

Recorded Speaking Tasks as Formative Assessment in Intermediate English Classes

Adrian WAGNER

Introduction

Intermediate English (中級英語) classes are compulsory classes for all second year students in The Faculty of International Studies and Liberal Arts (国際教養学部). Each student takes four 90 minute classes per week. Classes are divided into 4 skills as follows: Intermediate English Reading, Intermediate English Listening, Intermediate English Writing, and Intermediate English Oral Communication. The curriculum is designed to foster practical English ability with a balance of receptive and productive skills. A possible pattern would be that a student would take two classes of Intermediate English Reading and two of Intermediate English Oral Communication in the first semester, and in the second semester, two classes each of Intermediate English Listening and Intermediate English Writing. These two classes of each skill may be with the same teacher or different teachers. However, each class must be treated separately and assessed inde-

Keywords: speaking assessment, formative assessment, self-assessment, language learning plateau, curriculum design

pendently of the other. For example, if a group of students had the same teacher for Intermediate English OCa on a Tuesday and OCb on a Wednesday, in principle the two classes the assessment tasks and grades of each should in no way inform those of the other.

As of 2018, due to Ministry of Education, Culture, Sports, Science and Technology (文部科学省) regulation, all of the individual Intermediate English classes share a common syllabus. However, due to differences in ability level etc. individual teachers are allowed a degree of flexibility in their interpretation of the common syllabus, such as the freedom to select their own textbooks and specific methods of assessment.

Below is the course summary of Intermediate English Oral Communication:

This second year course builds upon the skills developed in first year English courses.

The goal of this class is for students to increase their fluency and accuracy in spoken English to the “B1” level of CEFR. In class, students will have time and opportunity to practice spoken interaction in English through a range of tasks and topics.

Below is the list of class goals and description of attainment goals and class themes:

Students will be able to:

- *use appropriate grammar, vocabulary and English phrases essential for communicating about common and familiar topics such as hobbies, work, interests, preferences, experiences etc.*
- *express and briefly explain their opinions about various topics.*
- *speak at length about familiar topics, connecting utterances into longer*

passages of speech

Students will work toward:

- *clear and comprehensible pronunciation and natural patterns of intonation*
- *conversing about more abstract or academic topics*
- *using spoken English appropriate for professional and formal contexts*

With the general curriculum and individual syllabus calling for class content targeted towards a specific skill, it should follow that assessment should also reflect effort, improvement, and achievement in that particular skill. Whatever the class, assessment should be something that both teachers and students take seriously. For the receptive skills, tests focusing on comprehension of texts and recordings would be the most appropriate form of test. In the writing class a teacher could assign essays, or tests focusing on grammar and structure of essays. However, practical and reliable ways to assess ability of second language oral communication skills (speaking ability) are not so easy to find. Keeping in mind the syllabus of the class and the characteristics of the learners, I decided to use the facilities of the language learning laboratories to have students make recordings of themselves performing various speaking tasks and use this as the basis for assessment.

The assessment system discussed in this essay is an attempt to find a method of assessment that matches the design and goals of the Intermediate English curriculum and inspires class participants to greater achievement. For this purpose, the assessment method was used over two semesters with two different groups of students. Their reactions and feelings about this method of assessment were collected via an anonymous survey at the end of each semester.

Literature Review

Assessment of Speaking Ability in Second Language Classes

There is a plethora of research conducted regarding the theoretical underpinning of any testing in second language ability and any educator wishing to inform their own assessment practices will be drawn into at times philosophical debates such as what it means to know a language (Spolsky, 2007) and other controversial topics such as the sometimes intangible nature of development or scaling of proficiency in second acquisition.

Alderson (2007) writes:

the levels of attainment expected of (typically school-based) language learners, are often vague, ill-defined, lack any empirical base, and bear little relation to theories of second language acquisition. In short, it is far from clear exactly what changes as learners develop and therefore what diagnosis of second language development (or lack of it) should be based on, or how diagnostic tests might be validated. (p. 21-22)

Furthermore, a series of other factors, not related to the language proficiency of learners affect test performance and cause doubts regarding the reliability and validity of any test. Bachman (1990), writes that performances on language tests can be influenced by personal attributes of learners such as such as “cognitive style and knowledge of particular content areas, and group characteristics such sex, race, and ethnic background” (p. 164).

Next, there are concerns of bias in any rating of language tests, particularly speaking assessment. The study of Lumley, and McNamara (1995) shows that even trained and certified assessors show inconsistency in their evaluations and level of severity. As a teacher assessing our own students, who we

have met and interacted with many times on a social level, there are surely biases on a personal level that could influence any judgment.

Overall, even in the best circumstances, any evaluation of speaking performance, and the score or grade it elicits, is likely to be somewhat arbitrary and erroneous.

Bringing the discussion of speaking assessment to a more practical place, Ginther (2012) writes that:

Speaking is seen by language testers as the most difficult of the four language skills to assess. Assessing speaking requires that we either observe a “live” oral performance or that we capture the performance by some means for evaluation at a later time. A method for elicitation must be selected, rating scales must be developed, and interviewers and/or raters must be trained. (p. 1)

The type of test that Ginther is describing is often called direct testing as opposed to indirect testing styles. Indirect tests evaluate the skills and abilities that underlie an examinee’s performance by eliciting performance on item types such as a multiple-choice main idea items to measure reading comprehension. Common styles of direct test are the one on one interview style as employed by the IELTS test and paired speaking tests such as employed by the Cambridge Suite of exams. Other variations could be the recorded responses used by the TOEFL iBT test, or even presentations.

Speaking of large scale- standardised tests such as IELTS, the study of Wagner (2015) noted, commercial and standardised language tests can be a valid reference or starting point for thinking about test styles, grading rubrics etc. However, it can be difficult to scale and adjust these tests for use in a school or university classroom setting.

The conclusion that must be reached is that there is no perfect way of testing the spoken proficiency of language learners, and teachers who wish to do so must use the information available to develop methods that are as practical and fair as possible in their educational context.

Formative Assessment and Self-Assessment in Second Language Classes

Not only in the field of second language acquisition, but in any educational context, in-class assessment can generally be classified as either summative or formative. Summative assessment is given at the conclusion of a course of study and seeks to evaluate overall achievement and attainment and retainment of skills or knowledge. Taras (2005) writes that summative assessment provides, “a judgement which encapsulates all the evidence up to a given point. This point is seen as a finality at the point of the judgement” (p. 468). It represents an end point and the information gleaned by both the learner and the teacher serves to report upon achievement but not enhance it.

On the other hand, formative assessment generally occurs during the course of study, and occurs numerous times. The term, ongoing assessment is also used but a key difference between formative and summative assessment is not only timing and frequency. As Taras (2005) writes:

... for an assessment to be formative, it requires feedback which indicates the existence of a ‘gap’ between the actual level of the work being assessed and the required standard. It also requires an indication of how the work can be improved to reach the required standard. (p. 468)

The imperative is that the assessment, or the information gleaned from it, must be used to inform class content and teaching practices. Learners themselves, should also be a part of the process.

Numerous studies, such as that of Ross (2005) and Alvarez, Ananda, Walqui, Sato, and Rabinowitz (2014) argue that formative assessment, if carefully planned and administrated can play a vital role in second language education. Generally, it can be seen as making learners more consciously involved in the learning process.

The practice of self-assessment is also linked to formative assessment as it can be seen as a way to foster learner autonomy, aiding learners take a more active and critical role in their own progress and achievements. Obviously, taking that concept to the extreme and having students become solely responsibility for their own assessment and grading can be seen as unfair and irresponsible educational practice.

While acknowledging the shortcomings of self-assessment in language learning in a formal context, Brown and Hudson (1998) list the following advantages of incorporating self- assessment:

First, self-assessments can be designed to be administered relatively quickly. Second, they inevitably involve students directly in the assessment process. Third, in turn, such involvement may help students understand what it means to learn a language autonomously. Finally, both the students' involvement and their greater autonomy can substantially increase their motivation to learn the language in question. (p. 666)

Learner Autonomy is integral for long-term success in second language acquisition and for this reason alone, it is worth considering incorporating some form of self-assessment into any curriculum.

Plateaus in Language Learning

When designing a syllabus for and considering the needs of intermediate

language learners, it is useful to be aware of the concept of plateaus in language learning. Not only limited to second language acquisition, the feeling of reaching a plateau or slump, when initially smooth improvement or progress is felt to slow down or cease completely, is commonly reported by people studying other subjects, learning sports, and on diets etc. Anecdotal evidence from both learners and educators suggests this often happens at the intermediate level. Richards, (2008) writes:

...once learners have arrived at an intermediate level of language learning, progress does not always appear to be so marked, and making the transition from intermediate to the upper-intermediate/advanced level sometimes proves frustrating. Some may feel they have arrived at a plateau and making further progress seems elusive, despite the amount of time and effort they devote to it. (p. 1)

Particularly, Richards (2008) gives five characteristics of learners who may have come to a learning plateau:

1. There is a gap between receptive and productive competence.
2. Fluency may have progressed at the expense of complexity.
3. Learners have a limited vocabulary range.
4. Language production may be adequate but often lacks the characteristics of natural speech.
5. There are persistent, fossilized language errors. (pp. 2-3)

Particularly, the first, third and fifth of these characteristics informed the planning for this Intermediate English class and the style of assessment. Considering Japanese learners, and the predominantly grammar and reading based style of English education in schools in Japan, the gap between receptive and productive competence is an issue that demands attention. As for the lack

Recorded Speaking Tasks as Formative Assessment in Intermediate English Classes of vocabulary range, a large portion of the textbook chosen and the class content focused on building vocabulary, particularly academic vocabulary and collocations. Finally, when assessing the recorded speaking tasks, I made efforts to identify and correct persistent repeated errors, particularly in grammar and pronunciation. These factors will be discussed in more detail in following sections of this paper.

The Class and Study Participants

The research was conducted in two classes of Intermediate English Oral Communication. One in the spring semester of 2018 and one in the spring semester of 2019. In both classes, students had two different teachers for their two weekly Intermediate English Oral Communication classes. Meaning, I met the students once per week for the 15 weeks of the semester. The 2018 class contained 13 students and the 2019 class contained 16 students. As there were no practical differences in the way the class and assessment were conducted or other shared differences between the groups distinct to their year level, the 2018 and 2019 cohorts will be treated together as a single group for the purpose of this research.

At the time of participating in the class, all students were in the first semester of their second year of university, the majority of students were Japanese and completing Intermediate English as a compulsory course of the Faculty of International Studies and Liberal Arts. However, there was a vast range of differences in ability levels and motivation. Approximately half of the students had experience of studying abroad or plans to study abroad while at university. Some students attended every class and submitted all assessment tasks on time, while others had numerous absences and failed to submit tasks on time.

The textbook chosen for the class was *Unlock Level 3 Listening and Speaking Skills Student's Book and Online Workbook* by Sabina Ostrowska.

The first five chapters of the book were covered in the semester. The themes of each chapter were: 1. Animals, 2. Customs and Traditions, 3. History 4. Transport, and 5. Tradition. In general, the topics and content of the book were quite academic in nature. Listening exercises were often based on academic lectures and the textbook included a large amount of academic and formal vocabulary and phrases. While the level was quite high and students did struggle with some of the content in classes, I thought this was appropriate considering the latter part of the stated goals for Intermediate English Classes, which included preparing students for: "... conversing about more abstract or academic topics and using spoken English appropriate for professional and formal contexts". Also, the stated level of the textbook was CEFR B1, the target specifically listed on the Intermediate English Curriculum

A key component of this textbook series is the incorporation of end-of-unit tasks. Completing these end-of-unit academic tasks necessitated students engaging with the textbook materials on a thematic level; to both understand the issues presented and to formulate their own opinions. Practically, they were instructed to use target phrases, vocabulary and grammatical structures from the unit to complete the task. These speaking tasks were used adapted for use as assessment tasks.

Assessment Method

As mentioned above, five chapters were covered over the 15 week semester. So, following a pattern of formative/ongoing assessment, the semester was divided into five cycles. There was one cycle per chapter of the text

Recorded Speaking Tasks as Formative Assessment in Intermediate English Classes

consisting of two regular lessons and one class of assessment. This meant that a substantial amount of class time (one third) was used for assessment tasks. However, considering that the assessment tasks were devised as revision and learning tasks, this was seen as not overly excessive.

So, in every third lesson the venue of the class was moved from the regular classroom to the language laboratory equipped with computers and recording software. First students took a regular exam, beginning with either one or two listening tasks, followed by grammar and vocabulary tasks. This paper part of the test was closed book and designed to take about 45 minutes of the total 90 minute class time. After a student had completed the first task of the test, they were given the speaking task. This part of the test was open book, with students told to use their textbooks to aid in the speaking task. Also, as it was hoped that doing the speaking tasks would be a learning activity as well as an assessment task, students were permitted to use dictionaries and the Internet to research the topic for ideas and specific examples to talk about. They were instructed to record their speaking task and email as an attachment by the end of the lesson. Students were encouraged to record as many times as they like until they were satisfied with their achievement or until they ran out of time. Generally, students completed the tasks in the computer room. However, some students who said that they were shy or unable to concentrate in the classroom chose to leave the classroom and record their speaking tasks on their smartphone in private. After the class students were sent a link to a Google Form which instructed them to re-listen to their submission and conduct self-assessment. The flow of the assessment day is summarised in the table below. Then, using the recorded submission and self-assessment form, students were given personalised feedback written by the teacher, highlighting

things they did well along with errors in grammar, vocabulary, punctuation etc. This was given to students two weeks after the day of assessment. All five Assessment Tasks and the Self-Assessment Form can be found as appendices. A summary of the assessment flow is found in Table 1 below.

Table 1

Assessment Flow

<u>Time</u>	<u>Activity</u>
First Half of Lesson	Listening, vocabulary and grammar test (closed book)
Second Half of Lesson	Recorded speaking tasks (open book)
After class (within one week)	Students submit self-assessment via Google Forms
Two weeks after submission	Students are given personalised feedback on paper

Regarding characteristics of learners whose progress has plateaued, Richards (2008) offers the following advice for combating the gap between receptive and productive competence. Some of his advice for educators to guide students towards progress in this area, borrows heavily from *Output Hypothesis*, first proposed by Swain (2008). This theory purports that not only comprehensible input, seen by researchers such as Krashen (1985) as the sole integral condition for successful language learning, but also comprehensible output also plays a key part in the second language acquisition process. Specifically, Richards suggests that:

Carefully structured and managed output (the output hypothesis) is essential if learners are to acquire new language. Managed output here refers to tasks and activities that require the use of certain target-language forms that “stretch” learners’ language knowledge and that consequently require a “restructuring” of that knowledge. The output hypothesis suggests the rather obvious notion that practice in using

target language forms is necessary for learners to acquire new target language. (p. 6)

The assessment activities were designed specifically to necessitate this kind of focused practice.

Next, the introduction of the self-assessment to the assessment task, mirrors Richard's advice for combating fossilised errors. He suggests that it is beneficial that learners become, "active monitors of their own language production through listening to recordings of their own speech and through having others monitor their speech for fossilized errors in focused listening sessions" (p. 20). The responses from students did indeed show that this created opportunity, in conjunction with the personalised feedback given by the teacher, was somewhat effective in this area.

Grading of Assessment Tasks

As there was a vast difference in ability levels of individual students in both cohorts, it was difficult to grade these recorded speaking tasks with a standard rubric, either analytical or holistic. Regarding the quality of the final product, a sound achievement for one student could be seen as a poor achievement for one of the higher level students or an excellent achievement for one of the lower ability level students. The goal was to foster motivation and development of ability, so these tasks were not graded like a standardised test such as IELTS or TOEFL.

Firstly, the written test, was sufficient to show which students had studied for the test and memorised new vocabulary, collocations and grammar constructions from the textbook and represented half of the total score for the test. Equal points were then allocated for the speaking task. For example, if

a test had a total of 50 points for the written section, 50 points were allocated for the recorded speaking task, making a total of 100 points. If 35 points were allocated for the written section, the speaking task would be a worth 35 points for a total of 70 points. As for scoring the recorded speaking tasks, before the teacher listened to it and read the self-assessment, each student started with full points and the product was assessed in a decremental manner. Students would lose points for not fulfilling the task instructions such as not using target phrases or going under time. Points were also lost for submitting self-assessment late or not completing satisfactorily.

As for the final grade of the class, each of the five tests were calculated to 15% for a total of 75%, with the remaining 25% allocated for attendance, participation and weekly homework. Over the two years, the lowest assessment score was 35% from 75% and the highest was 73%. The median assessment score over the two years was 58.5%.

Data Collection

In both cohorts, data was collected via an anonymous questionnaire conducted after the final class through Google Forms. Students were informed that the questionnaire was not part of assessment and would not affect their final grade in any way. It was also to be completed on a voluntary basis. Out of a total of 29 students, 27 students completed the final questionnaire. The data from both years was combined for analysis. The questionnaire consisted of 17 questions. 16 were Likert Scale questions. With the exception of Question 1, all of these Likert Scale questions had five response choices: *Strongly agree*, *Agree*, *Neither agree nor disagree*, *Disagree*, *Strongly disagree*. Finally, there was a question at the end where students could freely write

Recorded Speaking Tasks as Formative Assessment in Intermediate English Classes

their opinions or feelings about the recorded speaking tasks and self-assessment. The questionnaire was provided in both Japanese and English and is included as Appendix A.

Results and Analysis of Data

The overall goal of the questionnaire was to see to what degree and in what particular ways students found this process and method of assessment useful in improving their English ability and self-awareness as learners.

Responses to the end of semester questionnaire from the 2018 and 2019, were compared, but as there were no noticeable differences or biases, the data was combined. For ease of understanding, responses to the questions will be analysed out of order, with questions regarding similar themes discussed together. Responses to the final question, where students could write freely, are included in the relevant sections. Responses in English are unedited and responses in Japanese are translated into English.

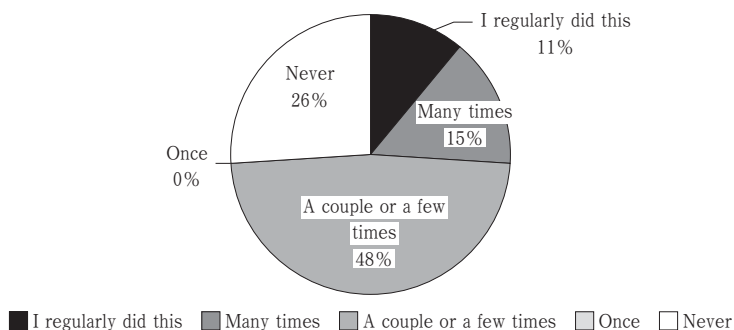
Familiarity with Task Style

The first question sought to discover what experience students had with recording their own speaking as a way of practicing English. The results below showed that the majority of students had no experience or limited experience of this.

As only 26% of students chose *I regularly did this or Many times*, it shows that this was not a regular part of the English study regime of most students. So, the assessment tasks provided a sense of doing something new, or at least some novelty that differentiated this class from others they had taken.

Graph 1

Question 1. Before this class have you had the experience of recording your speaking and listening to your own English?

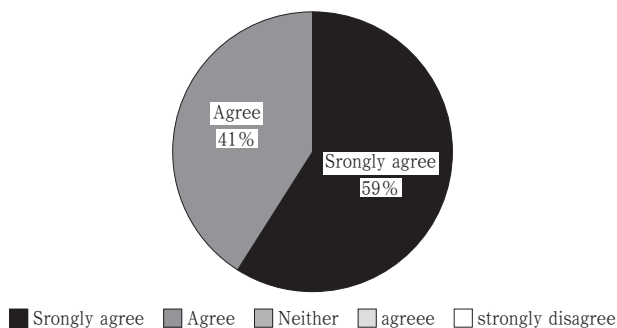


Connection with Class Content

As any assessment should have an obvious and direct connection with what was taught and what is expected to have been learned, Questions 3, 4, and 5 gauged students perception of the relevance of assessment tasks to class content. Responses to these questions are presented in the graphs below.

Graph 2

Question 3. I felt there was a direct connection between the speaking tasks and class activities/class textbook.

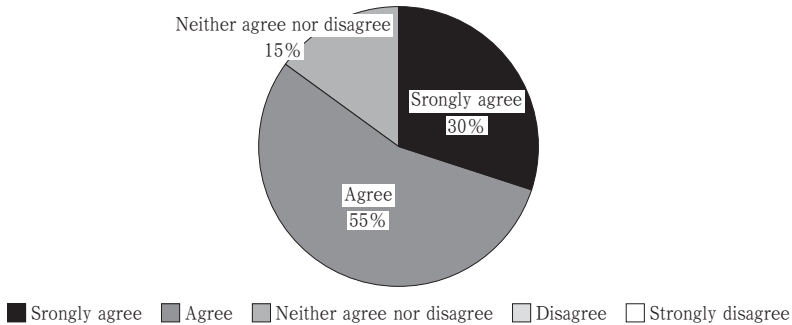


The student responses clearly showed that students not only felt a connec-

Recorded Speaking Tasks as Formative Assessment in Intermediate English Classes

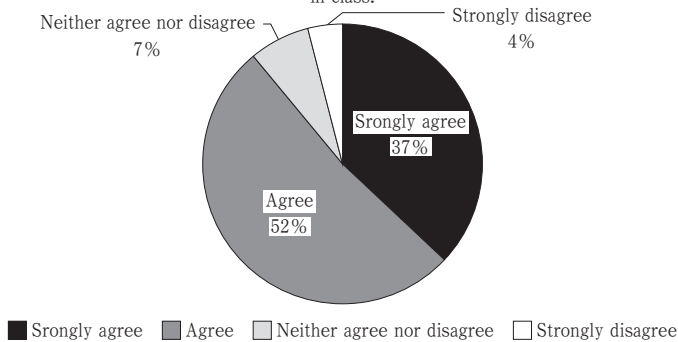
Graph 3

Question 4. Doing the speaking tasks and self-assessment caused me to revise the class content and textbook.



Graph 4

Question 5. Doing the speaking tasks and self-assessment helped me to remember grammar, vocabulary, phrases and pronunciation advice learned in class.



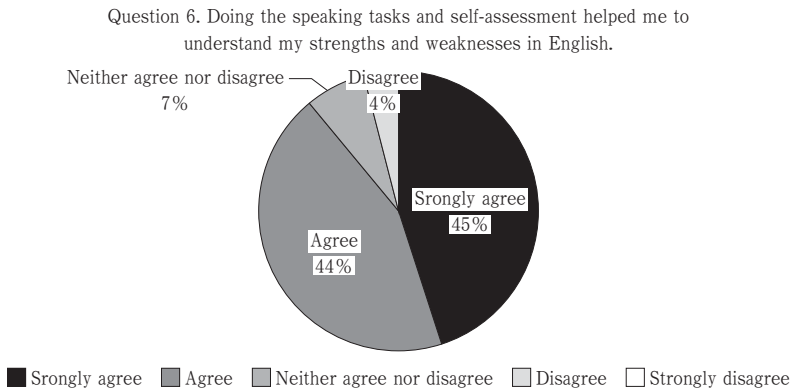
tion between the assessment tasks and class content, but also that the task was a catalyst for memory and revision of what had been learned in previous weeks.

Students' Awareness of Strengths, Weaknesses and Mistakes in Language Production

As both questions refer to students' ability to reflect on their ability and assess their own achievements, the responses to Questions 6 and Question 13 will be discussed together. The results show that the doing the recorded speaking tasks with the follow up of self-assessment made students notice their own mistakes and think about their strengths and weakness.

Probably, the feedback from the teacher was also a part of this. Students had a tendency to make similar mistakes, repeatedly in different assessment tasks. Some examples of repeated and possibly fossilised were mispronunciation of certain sounds or errors in forming and using particular grammatical structures.

Graph 5

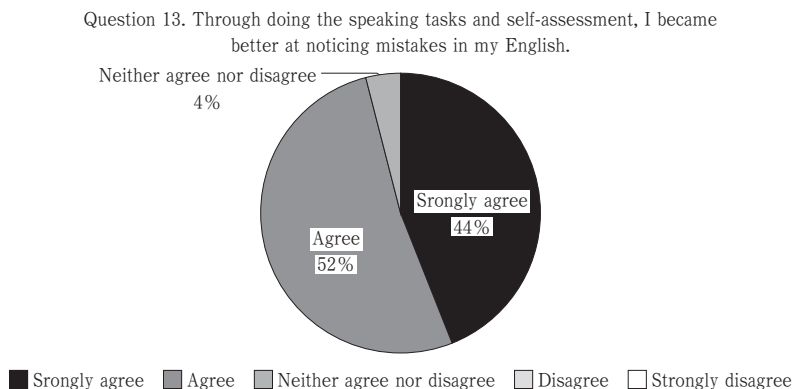


Probably, the feedback from the teacher was also a cause of almost 90% of the students selecting *Strongly agree* or *Agree* as their response to Question 6. Having their most frequent and prominent errors pointed out to them on numerous occasions would hopefully increase their sensitivity regarding those

Recorded Speaking Tasks as Formative Assessment in Intermediate English Classes

points of language production.

Graph 6



As for Question 13, students were told that identifying errors was an important part of the self-assessment stage of the assessment task and therefore became more skilled at doing this over the course of the semester. Particularly, the self-assessment provided extrinsic motivation to listen carefully to their recordings and identify errors.

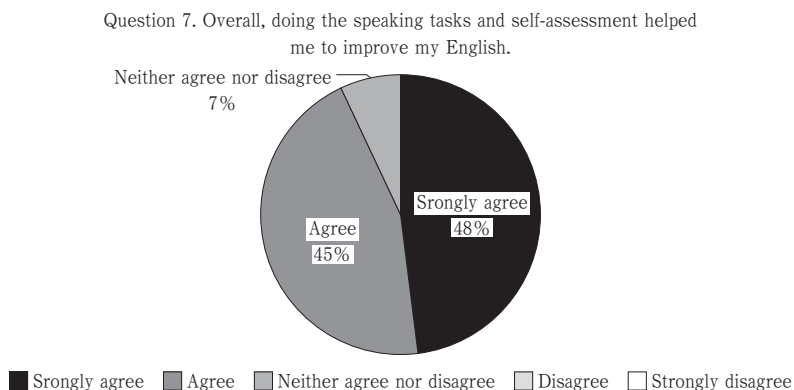
Student response: *I had not had many opportunities to hear myself speaking English before, so I became more conscious of my pronunciation and way of speaking* (translated from Japanese).

Improvement of English Ability

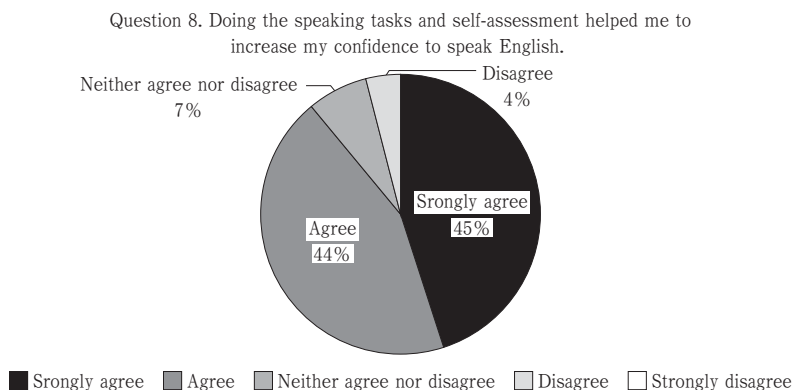
Of course, the ultimate goal of any course of study is to increase participants' knowledge or ability in the subject. As the assessment in this course was practical and formative, a key part of the assessment design was to help students feel a sense of achievement and improvement over the semester. Data from the end of semester questionnaires indicate that this was indeed the

case. Regarding general increase of ability and confidence, the majority of students responded that they felt an improvement in their overall English ability.

Graph 7



Graph 8

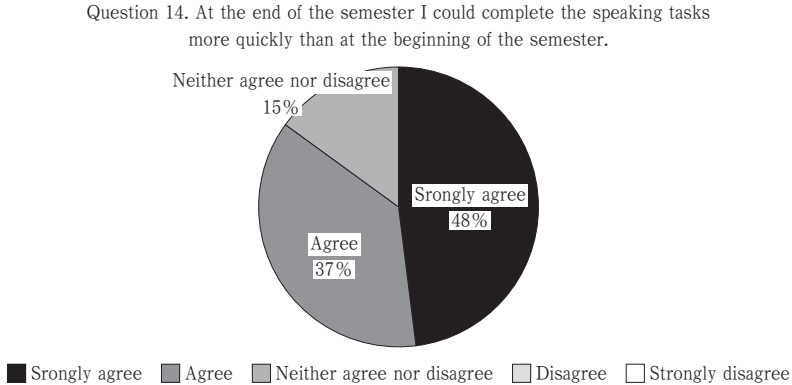


Also, overall, most of the students felt an increase in their confidence to speak English. Students remarked in the class that the assessment tasks were generally more complicated than the speaking tasks they had done in other

Recorded Speaking Tasks as Formative Assessment in Intermediate English Classes

classes and required more preparation and effort. Especially, at the beginning of the semesters, students' responses to the individual self-assessment tasks indicated that they often wrote a full script for their monologue or dialogue and re-recorded numerous times. Over the course of the semesters, more students reported writing notes rather than a full script and recording on average less times before they were satisfied with the final product. This observation is supported by the students' responses to Question 14. Which asked about the speed at which they could complete the tasks.

Graph 9

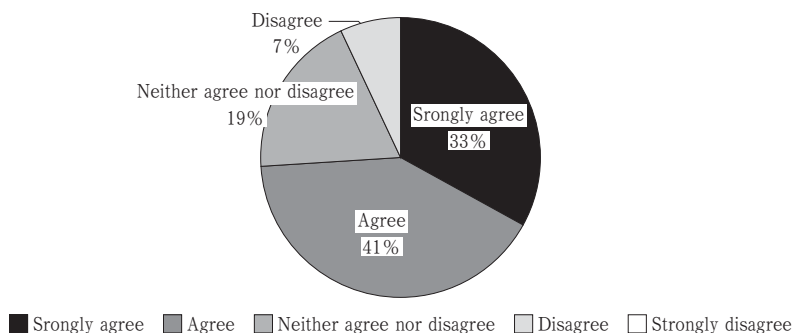


Out of the 27 students, only 4 did not report a perceived increase at their ability to complete the tasks faster over the semester.

Generally, students felt less slightly less improvement in grammar and vocabulary knowledge than in speed of completion and overall confidence. However, the majority still reported a sense of improvement in this area. Perhaps, this was due to the rules of the assessment tasks which directed students to use key vocabulary and certain grammatical constructions, which were often new to students, featured in the textbook. Therefore students had

Graph 10

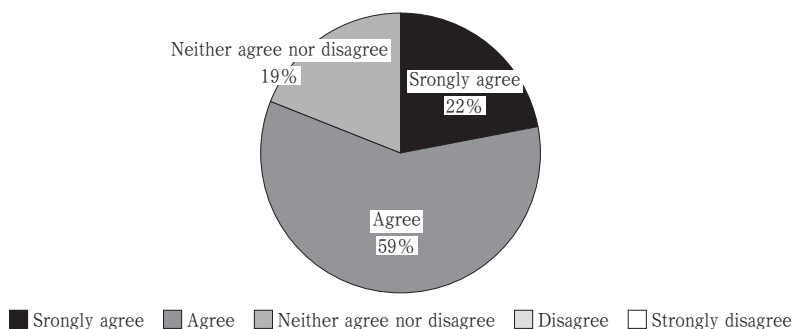
Question 9. Doing the speaking tasks and self-assessment helped me to improve my grammar.



to understand the meaning of these words and phrases and incorporate them into their own speech, triggering learning and hopefully memorisation.

Graph 11

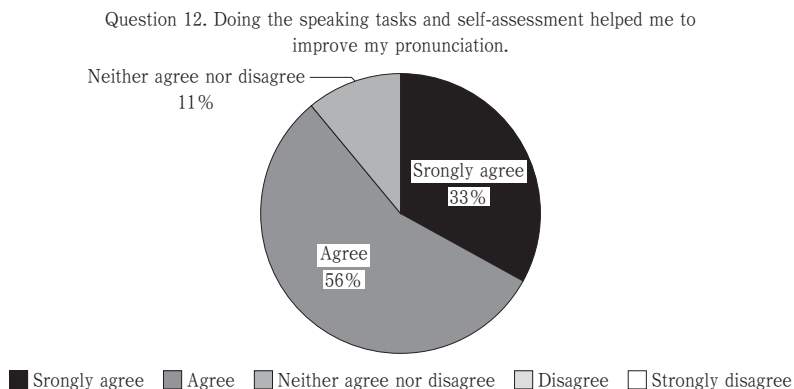
Question 10. Doing the speaking tasks and self-assessment helped me to improve my vocabulary.



As indicated in the responses to Question 1., recording and listening to their own speech in English was not a regular way of practising English for the students. Therefore, students had numerous chances to listen and objectively assess their own pronunciation. In fact, errors in pronunciation, particularly

Recorded Speaking Tasks as Formative Assessment in Intermediate English Classes

Graph 12



sound distinctions that are difficult to Japanese learners such as *l/r*, *θ/s*, */z*, were some of the most common errors identified by students in their self-assessments and pointed out in the feedback from the teacher. Surely, this led to an overall increase of consciousness of pronunciation, particularly problematic sounds.

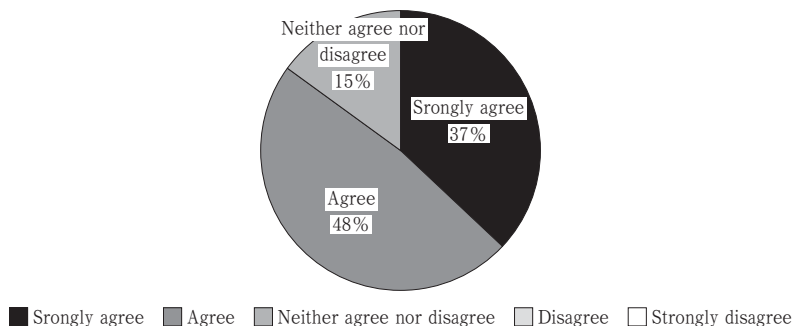
Finally, more than one third of the students chose *Strongly agree* when asked about improvement of their spoken fluency.

Often, during in-class speaking activities, it can be struggle for teachers to have Japanese learner speak at length on any topic. Due to factors such as general shyness, concerns about grammatical accuracy, lack of vocabulary or general lack of knowledge or even a clear opinion on a subject, students are often reticent to give more than short answers. In completing the recorded speaking tasks, students were forced out of their comfort zone, and made to speak at length and in depth about a range of topics.

Overall, considering one of the key considerations of this class was to cater for intermediate learners who may well have arrived at a learning plateau, the

Graph 13

Question 11. Doing the speaking tasks and self-assessment helped me to improve my fluency.



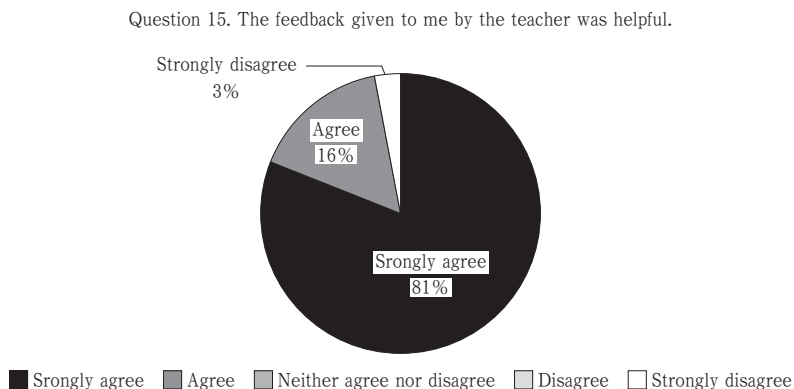
results of the questionnaire were overwhelmingly positive. Indeed, students claim to have felt that they got better at speaking English.

Feedback from the Teacher

The most time consuming and difficult part of this assessment design from a teaching and administrative standpoint was providing personalised feedback to each student for each of the speaking tasks. This required a lot of time to listen to each student's recording numerous times and point out errors, with corrections, explanations and advice. Deciding which errors to highlight and which to ignore was also difficult. Receiving feedback with too many errors identified could be overwhelming and demoralising to the students. However, from a teaching perspective, the types of errors and frequency of the errors provided a lot of useful information to the teacher in re-teaching points from previous classes and informing future class content. The graph below shows that almost all of the students appreciated the feedback and felt it was useful.

Student response: *I think speaking task and self-assessment will help me*

Graph 14



notice where I have missed about grammar or vocabulary and pronunciation. And teacher also checks it and find some mistake in the speaking task and feedback for my next class. I think it is very useful for me to improve my English skill.

Enjoyment and Continuation

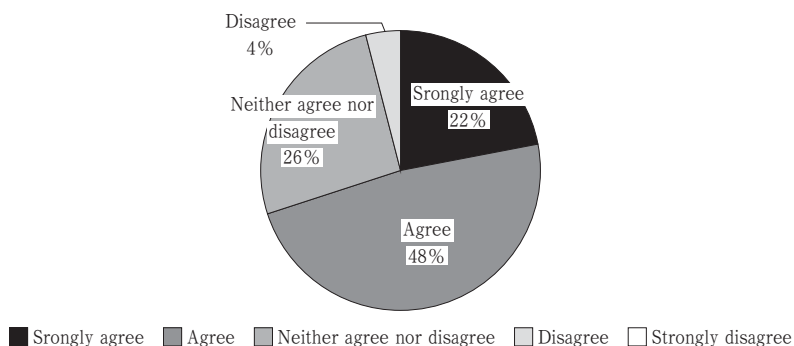
While “fun” may not be a stated goal of most courses at a university, enjoyment of the learning process is an important consideration and can have a strong effect on the motivation of students. Question 2. gauged the enjoyment level of students.

The responses show mixed results. While almost half of the students chose the response *Agree*, less than a quarter chose, *Strongly agree*.

However, a much higher percentage of students expressed a desire to continue recording themselves speaking as a way to improve their English ability in the future. It could be inferred that due to the challenging nature of the tasks, although some students did not particularly enjoy it, they saw these tasks as useful for their progression as an English user.

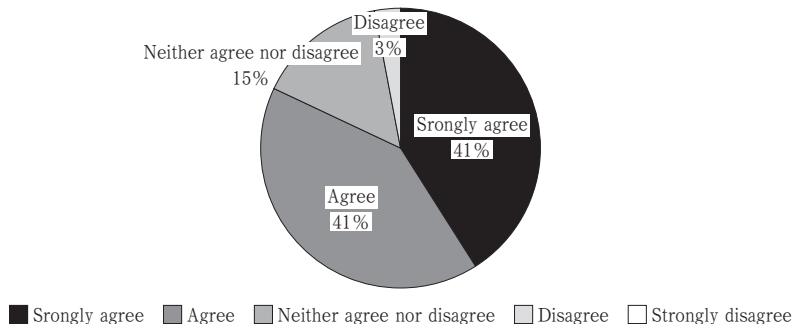
Graph 15

Question 2. I enjoyed doing the speaking tasks and self-assessment.



Graph 16

Question 16. In the future, I want to continue recording myself speaking English as part of my English study.



While the results displayed on the graph appear to be encouraging, they must be taken with the caveat that no information is available as to whether students are indeed continuing to record their own speech as a form of independent study.

Student response: *I really appreciate having us think many kinds of topics in class and your speaking feedback. Especially feedback, I could understand my*

Recorded Speaking Tasks as Formative Assessment in Intermediate English Classes

weak points. So I hope to do a kind of the activity such as feedback in autumn class as well.

Limitations of This Study

This action research was quasi-scientific in nature. It comprised of a relatively small group of participants/respondents. Also, there is no “control group” or point of comparison to how the same groups of students react to different styles of speaking assessment. Furthermore, all of the data was self-reported, which creates doubts regarding its reliability and validity. Although students voluntarily completed the questionnaire, and were assured that it was completely anonymous, they may have had lingering concerns that their responses could have an effect on their final grades for the course. Google Forms does have the capacity to record the email addresses (which are student numbers) of respondents. This function was switched off and it was not possible to identify individual respondents. However, there this the possibility that students were not completely convinced that this was the case. Therefore, due to the factors mentioned above, the conclusions of this research could not be considered generalisable. They are meant only as a guide or point of reference for educators working in similar contexts with an interest in curriculum and assessment design who want their tests to give “results” other than a number or a grade.

Conclusion

From the data collected, and the reactions of students observed by the teacher the combination of recorded speaking tasks and self-assessment was seen as a beneficial and meaningful activity by students. This form of assess-

ment is more practical and logistically easier to administer than other forms of speaking assessment such as individual interview style tests. Perhaps the most beneficial aspect, and the part most appreciated by students is it gives students the opportunity to hear and assess their own production of language and gives the teacher the ability to provide useful personalised feedback.

References

- Alderson, J. C. 2007. The challenge of diagnostic testing: Do we know what we are measuring? in J. Fox, M. Wesche, D. Bayliss, L. Cheng, and C. E. Turner (eds): *Language testing reconsidered* (pp. 21-39). Ottawa: University of Ottawa Press.
- Alvarez, L., Ananda, S., Walqui, A., Sato, E., & Rabinowitz, S. (2014). *Focusing formative assessment on the needs of English language learners*. San Francisco, CA: WestEd.
- Brown, J. D., & Hudson, T. (1998). The alternatives in language assessment. *TESOL Quarterly*, 32(4), 653-675.
- Ginther, A. (2012). Assessment of speaking. In C. A. Chapelle (Ed), *The Encyclopedia of Applied Linguistics*. Oxford: Wiley-Blackwell.
- Krashen, S. D. (1985), *The input hypothesis: Issues and implications*. New York, NY: Longman.
- Lumley, T., & McNamara, T. F. (1995). Rater characteristics and rater bias: Implications for training. *Language Testing*, 12(1), 54-71.
- Ostrowska, S. (2014). *Unlock Level 3 Listening and Speaking Skills Student's Book and Online Workbook: Listening and Speaking Skills Student's Book + Online Workbook*. Cambridge University Press.
- Richards, J. C. (2008). *Moving beyond the plateau*. New York, NY: Cambridge University Press.
- Ross, S. J. (2005). The impact of assessment method on foreign language proficiency growth. *Applied linguistics*, 26(3), 317-342.
- Spolsky, B. (2007). On second thoughts. In J. Fox, M. Wesche, D. Bayliss, L. Cheng, C. Turner & C. Doe (Eds.), *Language testing reconsidered* (pp. 9-18).

Ottawa, Canada: University of Ottawa Press.

Swain, M. (1985). Communicative competence: some roles of comprehensible input and comprehensible output in its development. In S. Gass, and C. Madden (Eds), *Input in second Language Acquisition* (pp.235-253). Rowley, MA: Newbury House.

Taras, M. (2005). Assessment-summative and formative-some theoretical reflections. *British journal of educational studies*, 53(4), 466-478.

Wagner, A. (2015). Designing speaking tests for Freshman English classes: looking to the commercial sector for inspiration. *英語教育研究所紀要 (CELE Journal)*, 23, 66-82.

Appendix A

Speaking Task and Self-Assessment Survey

スピーキング課題と自己評価に関するアンケート

* Required

The information collected in this survey will be used for research purposes. Names and students numbers will not be recorded. This information will not be used for grading.

このアンケートの回答内容は、研究目的のためにのみ使用されます。学生番号・氏名は記録されず、成績に反映されることはありませんので、正直に回答してください。*

Continue 次へ進む

Question 1. Before this class have you had the experience of recording your speaking and listening to your own English? 質問1：このクラスを受講する前、自分が英語を話しているのを録音し、聞いたことがありましたか？ *

I regularly did this. 定期的にしていました

Many times 何度もしたことがあった

A couple or a few times 数回したことがあった

Once 一度だけした

Never ない

Question 2. I enjoyed doing the speaking tasks and self-assessment. 質問2：

Speaking Task（スピーキング課題）と Self-assessment（自己評価）に、興味を持って取り組みましたか？ *

Strongly Agree 強くそう思う

Agree そう思う

Neither agree nor disagree どちらとも言えない

Disagree あまりそう思わない

Strongly Disagree そう思わない

Question 3. I felt there was a direct connection between the speaking tasks and class activities/class textbook. 質問3：Speaking Task（スピーキング課題）は、クラス内で使用された教科書や行われたアクティビティに、直接関係があると感じましたか？ *

Strongly agree 強くそう感じた

Agree そう感じた

Neither agree nor disagree どちらとも言えない

Disagree あまりそう感じなかった

Strongly disagree そう感じなかった

Question 4. Doing the speaking tasks and self-assessment caused me to revise the class content and textbook. 質問4：Speaking Task（スピーキング課題）と Self-assessment（自己評価）は、クラスや教科書の内容を復習するのに役立ちましたか？ *

Strongly agree とても役立った

Agree 役立った

Neither agree nor disagree どちらとも言えない

Disagree あまり役立たなかった

Strongly Disagree 全く役立たなかった

Question 5. Doing the speaking tasks and self-assessment helped me to remember grammar, vocabulary, phrases and pronunciation advice learned in class. 質問5：Speaking Task（スピーキング課題）と Self-assessment（自己評価）は、クラス内で学んだ英語の文法・語彙・表現・発音を暗記するのに役立ちましたか？ *

Strongly agree とても役立った

Agree 役立った

Recorded Speaking Tasks as Formative Assessment in Intermediate English Classes

Neither agree nor disagree どちらとも言えない

Disagree あまり役立たなかった

Strongly Disagree 全く役立たなかった

Question 6. Doing the speaking tasks and self-assessment helped me to understand my strengths and weaknesses in English. 質問6：Speaking Task（スピーキング課題）と Self-assessment（自己評価）は、自分の英語の得意な部分と苦手な部分を理解するのに役立ちましたか？ *

Strongly agree とても役立った

Agree 役立った

Neither agree nor disagree どちらとも言えない

Disagree あまり役立たなかった

Strongly Disagree 全く役立たなかった

Question 7. Overall, doing the speaking tasks and self-assessment helped me to improve my English. 質問7：Speaking Task（スピーキング課題）と Self-assessment（自己評価）は、英語の上達に役立ちましたか？ *

Strongly agree とても役立った

Agree 役立った

Neither agree nor disagree どちらとも言えない

Disagree あまり役立たなかった

Strongly Disagree 全く役立たなかった

Question 8. Doing the speaking tasks and self-assessment helped me to increase my confidence to speak English. 質問8：Speaking Task（スピーキング課題）と Self-assessment（自己評価）は、英語を話す自信の向上に役立ちましたか？ *

Strongly agree とても役立った

Agree 役立った

Neither agree nor disagree どちらとも言えない

Disagree あまり役立たなかった

Strongly Disagree 全く役立たなかった

Question 9. Doing the speaking tasks and self-assessment helped me to improve my grammar. 質問9：Speaking Task（スピーキング課題）と Self-assessment

(自己評価)は、文法力を伸ばすのに役立ちましたか？ *

Strongly agree とても役立った

Agree 役立った

Neither agree nor disagree どちらとも言えない

Disagree あまり役立たなかった

Strongly Disagree 全く役立たなかった

Question 10. Doing the speaking tasks and self-assessment helped me to improve my vocabulary. 質問10: Speaking Task (スピーキング課題)と Self-assessment (自己評価)は、語彙力を伸ばすのに役立ちましたか？ *

Strongly agree とても役立った

Agree 役立った

Neither agree nor disagree どちらとも言えない

Disagree あまり役立たなかった

Strongly Disagree 全く役立たなかった

Question 11. Doing the speaking tasks and self-assessment helped me to improve my fluency. 質問11: Speaking Task (スピーキング課題)と Self-assessment (自己評価)は、よりスムーズに英語を話すことに役立ちましたか？ *

Strongly agree とても役立った

Agree 役立った

Neither agree nor disagree どちらとも言えない

Disagree あまり役立たなかった

Strongly Disagree 全く役立たなかった

Question 12. Doing the speaking tasks and self-assessment helped me to improve my pronunciation. 質問12: Speaking Task (スピーキング課題)と Self-assessment (自己評価)は、発音の上達に役立ちましたか？ *

Strongly agree とても役立った

Agree 役立った

Neither agree nor disagree どちらとも言えない

Disagree あまり役立たなかった

Strongly Disagree 全く役立たなかった

Question 13. Through doing the speaking tasks and self-assessment, I became

Recorded Speaking Tasks as Formative Assessment in Intermediate English Classes

better at noticing mistakes in my English. 質問13: Speaking Task (スピーキング課題) と Self-assessment (自己評価) を通して, 自分の英語の間違いにより気づくことができるようになったと思いますか? *

Strongly Agree 強くそう思う

Agree そう思う

Neither agree nor disagree どちらとも言えない

Disagree あまりそう思わない

Strongly Disagree そう思わない

Question 14 At the end of the semester I could complete the speaking tasks more quickly than at the beginning of the semester. 質問14: 学期の始まりに比べると, 学期の終わりには自分が Speaking Task (スピーキング課題) を完成するスピードが早くなったと感じますか? *

Strongly Agree 強くそう感じる

Agree そう感じる

Neither agree nor disagree どちらとも言えない

Disagree あまりそう感じない

Strongly Disagree そう感じない

Question 15. The feedback given to me by the teacher was helpful. 質問15: 教師から各学生に渡されたフィードバックは役立ちましたか? *

Strongly agree とても役立った

Agree 役立った

Neither agree nor disagree どちらとも言えない

Disagree あまり役立たなかった

Strongly Disagree 全く役立たなかった

Question 16. In the future, I want to continue recording myself speaking English as part of my English study. 質問16: 英語の勉強方法の一つとして, これからも自分が話した英語の録音を続けていきたいと思いますか? *

Strongly Agree 強くそう思う

Agree そう思う

Neither agree nor disagree どちらとも言えない

Disagree あまりそう思わない

Strongly Disagree そう思わない

Question 17. Please write your feelings or opinions about the speaking tasks and self-assessment. 質問 17 : Speaking Task (スピーキング課題) と Self-assessment (自己評価) について, 感じたことや意見があれば, どのようなことでも良いので自由に書いてください。

Appendix B

Speaking Assessment Tasks

Speaking Task 1

Using animals for entertainment should be banned.

Do you agree or disagree with this statement?

Rules

- Say your opinion clearly.
- Support your opinion with three reasons and examples.
- Use signposting language (page 28) to organize your ideas.
- Speak for 1-2 minutes
- Record your voice and send it to me as an attachment (添付).
- Complete the self-assessment after the lesson.

You can:

- ✓ Write notes.
- ✓ Practice and record as many times as you want.
- ✓ Use your textbook.
- ✓ Use the Internet/ Dictionaries etc.

Speaking Task 2

Discussion Customs and Traditions

Record a discussion between two people. The discussion should be two min-

utes long.

- The people have opposite opinions.
- You will play both people.
- You should say “A” or “B” before you change character.
- Each person should speak at least two times.
- Use phrases from pages 44-47 to help you.
- Good luck !

Topic Choices:

- Big traditional wedding ceremonies are a waste of time and money.
- Elementary school students should not use social media.
- Shopping online is much better than going to real shops.
- Children have to participate in traditional festivals.

Speaking Task 3

Give a Presentation about a famous person from history.

Appendix B

Your presentation should be between one and two minutes.

- ✧ Use the outline on page 65 to help you
- ✧ Include some personal opinions not only facts
- ✧ Use synonyms to avoid repetition
- ✧ Speak with confidence
- ✧ Be careful about pronunciation of times, years and dates

Speaking Task 4

- ✧ Have a discussion about using smartphones while walking.
- ✧ Choose two characters from below and record a conversation between them.
- ✧ Incorporate personal experiences and opinions into the discussion.
- ✧ Propose ideas to fix the problem.

- ✧ The conversation should be about 2 minutes.

<p><u>Character A</u></p> <p>You are a nurse at a hospital. You have noticed that accidents caused by using smartphones while walking have increased. Pedestrians should not do this.</p>	<p><u>Character B</u></p> <p>You are student. You use your smartphone while walking all the time, even when you cross the road. You have never had an accident. You think it is OK.</p>
<p><u>Character C</u></p> <p>You are a parent who takes your children to school every morning. Every day, you see teenagers using smartphones while crossing the road. It should be banned.</p>	<p><u>Character D</u></p> <p>You are a business person. You are very busy and you often use your smartphone while walking. The government should punish people who look at their phones while crossing the road, but not people who are just walking.</p>

Speaking Task 5

You are members of a city council (市議会). The city has decided to develop a large piece of land. There are two different plans. Some members of the city council want to build a new shopping mall. Others prefer to build a park.

- ✧ Have a discussion about the two plans.
- ✧ Choose two characters from below and record a conversation between them.
- ✧ Talk about the advantages and disadvantages of these two plans.
- ✧ Use the phrases and vocabulary from pages 102 and 103 to organize your ideas.
- ✧ The conversation should be about 2 minutes.

Recorded Speaking Tasks as Formative Assessment in Intermediate English Classes

<u>Character A</u> You are a parent. You want your children to spend more time outdoors and be active.	<u>Character B</u> You own a small convenience store in the city and you are worried that a new shopping mall will be bad for your business.
<u>Character C</u> You are a university student. You think the mall will be a good place for young people to hang out and also be a good place for students to get a part time job.	<u>Character D</u> You are the owner of a fast food chain. You want to open a new restaurant in the mall and expand your business.

Recorded Speaking Tasks as Formative Assessment in Intermediate English Classes

Adrian WAGNER

This paper reports upon an action research project to incorporate formative assessment into two intermediate level English oral communication classes for second year students at a university in Japan. Recorded speaking tasks were incorporated into the class at regular intervals throughout the semester. These tasks were closely based on the class and textbook content. After completing each task students re-listened to their submissions and completed a self-assessment report. Then, the teacher would listen to all submissions and provide a detailed feedback form identifying errors and giving personalised feedback. At the end of the semester, students completed a questionnaire designed to measure their reactions to and perceptions of learning outcomes of this assessment style. Overall, the students reported that doing the challenging speaking tasks, analysing their own language production, and receiving personalised feedback was a valuable learning experience. The results suggest that formative assessment is a valid tool for language teachers who want their assessment tasks to not merely provide justification for a final grade but also to motivate students and enhance learning outcomes.