KITA）：幼保一元化された施設
・学童（Kinderhort）：基礎学校１年生から14歳未満児童までを対象とした施設[ii]

２.３　保育をめぐる諸問題

　2008年12月に子供援助法が施行されたことで、３歳未満児に向けた保育サービスの利用率は向上し、半日保育から全日保育への移行が行なわれつつある。2013年８月には、３歳未満児を対象とした「保育請求権」が認められ、３歳未満児のための保育施設の増設が急速に進められた。なお、この数年で議論されている問題としては、以下の３点が挙げられる。

（1）　施設の増設に伴う保育者不足

　2014年７月16日付の Die Welt 誌によれば、連邦全体でさらに12万人の保育者が必要という試算が出ている。保育者一人あたりの担当する３歳未満児の理想的な割合は１:３であるのに対し、旧西ドイツ地域では平均3.8人、旧東ドイツ地域では平均6.3人であった。３歳児以上の場合にも同様の傾向が見られ、保育者１人あたり理想とされる幼児の人数基準を満たしているのは、ブレーメン市と BW 州のみであった。

（2）　移民背景を持つ子供の就園率増加

　従来、外国籍を持つ子供の就学前教育・保育への参加率は、ドイツ国籍の子供より低く、就学前の段階から教育格差が始まっているとの指摘もあったが（小玉2008:78-79）、近年、自宅で保育されていた移民背景をもつ子供も、就学前に施設に通う傾向にある（VBW 2016）。

表１：家庭外保育施設で保育を受けている、あるいは昼間保育サービスを受けている子供の割合

	満１歳から３歳			満３歳から６歳		
	全　体	移民背景を持つ幼児[iii]	移民背景を持たない幼児	全　体	移民背景を持つ幼児	移民背景を持たない幼児
2009年	20％	11％	25％	92％	84％	96％
2014年	32％	20％	38％	94％	85％	98％

　３歳未満児で保育サービスを受ける子供の割合の目標は35％と設定されており、全体での結果は目標値に近づいている。移民背景を持つ３歳未満児で、保育サービスを受ける割合がほぼ２倍に増えているのは、特筆すべき傾向である。なお、３歳から６歳児に関しては、就学前に何らかの保育施設に通う、あるいはサービスを受けるということが定着している。次に、移民背景を持つ子供のうち、ドイツ語が家庭内言語ではない子供の割合を示す。

表２：家庭外保育施設に在籍している乳幼児および児童（０歳〜14歳）の中で、ドイツ語が家庭内言語でない子供の割合

	在籍児童のうち ０〜25％	在籍児童のうち 25〜50％	在籍児童のうち 50〜75％	在籍児童のうち 75〜100％
2009年	33.7％	36.5％	22.4％	8.4％
2014年	30.4％	36.0％	23.2％	10.5％

　表２では、表１よりも対象とする子供の年齢が拡大されているが、在籍児童のうちドイツ語を家庭内で使用しない子供の割合が50％を超える保育施設は、数パーセントではあるが上昇している。調査では地域差や年齢別の傾向までは把握できておらず、検討の余地がある。

　連邦全体で、ドイツ語を家庭でほとんど使っていない子供が大半を占める保育施設が10.5％存在するいうことは、保育施設で言語支援策を行なう必要性が高いだけでなく、日常の場面でドイツ語のみを使用すること自体が困難である保育施設も増加していることが示唆される。

（3）　保育者養成課程の見直し

　ドイツの保育者は社会的教育者（Sozialpädagoge）という位置づけで、小学校段階以上の教員養成とは異なる養成システムで教育される（吉岡2007）[iv]。そのため、最終学歴が専門学校卒である保育者が大半を占めており、BW 州では大学で社会教育学を修めた保育者が全体の14％に過ぎない（同州文化省の発表による）。Vitzthum（2014）によれば、在学中に行なわれる実習の比重が高く、離職率は極めて低いが、専門学校在学中の中退率が60％に達する学校もある。そのため、1）保育者の給与水準の改善[v]、2）専門学校で行なわれている教育の質の向上[vi]、3）新卒保育者の採用基準の一定化、が求められている（Mix 2013）。今後、一層多様化する幼児（とその家族）を理解し、就学前教育を行なっていくために、専門性を持った保育者、および常に新しい保育事情を学び続けることのできる姿勢を持った保育者の養成が重要とみられる。

３．就学前における言語支援策 ―BW 州を例に―

Statistisches Bundesamt（2016）によれば、BW 州の人口の27％（約295万人）が移民背景を持っている。そして就学前教育、保育施設とサービスの量・質を調べた調査で、BW 州は一位を獲得した[vii]。この背景にあるのは、BW 州の文化・青年・スポーツ省が保育政策を積極的に行なうなど、保育の質の向上への努力である。保育者向けの研修・再研修を充実させ、離職率も連邦で最も低い水準を保っている。そして以前は学士号を持つ保育者が少なかったが、同州の政策により、学士号を持つ保育者は全体の25％に上昇した。以下、実際に行なわれてきた言語支援策を紹介する。

３.１　Sag'mal was ― Sprachförderung für Vorschulkinder
　　　（ねえ、何か言って ― 就学前の子供のための言語支援策）

このプログラムは2003年から2010年にかけて、ドイツ語を第二言語として獲得する３歳未満児を担当する保育者に向けて開発された。1）Mit Kindern im Gespräch（子供との会話）、2）Dialoge mit Kindern führen（子供との対話を導く）、3）Sprache macht Stark!（言葉が強くする！）という３つの下位プロジェクトから成る。

このような政策が立ち上がった背景には、以前より保育者養成校のカリキュラムに、言語獲得・習得などを学ぶ科目が極めて少ないと指摘されていたことが挙げられる。この点を補うために、保育者が、1）自身の言語習得経験・学習経験に気づき、振り返る、2）理論を学び、ドイツ語を母語としない幼児の言語発達を適切に見極め、支援を行えるようになる、3）言語獲得に関する知識を広める伝達者となる、ことを目的として行なわれた。

３.２　Sprache macht Stark! Sprachbrücke Familie-Kita
　　　（言葉が強くする！ 家庭と保育施設の間に架かる言葉の橋）

このプログラムは2006年から2009年にかけて、マンハイム大学とマンハイム市及びルートヴィヒスハーフェン市の保育施設、BASF 社の協力のもと実施された。報告書、Tracy（2009）を基にその詳細を述べ、実際に行なわれた活動の例を付録に挙げる。

・対象とする幼児：２歳児から４歳児。家庭でドイツ語以外の言語のみに触れていた幼児、あ
　　　　　　　　　るいは教育から離れた環境で育ったため、言語能力が年齢相当と見なされ
　　　　　　　　　ないドイツ語を母語とする幼児。
・期間：１つのテーマにつき２週間を１クールとして、１年間行なう。
・対象とする保育施設：2006年は５ヶ所から開始、2009年までに39ヶ所で実施された。

・目的　1）保育施設に入る前の幼児に、基礎的なドイツ語能力を身に着けさせる。

　　　　2）言語支援策を理解した保育者を一園に2名ずつ配置するための研修も兼ねる。

・活動の内容：3つの活動を連続して行なう。

　　a）ミニグループ活動：幼児4人1組でグループを組み、週3回1時間ずつ行なう。1テーマにつき、15〜20の語彙や表現を重点的に扱う。保育者は擬音語や幼児語を使わず、幼児がドイツ語を使う楽しさに気づくよう促す。

　　b）親子グループ活動：週1回1時間半〜2時間程度行なう。自宅で、第一言語を使っても出来る活動を多く行なう。保育者は幼児が多文化・多言語で育つことの意味を考え、ドイツ語を母語としない親が、今後どのように子供のドイツ語習得を支援できるかを助言する。

　　c）日常での支援：a）とb）で学習しているテーマで扱われたテーマと語彙を他の保育者と共有し、なるべく日常の中で幼児が語彙や表現に自然に触れられる機会を増やす。

3.3　Sprache macht Spaß!（言葉は楽しい！）

　本プログラムは、3.2で得られた知見を活かし、勤務先で言語支援プログラムを遂行する保育者の養成を目的として、2012年から2015年まで行われた[viii]。言語支援に取り組む保育者に必要となるテーマ別のモジュール制で行なわれ、理論と振り返りの合計40コマからなる。本項では、この中から Modul 3、5、9 を取り上げる。

Modul 3：言語モデルとしての保育者

学習目的：「言語モデルとしての保育者」とは何かを考える。幼児が第二言語で話すことの楽しさを知るために、保育者が何をすればよいのかという問題に取り組む。

理　　論：言語獲得理論、学習理論などを学ぶ。同時に子供の発話を「聞くこと」、子供と大人の相互作用についても考えを深める。

実　　践：声の出し方、言語を発する時の身体の使い方を学び、唇や舌などを重点的に動かす練習を通して、幼児が言語を獲得するプロセスを体験する。

Modul 5：子供の言語発達を観察する：ポートフォリオ入門

学習目的：幼児の言語発達を観察し、報告するためのポートフォリオを作る。

理　　論：各幼児の言語発達の観察方法、およびポートフォリオの作り方を学ぶ。

実　　践：実際に幼児1人の言語発達を、3週間にわたって観察したのち、ポートフォリオを作成し、発表する。

図1：ポートフォリオの例

Das bin ich（これがわたし）
_____

Name（名前）   Datum（日付）

	Ich heiße（名前）
	So alt bin ich gerade（年齢）
So sehe ich aus （写真を貼る）	So groß bin ich gerade（背の高さ）
	So viel wiege ich（体重）
	Meine Haarfarbe（髪の色）
	Meine Augenfarbe（目の色）

Das möchte ich über mich erzählen（こんなことを話したい）
_____
_____
_____
_____

Diese Wörter kann ich schon（使える言葉）

Datum（日付）	Wort（語彙）	Wortbedeutung（語の意味）
_____	_____	_____
_____	_____	_____
_____	_____	_____

Modul 9：保育の観点から見た早期の多言語

学習目的：第二言語獲得（習得）において、第一言語が果たす役割を知る。他の言語や方言に
　　　　　対する保育者自身の態度を振り返る。

理　　論：第二言語獲得・習得研究で得られた知見を知り、外国語を学ぶ態度について改めて
　　　　　考える。第一言語・第二言語・母語の違いを知る。自分に合った言語学習の方法が
　　　　　何かを振り返る。さらに言語バイオグラフィの作り方を学ぶ。

実　　践：保育者自身の言語バイオグラフィを作成する。さらに幼児用のバイオグラフィには
　　　　　何が必要となるかを参加者と話し、アイディアを交換・共有する。

図２：自分の言語態度を振り返るためのシート

Meine Sprachbiografie （言語バイオグラフィ）	Erst- oder Muttersprache(n)（第一言語もしくは母語）： Zweitsprache(n)（第二言語）： Fremdsprache(n)（外国語）： Dialekt(e)（方言）： Besonderheiten（特記事項）：
Mein Sprachverhalten mit Fremdsprache(n), Dialekt(e) （外国語や方言の言語態度）	z.B: Gestik, Mimik, Sprechtempo, Sprechmelodie, Grammatik （ジェスチャー、ミミック、テンポ、メロディー、文法など）

４．考察

　本稿ではドイツを例に就学前教育における諸問題と、言語支援策を検討したが、ここでは東西ドイツでの「幼児教育」「保育」に関する考え方、保育施設の違いから、統一後の制度一本化の難しさが浮き彫りとなった。そして、保育の責任は私的から公的な領域へ移行されつつあり、家庭外で教育を受ける子供の中には、ドイツ語を家庭内言語としない子供も多く含まれるようになった。彼らも就学前の段階で、基礎的なドイツ語力も身に着ける必要がある。日々刻々と変化していく状況に対して、専門家や州の文化省などが連携し、保育者に向けた再研修や資格取得の機会を提供することで保育者自身のスキルを向上させてゆくべきである。

　日本の場合には、すでに高卒以上の学生を対象とした保育者養成校や大学で、発達心理学と教育心理学の授業が行われ、日本語を母語とする幼児の認知発達に関する知識が教授されている。本研究で得られた知見を基に、今後、外国にルーツをもつ子供の日本語支援を行なう保育者に必要な点を提案することができるだろう。具体的には、保育者自身の言語学習経験や態度を振り返ること、幼少期から複数の言語を身に着けていくことに関する知識を得て、多様な文化背景を持った子供とその家庭を理解できるだけの開かれた態度を身に着けていくことである。

　今回は、ドイツの一州の取り組みを紹介するにとどまったが、今後の課題としては、日本国内における就学前の子供たちに向けた日本語支援の問題がある。いくつかの市町村の言語支援策を調べたところ、日本語を体系的に学ぶということに加えて、「育児相談」「託児」といった側面もやや強いように見受けられた。識字教育も含め、外国にルーツを持つ就学前の子供が日本語をどのように学ぶのか、どの程度最新の研究成果などが活かされたカリキュラムが行われているのかなど、引き続き検討していきたい。

注

i 　児童昼間保育は、日本でいう保育ママ・パパサービスを指し、保育者と両親が私的な契約を結び、委託するという形で行なわれる。

ii 　ドイツの場合、初等教育は基礎学校において４年間（一部の州は６年間）行われる。中等教育は生徒の能力と適性に応じて、基幹学校（卒業後に就職して職業訓練を受ける者が主に進学する５年制学校）、実科学校（卒業後に職業教育学校に進む者や中級の職につく者が進む６年制学校）、ギムナジウム（大学進学希望者 が主として進む９年制学校）が設けられている。

iii 　移民背景を持つ幼児＝両親の少なくとも片親がドイツ以外の出身である、という定義に基づく。

iv 　保育者を目指す学生の多くは、基幹学校ないしは実科学校を卒業後、２〜３年制の専門学校、あるいは５年制の保育者養成を目的とした、高度職業専門学校に進学する。専門学校では週３〜５日の実習が半年間、高度職業専門学校では後半の２年間が実習に充てられる。

v 　就学前教育に従事する保育者の給与は、一般的な小学校教員の50〜60％である。

vi 　教育内容に関しては州により偏りがみられる。教授する立場にある保育者の現場での経験に基づく授業が行われがちで、必ずしも時代の流れやニーズに即していないことがあると指摘されている。

vii 　ベルテルス財団の発表による（2015年９月17日付）

viii 　なお、2016年現在も改良されたプログラムが継続中である。

参考文献

小玉亮子（2008）.「PISA ショックによる保育の学校化「境界線」を越える試み」泉千勢、一見真理子、汐見稔幸（編）『世界の幼児教育・保育改革と学力』明石書店，pp.69-88.

斉藤純子（2014）.「ドイツの保育制度―拡充の歩みと展望―」『レファレンス』2014年２月号，pp.29-62.

豊田和子（2011）.「ドイツの幼稚園における「教育の質」をめぐる議論と成果―Tietze ら（ベルリン自由大学研究グループ）を中心に」『保育学研究』48号３巻，pp.269-280.

吉岡真佐樹（2007）.「教育福祉専門職の養成と教育学教育―ドイツにおける教育福祉専門職養成制度の発展と現状―」『教育学研究』第74巻，第２号，pp.88-101.

Baden-Württemberg Stiftung（2015）. *Sprache macht Spaß. Curriculare Grundlage: Spracherwerb und Sprachentwicklung von Kindern unter drei Jahren in der Kindertagespflege.*
https://www.bwstiftung.de/uploads/tx_news/150717-tm-sprache-macht_spass-download.pdf
（accessed 2016-08-01）

Baden-Württemberg Stiftung（2016）. *Sag mal was. Das Programm.*
http://www.sagmalwas-bw.de/startseite/（accessed 2016-08-01）

Landesstiftung Baden-Württemberg（2006）. Sag' mal was – Sprachförderung für Vorschulkinder – Eine Projektdarstellung –

Mix, Ulrike（2013-07-27）. Mängel bei der Erzieher-Ausbildung – Studie bescheinigt der Lehre eine schlechte Qualität- *Die Deutschlandfunk.*
http://www.deutschlandfunk.de/maengel-bei-der-erzieher-ausbildung.1180.de.html?dram:article_id=255370（accessed 2016-08-05）

Pressemeldung Ministerium für Kultus, Jugend und Sport Baden-Württemberg（2015）. Qualität in der frühkindlichen Bildung ist für Baden-Württemberg ein Top-Thema.*bildungsklick.*
https://bildungsklick.de/fruehe-bildung/detail/qualitaet-in-derfruehkindlichen-bildung-ist-fuer-baden-wuerttemberg-ein-top-thema/（accessed 2016-08-05）

Statistisches Bundesamt（2016）. *Destatis.* https://www.destatis.de/DE/Startseite.html
（accessed 2016-08-11）

Tracy, Rosemarie（2009）. *Sprache macht stark! Offensive Bildung.* Cornelsen Verlag.

Vereinigung der Bayerischen Wirtschaft e.V.（2016）. *Integration durch Bildung. Migranten und Flüchtlinge in Deutschland.* Waxmann.

Vitzthum, Thomas（2014-07-16）. In deutschen Kitas fehlen 120.000 Erzieherinnen.*Die Welt.*
https://www.welt.de/politik/deutschland/article130538422/In-deutschen-Kitas-fehlen-120-000-Erzieherinnen.html（accessed 2016-08-05）

付録：Sprache macht stark! プログラム例

テーマ：わたしの身体（Mein Körper）

形　態	時　間	内　　　　容
1週目（基本的な身体部位の語彙を学ぶ）		
a）ミニグループ活動	1時間目	厚紙で人形を作る。
		身体の部位に関する歌を歌う。
	2時間目	身体の部位に関する歌を歌う。
		手や足が冷たい、温かいといった表現を学ぶ。
		2人1組で毛布に隠れ、1人ずつ身体のどこかを外に出す。他の2人が中に入っている子供のどちらの手足かを当てる。
	3時間目	身体の全体像を描く。
b）親子グループ活動		親子で Peter Hampelmann（手足が動く人形）を作る。
c）日常での支援		ミニグループでの活動をポスターにまとめる。 幼児が親と作成した「手足が動く人形」を使い、全員で語彙を復習する。
2週目（身体部位の語彙を拡充する）		
a）ミニグループ活動	1時間目	身体の部位に関する歌を歌う。
		顔を扱って、部位の名前を覚える。
		1週目の終わりに親子で作った人形を使って復習する。
		新しく学習した身体部位の名前を入れて、歌を歌う。
	2時間目	身体の部位に関する歌を歌う。
		病気がテーマになっている絵本を読み聞かせる。
		鼻水が出ている自分の顔を描いて、表現を学ぶ。
	3時間目	身体の部位に関する歌を歌う。
		手に絵の具をつけて、自由に自分の絵を描く。 手洗いの表現も学ぶ。
b）親子グループ活動		ギプスを使ってお面を作る。 各国の手遊び、歌などを紹介し合う。
c）日常での支援		身体の部位に関する歌を歌う。

2016年度『慶應義塾 外国語教育研究』第13号投稿規程

１．投稿資格：原則として塾内の教員・職員・研究員（常勤・非常勤を問わない）。共同執筆
　　　　　　者についてはこれ以外の者も可とするが、その場合も投稿筆頭者は原則として
　　　　　　塾内の教員・職員・研究員（常勤・非常勤を問わない）とする。
２．投稿論文の種類：以下の内容のものを掲載対象とし、未発表のものに限る。なお、１）研
　　究論文は特別寄稿を掲載することもある。
　　　１）研究論文
　　　　　・テーマが言語教育あるいはそれに深く関係するものであること
　　　　　・独創性を有する実証的または理論的研究の成果であること
　　　　　・先行研究・関連研究を十分に踏まえていること
　　　　　・他の研究者の検証にも耐えうる、客観性を有すること
　　　２）調査・実践報告
　　　　　・言語教育あるいはそれに深く関係する分野における調査もしくは実践の報告であ
　　　　　　ること
　　　　　・調査・実践内容について具体的かつ明確な記述がなされていること
　　　　　・得られた知見の応用可能性や実践面での問題点について批判的に論じられている
　　　　　　こと
　　　３）研究ノート
　　　　　・テーマが言語教育あるいはそれに深く関係するものであること
　　　　　・未だ論文の形には至らないが、実証的または理論的研究の中間的報告であり、着
　　　　　　想に独創性がみとめられること
　　　　　・他の研究者の検証にも耐えうる、客観性を有すること
３．アブストラクト（概要）
　　　１）研究論文および２）調査・実践報告については、タイトルの下、本文の前に以下の
　　　要領でアブストラクトを記載すること。３）研究ノートについては、アブストラクトは
　　　不要。ただし、応募用紙の５）概要（原則として和文800 〜 1,000字）については、い
　　　ずれの場合も記載すること。
　　　　　・本文が和文・欧文の場合：欧文で150 〜 200語程度
　　　　　・本文が和文・欧文以外の言語の場合：欧文で150 〜 200語程度、あるいは和文で
　　　　　　800 〜 1,000字程度
４．書式・長さ
　　　和文・欧文とも横書きとし、Ａ４用紙に、和文の場合「明朝体」、欧文の場合は「Times」
　　　か「Times New Roman」か「Century」とする。文字の大きさは12ポイント。ページ
　　　番号を記載すること。
　　　脚注・付録・図表、参考・引用文献リストなども以下の字数・語数に含む。ただし、ア
　　　ブストラクトは除く。

字数・語数のカウントは、例えば Word の場合、「校閲」→「文字カウント（字数の場合、スペースを含める）」機能を使用し、テキストボックス、脚注、文末脚注を含めること。

1）研究論文

・和文の場合、20,000 〜 24,000字程度

・欧文の場合、7,000 〜 8,000語程度

・その他の言語の場合は和文に準ずる

2）調査・実践報告

・同上

3）研究ノート

・和文の場合、8,000 〜 10,000字程度

・欧文の場合、2,800 〜 3,500語程度

・その他の言語の場合は和文に準ずる

5．使用言語

特に定めない。

6．応募用紙の添付

所定の応募用紙に、以下1）〜7）の内容を記入し、原稿に添えて提出する。

応募用紙は当センター、Web サイトからダウンロード可能

1）氏名、所属、職位、（担当外国語）

2）連絡先住所、電話番号、e-mail

3）論文の種類、使用言語

4）論文タイトル、総文字数（欧文の場合は総語数）

5）概要　原則として和文800 〜 1,000字とする。

6）キーワード5語（日本語）

7）提出時には応募用紙の全ての項目に記入すること。

7．書式上の注意

引用や参考文献一覧表の形式については、執筆者の分野における標準の形式（例：APAスタイル、MLA スタイルなど）に従うこと。

注は本文の末尾にまとめて付けること。

氏名や所属、住所等は応募用紙にのみ記入すること。アブストラクト（概要）や本文（タイトルの下）に氏名は記入しない。

8．採録の決定および通知

査読を行なったうえで採否を決定する。査読の結果によっては、修正または論文の種類の変更（研究論文から研究ノートへの変更等）を要請することもある。

結果は執筆者に通知する。

9．採録決定後の校正

校正は再校まで執筆者が行なうこととする。校正は誤字・脱字および査読者の指示による修正のみとし、原稿の変更はできない。

10. 論文の公開
1）採録論文は Web 上での公開（慶應義塾大学の運用する Web サイトおよびデータベースへの登録と公開）を前提とする。執筆者によって公開が許諾されない論文は採録しない。
2）文字データ以外に、他から転載された写真・図版等のデータが採録原稿に含まれる場合は、執筆者自身が著作権についての処理を済ませていることを前提とする。
3）採録論文の著作権については、執筆者によって以下の事項が承認されていることを前提とする。
3－1）執筆者は、『外国語教育研究』編集委員会が発行する『外国語教育研究』に掲載された論文等の著作権を当委員会に委託し、当委員会が論文等の編集著作権および出版権を保有する。
3－2）上記3－1に基づき、当委員会は執筆者の論文等を原文のままの形において印刷物または電子媒体により再出版または再配布する権利を保有する。
3－3）執筆者は、論文等を他の印刷物または電子媒体に転載する場合には、当委員会にその旨、通知する。
3－4）執筆者の論文等を要約して印刷物または電子媒体により再出版または再配布する場合は、原則として当委員会は執筆者から事前の同意を得るものとする。
11. 原稿提出締切
2016年9月16日（金）必着とする。
以下1）・2）を印字したものを郵送にて、文書ファイル（MS Word）を e-mail 添付にて提出のこと。なお、提出された原稿は返却しない。また文書ファイルは印字した原稿と同一のものとする。
1）原稿　　　3部（用紙はA4サイズに限る）
2）応募用紙　1部
12. 原稿提出先
（郵送）　〒108-8345　東京都港区三田2-15-45
　　　　　慶應義塾大学外国語教育研究センター三田支部
　　　　　『慶應義塾 外国語教育研究』担当
（e-mail）　flang-kiyo@adst.keio.ac.jp
13. 問合せ先
　　同　上
（電話）　03-5427-1601　内線22304
（e-mail）　flang-kiyo@adst.keio.ac.jp
※開室時間については、上記に問合せること。

以上

Submission Guidelines for the Journal of Foreign Language Education 2016 (vol.13)

1. Full and part-time faculty members of the Keio University system are eligible to submit papers to the journal. The leading author should be a full or part-time faculty member of Keio.

2. Papers may be written in any language.

3. Contributions include research articles, survey/practical reports, and research notes.
    a. Research articles and survey/practical reports in a European language must be within 7,000–8,000 words; articles in a non-European language including Japanese should be within 20,000–24,000 characters (including notes; but abstract not included).
    b. Research notes in a European language must be within 2,800–3,500 words; research notes in a non-European language including Japanese must be within 8,000–10,000 characters.

4. Format:
    – All submissions should conform to the conventions and guidelines of the relevant field (APA, MLA, etc.).
    – Submissions must be typed horizontally and double-spaced on A4 paper.
    – If written in Japanese, please use Minchotai font in 12 point. If written in a European language, please use Times, Times New Roman, or Century in 12 point.
    – Include page numbers.
    – Place all notes at the end of the paper.
    – Name must not appear in the submitted text other than in the submission form.

5. Two abstracts:
    a. The first abstract must be included in the paper following the title. If written in any European Language the abstract must be in 150–200 words. However, if the paper is written in any non-European language other than Japanese, the abstract may be written either in Japanese in 800–1,000 characters or in any European language in 150–200 words.
    b. The second abstract must be provided in Japanese in 800–1,000 characters in the submission form.

6. The deadline for submission is September 16 (Fri). Contributors should submit three copies of the paper, accompanied by an electronic copy as an e-mail attachment, CD-R or flash memory, as well as a submission form. The submission form is available for download at the Keio Research Center for Foreign Language Education website. (www.flang.keio.ac.jp)

You must complete all items in the submission form and send to:

Office of *Journal of Foreign Language Education*
Mita Office
Keio Research Center for Foreign Language Education
2-15-45 Mita, Minato-ku, Tokyo 108-8345 Japan

7. Papers will be reviewed by the Editorial Committee.

8. Printer's proofs will be provided twice. Major revisions other than those suggested by the Editorial Committee will not be accepted.

9. Papers will be made available for public view on the Keio Research Center for Foreign Language Education website. If a contributor borrows from another author's work/data, the contributor must obtain that author's consent concerning the copyright of photographs and/or illustrations, before the publication of the paper, normally at the first proof stage. The Editorial Committee holds the copyright and publishing rights of the published papers including the reprint right. If a contributor wishes to republish their paper elsewhere, the author must inform the Editorial Committee.

Inquiries:
Postal address as above.
tel.: + 81-3-5427-1601 ext. 22304
e-mail: flang-kiyo@adst.keio.ac.jp
URL: http://flang.keio.ac.jp/

執筆者紹介

大　竹　優　志 …………………… 慶應義塾高等学校 教諭

北　川　　　彩 …………………… 慶應義塾高等学校 教諭

プリエト ムニョス,　…………………… 慶應義塾大学
　　　　ベアトリス 商学部・総合政策学部 講師（非常勤）

シェイ デビット …………………… 慶應義塾大学
商学部 准教授

折田（由川）京子 …………………… 慶應義塾高等学校 教諭

伊　藤　　　扇 …………………… 慶應義塾幼稚舎 教諭

矢田部　清　美 …………………… 慶應義塾大学
外国語教育研究センター 助教

皆　川　泰　代 …………………… 慶應義塾大学
文学部 准教授

須　山　哲　治 …………………… 慶應義塾大学
文学部 准教授

山　下　一　夫 …………………… 慶應義塾大学
理工学部 准教授

吉　川　龍　生 …………………… 慶應義塾大学
経済学部 准教授

嚴　　　　　馥 …………………… 慶應義塾大学
法学部・総合政策学部 講師（非常勤）

牛　山　さおり …………………… 慶應義塾大学
外国語教育研究センター 兼任研究員