



Research Article

Learning Japanese Using Sail App at A British University

Junko Winch^{1*}, Ayumi Asabe², Shingo Ono²

¹Department of Language Studies, School of Media, Arts and Humanities, University of Sussex, Falmer, UK

²Helte Co Ltd, Chiba, Japan

*Correspondence to: Junko Winch, PhD, Lecturer, Department of Language Studies, School of Media, Arts and Humanities, University of Sussex, Falmer, Brighton BN1 9RH, UK; E-mail: J.Winch@sussex.ac.uk

Received: June 30, 2023 Revised: July 21, 2023 Accepted: August 17, 2023 Published: September 25, 2023

Abstract

Objective: This study investigates if the Sail App Japanese speaking at least once a week for 25 minutes with native speakers enhance their speaking and listening ability in three months.

Methods: Fourteen undergraduate students (Ab Initio, Intermediate and Advanced Japanese level) who studied Japanese language at University of Sussex in the 2020/21 used Sail App to practice speaking Japanese with native Japanese speakers who were registered Sail App users living in Japan. Sussex students were asked to have a total of 6 speaking and listening skills evaluations by 5 Japanese native speakers who were also Helte Co Ltd staff. The 5 Japanese native speakers evaluated 14 students on 4 occasions (the initial baseline evaluation, the second month, and the third month and final evaluation). The students also self-evaluated of their listening and speaking skills on 2 occasions (the initial baseline evaluation and the final evaluation).

Results: After three months, only 1 student completed all 6 evaluations and 3 students completed 4 native speakers' evaluations by Helte Co Ltd staff. The majority of students' speaking ability remained unchanged. However, the perception discrepancies of speaking and listening skills between Sussex students and Japanese native speakers were found.

Conclusion: The majority of Ab Initio and Intermediate level students had a tendency to evaluate their listening and speaking skills higher than those by native Japanese speakers. As their level of language went higher, this discrepancy became smaller.

Keywords: higher education, Japanese language learning, speaking and listening, Sail App

Citation: Winch J, Asabe A, Ono S. Learning Japanese Using Sail App at A British University. *J Mod Educ Res*, 2023; 2: 9. DOI: 10.53964/jmer.2023009.

1 INTRODUCTION

1.1 Background of Japanese Language Teaching and Learning in the UK

Approximately three million people around the globe

study Japanese language^[1]. The number of undergraduate students who study Japanese at British universities is also increasing. According to the Association of University Language Communities (AULC) in the UK and Ireland

survey, a steady increase and uptake in Japanese language had been reported in the four previous years^[2]. AULC 2019 survey reports that Japanese language is the fourth most popular language in the UK after Spanish, French, and German^[3]. At University of Sussex, where this study was conducted, an increase of students who took Japanese was observed between the 2018/2019 and 2020/2021 academic years. Nearly 90 students have been registered to study Japanese in the academic years 2021/22 and 2022/23. Students study Japanese for various reasons ranging from wishing to work, study abroad or live in Japan, manga, anime, Japanese culture, Japanese singers / songs, Hello Kitty, interests in language studies etc.

1.2 Issue Under Consideration

Japanese language teaching and learning at British universities has limitations in enhancing students' speaking abilities within the formal university class hours. Especially the challenges for students who are studying Japanese at British universities, compared to those studying at Japanese universities, are limited opportunities to practise speaking Japanese. Students usually become very familiar with only their language tutor's Japanese pronunciation and speaking, which is not expected learning outcome, especially if they study Intermediate and Advanced level. As for students' listening skills and pronunciation studying at British universities who frequently watch Japanese anime and YouTube programmes, they are surprisingly good as they have been watching English subtitled Japanese TV programmes from their young age, which provides similar environment to how Japanese children would be brought up. Due to lack of speaking opportunities, some keen students pay Japanese language online speaking services in order to practise their Japanese. Despite this, their Japanese online speaking practice may not work or their self-study does not last. Some possible causes may include, firstly, students' expectation to improve their language skills is very short term. Secondly, these online study or language speaking app learning is currently considered as 'students' self-study' or 'independent learning' outside their formal class hours by both students and language teachers. Thirdly, the majority of language learning apps do not monitor learners' language skill progress.

1.3 The Sail App

The Sail app is an online Japanese speaking service which allows non-native Japanese speakers who wish to practise speaking Japanese for 25 minutes on one-to-one basis on-line video conversations in Japanese through mobile phone, computer, or tablet with Japanese native speakers residing in Japan. Students are asked to register, install the app and make an appointment with their registered Japanese native speakers. However, this Sail App Japanese learning is not part of the University's language class hours and students are asked to practise speaking Japanese at least once a week outside their formal

university class hours. Unlike other language learning apps, students' speaking progress and frequencies are monitored regularly by the Helte Co. Ltd staff who offer this service for free. To my knowledge, Sussex university is the first and the only university in the UK to use this Japanese language speaking Sail app, although some Japanese and American universities have already included this type of Japanese speaking learning as a part of their course requirements.

1.4 Research Questions (RQs)

The proposed study has a total of 4 RQs related to language skills and a non-language skill:

1) Language skills (listening and speaking)

RQ1. Were the students able to continue using Sail App in three months?

Japanese language learners use various mobile apps and websites. It is common that students' mobile app Japanese learning is not included in the university formal class hours. However, this presents a limitation to students' self-study as they are in addition to their formal university Japanese learning hours and these self-studies do not usually last.

RQ2. Are students' speaking and listening skills enhanced by the Sail App in three months?

RQ3. Are there any speaking and listening skills perception gaps between the native Japanese speakers and the undergraduate students?

2) Interests to Japan

RQ4. Did students increase their interests to visit Japan by using Sail App?

2 JAPANESE LANGUAGE TEACHING STANDARDS, PURPOSES AND LIMITATION

2.1 CEFR and JF Standard

Two Japanese language standards exist in Japanese language teaching, The Common European Framework of Reference for Languages (CEFR) and JF Standards for Japanese Language Education (JF Standard). CEFR "is now accepted as the international standard for language teaching and learning"^[4] not only English language teaching but also all languages including Japanese. CEFR may also serve as a quality assurance role in language education.

JF Standard, on the other hand, was developed by Japan Foundation (JF), which was established in 1972 by the Ministry of Foreign Affairs^[1]. JF Standard was modelled on the proficiency level-based can-do statements of the CEFR^[5] and modified to fit the Japanese context^[6].

Having two standards in Japanese language confuses the Japanese language teaching professionals on which rubrics to follow, CEFR or JF standard, as JF Standard is not identical to its European counterpart^[5]. Bučar et

al.^[7] rightly point out that JF Standard was developed partially in response to the CEFR^[4]. Ijichi^[8] also claims JF Standard as “JF version of the CEFR (p.444)”, but JF Standards may also be considered as an equivalent to CEFR-J in English language teaching context, details of which will be explained in section 2.3.

A series of Marugoto textbooks which JF has launched in 2010 may be a significant milestone in the Japanese language teaching as Marugoto is based on the JF Standard^[7]. Marugoto not only alleviated the double-standard confusion among Japanese language teachers, but also encouraged the Japanese language teachers to follow JF standard rather than CEFR on behalf of JF.

Furthermore, Japanese Language Proficiency Test (JLPT) is not congruent with CEFR standard^[9], either. JLPT is an internationally recognised qualification sponsored by JF and is the largest-scale JLPT in the world^[7] since 1984. However, JLPT’s focus is on reading, writing and listening skills in five levels (N1 to N5) except for speaking skills. According to Bučar et al.^[7], this reason is “due to the very large number of examinees in numerous locations around the world, the test does not include oral interviews, essay tasks, or other productive tasks (p.459)”. Although JLPT is claimed as “a reliable means of evaluating and certifying the Japanese proficiency of non-native speakers” by JF^[5], JLPT currently does not test speaking skills. In other words, the claimed reliability of JLPT language proficiency may be only valid for reading, writing and listening skills and we cannot assume and transfer students’ speaking ability to align with their reading, writing and listening skills. In any case, the impact which JLPT does not include speaking tests is significant in Japanese language teaching and learning the majority of native Japanese language teachers and non-native Japanese learners focus their teaching and learning on grammar, listening, reading and writing skills. This point may also relate to JF’s encouragement to Japanese language teachers to follow JF Standard rather than CEFR.

2.2 Purpose of CEFR

The purposes of CEFR are discussed by the stakeholders of CEFR, as each stakeholder’s purpose usually differs but may also overlap. CEFR stakeholders may include a) language learners, b) language teachers, c) directors of language course programme, d) language policy makers and institutions which offer to teach languages. Among these stakeholders, the following three parties are chosen as a focus of our discussion in this paper, i.e., a) language learners, b) language teachers and c) directors of the language course programme.

As for a) language learners, they usually set out with their expectations when they start learning a language, i.e. to be able to speak the language. If they perceive that

they made good progress in their language skills, this increases their self-confidence and they are more likely to continue learning the language. In order for language learners to perceive their positive language learning experience, they usually: (1) self-assess / diagnose the language skills before learning; and (2) monitor their progress throughout their language learning to perceive whether they could meet their expectations or achieve the proficiency level they anticipated at the beginning. From this language learners’ perspectives, the following two purposes of CEFR may exist: (1) to help learners to self-assess / diagnose their own language skills; and (2) to help language learners as a guide to monitor their progress.

The main purposes of CEFR may be for b) language teachers and their teaching activities. The purposes of language teachers’ CEFR include the following four: (1) assess individual students as well as overall and average students’ language proficiency level; (2) design their teaching syllabus; (3) plan and design their assessment; and (4) diagnose, assign and recommend learners to a suitable language level course.

As for c) directors of the language course programme, following five management related purposes of CEFR may be considered: (1) course accountability; (2) course reputation; (3) quality assurance; (4) management of all different language staff and standardise the language course level; (5) development of their overall language course curriculum. As for the (1) course accountability, accountability links to attractive concepts such as ‘quality’, ‘autonomy’ and ‘integrity’, all of which gives special attention and positive connotation^[10] regarding language education. As for the second point, (2) course reputation, some higher educational institutions assume that CEFR provides them “cutting-edge and globally conscious curricula”^[5]. Reputation also includes league tables and competition with other similar higher educational institutions to demonstrate their innovativeness. With regards to (3) quality assurance, quality assurance is credence properties^[11]. Credence properties are generally found in goods and services that are characterised with high level of complexity including teaching and learning process^[12], which may be realised when language learners graduate. Through British Institution-Wide Language Programme (IWLP), non-language major undergraduate students can have a degree with a language with specific CEFR proficiency in graduation (quality assurance) if a student completed studying a language for a certain number of hours. Students are entitled to add their degree descriptions linked to CEFR in their CVs. At University of Sussex, for instance, a student who majored in Biology and studied Japanese for three years through IWLP, he is awarded with “BSc Biology with Japanese CEFR B1+”. This is academic delay of gratification^[13], which is explained

as “postpone of immediately available opportunities to satisfy impulse in favour of pursuing a chosen important academic rewards or goals that are temporally remote but ostensibly more valuable”^[14]. If students know the usefulness of the language and the language course programme, they are more willing to delay gratification^[13]. Director of language course programme also find CEFR very useful to manage and to standardise all their language courses they offer in each language.

To review the purpose of CEFR, each stakeholder’s purposes are different but some of the purposes also overlap, which indicates importance of standard / CEFR in language studies.

2.3 Limitations of CEFR

The majority of the criticisms of CEFR in the literature is on its descriptors and its cultural challenge, which is the focus of our discussion in this section. Comparing criticism on CEFR between English language teaching and Japanese language teaching allows to identify whether limitations are consistent between English and Japanese language teaching.

2.3.1 Criticism on CEFR Descriptors

CEFR has six consecutive levels of language proficiency, ranging from A1 to C2 with 53 illustrative scales, listing language-independent descriptions of each proficiency level for a given skill or ability. In English language teaching context, these illustrative scales have become the most influential and the most heavily criticised aspect of the CEFR^[15,16] for their vagueness and inconsistencies, both within and across levels^[17,18] leaving room for dissimilar interpretations. Alderson^[17] claims CEFR descriptors ‘gaps and flaws’^[16]. More specific content-related criticism includes that the levels are not equidistant because they contain overlaps and inconsistencies^[19] and on the wording of the level descriptors^[17]. Weir^[20] maintains that “CEFR is not sufficiently comprehensive, coherent or transparent for uncritical use in language testing”^[20].

Similar criticisms are discussed in a Japanese language teaching context on the CEFR’s descriptors: CEFR’s descriptors cannot be accommodated within their existing curriculum and the cohort of students as they are^[5,7]. The wording of the level descriptors are also criticised as follows: “CEFR’s descriptors are not language specific and thus are more abstract”^[21].

From the comparison of the above criticism on CEFR descriptors, it may be possible to say that the criticisms from both English and Japanese language teaching are similar and consistent.

2.3.2 CEFR’s Cultural Challenge

Anderson^[17] gives two reasons why CEFR has cultural

challenges in English language teaching: (1) ‘being developed for the major European languages’^[17]; (2) ‘having a framework we could claim was genuinely European (despite in large part being based on an idea from the Anglo-American world of Communicative Language Teaching (CLT) and applied linguistics’^[17].

CEFR’s cultural incompatibility from non-European countries (especially Confucius countries) is reminiscent of CLT when CLT was introduced to these countries. CLT was also heavily criticised as culturally ‘conflict and incompatible’^[21].

In English language teaching context, adopting CEFR as the universal framework seems to present challenges. Anderson^[17] claims “as the experience with the use of CEFR grows, more and more problems arise” (p.661). Negishi^[22] maintains the challenges in Japan and how CEFR-J was developed as follows:

It is possible to use the CEFR descriptors for English language teaching in Japan, but that we need some modification. The fact that the population of Japanese EFL learners skews towards the lower levels suggests the need for the branching of the CEFR A1-B2 levels. The development process of the CEFR-J, a modified version of the original CEFR for English language teaching in Japan.

South Korea also tried to accommodate CEFR but found it difficult to adopt as a current form^[23,24]. Other Asian countries experienced similar cultural challenges and developed their own modified CEFR version including Malaysia’s CEFR-M, Vietnam’s CEFR-V, China’s Standard of English.

Shifting our focus to Japanese language teaching context, again, similar discussions occur. Mori and Mori^[25] claim that “ideas based on the instruction of commonly taught European languages cannot be easily or speedily adapted to Japanese Foreign Language contexts.” Bučar et al.^[7] also point out that descriptors used in the CEFR self-assessment grid are, thus, difficult to apply due to Japanese scripts (i.e., three writing systems), Japanese language specific characteristics (i.e., politeness) and cultural reasons^[5,7]. It is possible to say that the CEFR cultural incompatibility criticism are consistent in both English and Japanese language teaching context.

3 METHODOLOGY

This section discusses the overview of this study, RQs and instruments, participants, ethical issues, data collection and data analysis.

3.1 The Overview of This Study

The below [Figure 1](#) visually summarises how we proceed this study.

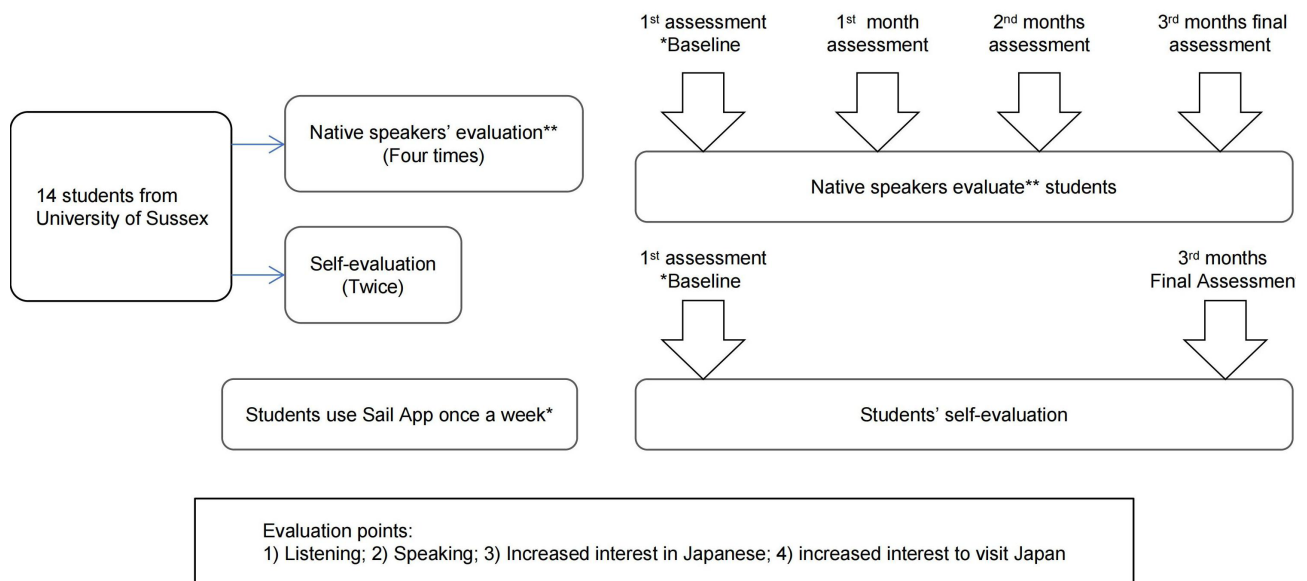


Figure 1. Overview of this study. *Students will have conversation practice at least once a week using the Sail App. There is no upper limit to the number of times students can use Sail App; **Four Japanese native speakers evaluated the students on a one-to-one basis for 25 minutes for three months. These Japanese native speakers received evaluation training in advance to ensure equal standards, consistency and fairness.

3.2 RQs and Instruments

RQ1 asks whether the students are able to continue for three months. A member of Helte Co Ltd monitors and keeps record of 14 students which provides the quantitative data.

RQ2 asks if students' speaking and listening skills are enhanced through the Sail App. Five members of Helte Co Ltd evaluate 14 students' speaking and listening skills on four occasions based on our own evaluation framework / rubrics (Appendix), which provides quantitative data regarding students' language skills. This evaluation framework / rubric (Appendix) is collaborative and innovative framework adapted from CEFR and JF Standard.

RQ3 asks if there are any perception gaps on their speaking and listening skills between the 5 Japanese native speakers from Helte Co Ltd and 14 undergraduate students. To collect this quantitative data, two types of evaluations are used: (1) evaluations by 5 native Japanese speakers and (2) Sussex students' self-evaluation. Both evaluations [evaluations (1) and (2)] use the same evaluation framework / rubrics (i.e., Appendix). However, evaluation (1) does not evaluate '3. Interest question' and '4. Cultural questions' in the Appendix.

RQ4 asks whether students increased their interests to visit Japan by using Sail App. This quantitative data is collected using students' self-evaluation on 3. Interest question' and '4. Cultural questions'.

3.3 Participants

A total of 19 participants who consisted of 14 undergraduate students (ranging from 19 to 22 years old and studied Japanese language from Ab Initio to Advance at University of Sussex in the 2020/21 academic year) and 5 native Japanese speakers from Helte Co. Ltd., who evaluated the 14 undergraduate students. University of Sussex offers three levels of Japanese: Ab Initio, Intermediate and Advanced. Ab Initio is for complete Japanese beginners. Students study Japanese for four hours a week for 11 weeks. These students took Japanese languages as a part of their degree and they study their own major through IWLP. Fourteen students' Japanese language levels are summarised in Table 1.

3.4 Ethical Issues

The researcher sought permission to conduct this study and was granted permission from the Director of the Language Centre in the School of Media, Arts and Humanities in February 2021. This research was also obtained from the Social Sciences & Arts Cross-Schools Research Ethics Committee. The ethical review application number of this research is ER/JW666/1. Informed consent was not required as this was a three-months pilot study. The researcher informed all students regarding: (1) purpose of procedure of the research; (2) a guarantee of anonymity and confidentiality of records; (3) students' participation is voluntary and that they may withdraw at any time (they can join other Japanese classes); and (4) an offer to provide a summary of finding. Students have had the opportunity to ask questions about this research.

Table 1. The Level of Participants' Japanese

Ab Initio	Intermediate	Advanced	Total
7 students	4 students	3 students	14 students

In the process of collection and analysis, nobody other than the researcher had access to the data. Small scale research including this study is vulnerable to this threat. However, identifying individual's names and their details were unnecessary for the study and no information about each individual was identified. Therefore, the students' confidentiality was protected in this study.

3.5 Data Collection

Data consisted of two quantitative data [(1) one-to-one native speakers' evaluation and (2) self-evaluation] and students' voluntary comment which formed as qualitative data. For example, the first topic 'Speech level' is under 'Listening skill' in the [Appendix](#). This is evaluated by both native Japanese speaker [i.e., (1) one-to-one native speakers' evaluation] and the students themselves [i.e., (2) self-evaluation]. The native speaker makes daily conversation with a student on one-to-one basis and evaluates the student's speech level from the following five options:

- 1: Can understand if spoken slowly, clearly and carefully articulated with long pauses.
- 2: Can understand if spoken slowly, clearly and carefully.
- 3: Can understand with slower than standard speed.
- 4: Can understand with near standard speed.
- 5: Can understand with standard speed.

Similarly, students also self-evaluate their speech level and choose the most suitable answer from the above five options. This allows not only to collect the data (quantitative ratings from 1 to 5) from both native Japanese speakers and students but also to identify the perception discrepancies between the native speakers and students by comparing the same data.

3.5.1 One-to-One Native Speakers' Evaluation

A total of 4 one-to-one evaluation results (i.e., baseline assessment in the first week of May 2021, 1st month assessment in the last week of May 2021, 2nd month assessment in the last week of June 2021 and the 3rd assessment in the last week of July 2021) per student were collected between 1st May 2021 and 31st July 2021. Four native Japanese evaluators carried out the students' speaking and listening evaluations in 25 minutes one-to-one evaluation which provided quantitative data using this study's language evaluation framework ([Appendix](#)'s '1. Listening' and '2 Speaking' only). These evaluators received training in advance to ensure standards, consistency and fairness. The five native speakers' evaluations were recorded. The results

of students who completed all four native speakers' evaluations are used to analyse RQ2.

3.5.2 Students' Self-Evaluation

Although it is normal expectation in language teaching that native speakers assess speaking and listening skills and evaluate non-native speakers' language skills, the researcher also find it important that students also monitor their progress using this study's language evaluation framework ([Appendix](#)). Students were asked to self-evaluate their speaking and listening skills as well as interest and culture questions on two occasions (in the first week of May 2021 and in the last week of July 2021), which enables to record students' self-perceived ratings into quantitative data. Students' comments were also asked if they wish to provide.

3.6 Data Analysis

All students data was entered into Excel sheet by a member of Helte Co. Ltd and manually analysed by the researcher.

3.6.1 Speaking and Listening (RQ1-RQ3)

To analyse data related to RQ1, all students' evaluation dates were recorded by the number of completed evaluations (i.e., completed all evaluations, completed three evaluations, completed two evaluations, completed one evaluation). Then, this information was organised by the three language levels (Ab Initio, Intermediate and Advanced). Only the results of students who completed all 4 evaluations with the native speakers were used to answer RQ2 (whether students enhanced their speaking and listening skills within three months) are used as it compares the results and find any improvement between the start and the end of the study.

To analyse RQ3 (any discrepancies perceived in evaluation between the non-Japanese undergraduate students and native Japanese speakers), however, all students who participated evaluations were used to calculate the average regardless of their number of evaluations they completed (as it does not compare results find and any differences between the start and the end of the study). Then, they were organised by their three language levels by Ab Initio, Intermediate and Advanced. This allowed to identify any perceived discrepancies by three levels.

3.6.2 Interests in Japan (RQ4)

Only the results of students who completed all self-evaluations ([Appendix](#)'s '3. Interest question' and '4. Cultural questions') are analysed as it compares results find and any differences between the start and the end of the study. The questions asked if the use of Sail App a) may affect to increase students' interests in Japanese society and culture and b) increased tolerance for other cultures are used.

4 RESULTS AND DISCUSSION

4.1 RQ1 (Were the Students Able to Continue Using Sail App for Three Months?)

Table 2 summarises the number of participants at the start and the end of this study. After 1 month, two students informed the researcher that they wished to withdraw from the research. The students were asked have a total of six evaluations: twice (the first assessment and after three months) by student themselves (i.e., self-evaluation) and four times (the first diagnostic assessment, the first month progress check, the second month progress check and the final third months assessment) with the native Japanese speakers.

Only 1 person completed both self-evaluations and a native speaker's evaluation for three-month period. The number of students who completed this research were drastically decreased as the participation to this study was voluntary and students could easily withdraw at any time. The poor completion rate may be also affected by the timing (i.e., the University summer break) when this study was conducted. The three assessments except for the basement were held after students finished their University Spring term.

Table 3 summarises the details on '(2) Native speakers' evaluation' in Table 2. It summarises the number of students who completed one, two, three and four evaluations by native speakers and their language levels.

Five students completed one native speaker's evaluation. This is one third of all participants. Students must have faced their language skills reality by speaking with the native speakers and also realised their perception gap - how students were optimistic about their language skills. Perhaps, some students may have discouraged or demotivated to continue the native speakers' evaluations.

Students' persistence as well as their language ability may be important factors for students to complete native speakers' evaluations, as the 3 out of 14 students completed four native speakers' evaluations are either Intermediate or Advanced students. The result of next section is based on these 3 students who completed all four native evaluations.

4.2 RQ2 (Is Speaking and Listening Skill Enhanced Through Sail App?)

With regarding to 3 students who completed all 4 evaluations, it may be difficult to expect students to enhance their listening and speaking skills in three months. Table 4 shows two results: (1) Two Intermediate students showed hardly any improvement on listening and speaking; and (2) The Advanced student result shows

that some of his speaking and listening skills deteriorated compared to when he started.

However, the results confirmed that the higher students' Japanese level was, the higher their quantitative ratings were. The Advanced student received overall higher quantitative ratings by native evaluator than those of two Intermediate students. There were three Advanced level students, all of whom had higher quantitative evaluation results (Advanced level students received quantitative rating of between 2 and 4 by native speakers' evaluations) compared with those of Ab Initio and Intermediate level students (Intermediate and Ab Initio level students received quantitative rating of between 1 and 2 by native speakers' evaluations).

4.3 RQ3 (Are There Any Perception Gaps on Their Speaking and Listening Skills Between the Native Speakers and Non-native Japanese Learners?)

Table 5 summarises the average evaluation rating of students' and native Japanese speakers' listening and speaking skill. This allows to compare and identify the discrepancies of native and non-native speakers' perception of speaking and listening. Students appeared to evaluate their language abilities higher than those by native speakers'. These discrepancies became wider among the Ab Initio than Advanced level students.

As their level of language became higher, the discrepancy became lower. In other words, students' perceptions may become closer and more realistic to the native speakers' language evaluation results as the higher their language levels becomes. Below is further details on each level.

4.3.1 Ab Initio Level

The differences between native speakers' and students' ratings are observed among Ab Initio Japanese level (Table 6) which indicates that perception discrepancies (listening 1.6 and speaking 1.6) is large. It should be also added that 2 out of 7 Ab Initio level students self-evaluated their listening and speaking skills 4 out of 5 which indicated their confidence in their language skills.

4.3.2 Intermediate Level

The difference between native speakers' and students' ratings are observed among Intermediate Japanese level (Table 7) and the discrepancies (listening 1.1 and speaking 1.1) of perception are slightly smaller than that of Ab Initio (listening 1.6 and speaking 1.6) but over one rating / rubric / descriptor difference is still large.

4.3.3 Advanced Level

The discrepancies between native and non-native speakers' ratings and this indicates discrepancies (listening 0 and speaking 0.7) of their perception is closest compared

Table 2. The Number of Students Who Completed Two Evaluations

	The Start of the Study When Student Registered	The End of Study Who Completed All Evaluations	Student Who both Completed Self-Evaluations and Native Evaluations
Self-evaluation	14 students	2 students	1 student
Native speakers' evaluations	14 students	3 students	

Table 3. The Details of Students Who Completed Native Speakers' Evaluations

Completed All 4 Evaluations	Completed 3 Evaluations	Completed 2 Evaluations	Completed 1 Evaluation	Completed None	Total
3 (2 Intermediate and 1 Advanced)	3 (1 Ab Initio, 1 Intermediate, 1 Advanced)	1 (1 Ab Initio)	5 (4 Ab Initio, 1 Advanced)	2 (1 Ab Initio, 1 Intermediate)	14

Table 4. Students' Language Skills Before and After Evaluated by Native Japanese Speakers

	Intermediate Student 1	Intermediate Student 2	Advanced Student
Listening speech speed rating	2→2	2→3	4→3
Listening understanding rating	1→1	1→1	4→3
Speaking communication rating	1→1	1→1	3→3
Speaking grammatical accuracy rating	1→1	1→1	2→2
Speaking vocabulary range rating	1→1	2→2	3→2

Notes: Rating 1 is the lowest and rating 5 is the highest.

Table 5. Rating of Evaluation of Students' and Native Speakers' Listening and Speaking Skills

	Ab Initio Level		Intermediate Level		Advanced Level	
	Students' self-evaluation	Native speakers' evaluation	Students' self-evaluation	Native speakers' evaluation	Students' self-evaluation	Native speakers' evaluation
Average listening rating*	3.0	1.4	3.1	2.0	3.7	3.7
Average speaking rating**	2.7	1.1	3.6	1.5	3.5	2.8

Notes: * 'average listening rating' is the average of 'speech speed' and 'understanding'; **'average speaking rating' is the average of 'communication', 'grammatical accuracy' and 'vocabulary range'.

Table 6. The Discrepancies between Native and Non-native Speakers' Evaluations of Speaking and Listening Skills in Ab Initio Level

	Students' Self-Evaluation	Native Speakers' Evaluation
Listening skills	3.0 so we look at rubric 3 which says 'I can follow if the speech was slightly slower'	1.4 so we look at rubric 1 which says 'students can follow speech which is very slow and carefully articulated with long pauses for them to assimilate'
Speaking skills	2.7 so we look at rubric 3 which says 'I can express personal opinions and exchange information on topics that are familiar and of personal interest'	1.1 so we look at rubric 1 which says 'students can ask and answer simple questions'

Table 7. The Discrepancies between Native and Non-Native Speakers' Evaluations of Speaking and Listening Skills in Intermediate Level

	Students' Self-Evaluation	Native Speakers' Evaluation
Listening speech speed rating	3.1 so look at rubric 3 which says 'I can follow if the speech was slightly slower'	2 which is 'students can follow speech if the speech was slow'
Speaking communication rating	3.6 so we look at rubric 4 which says 'I can communicate on major abstract topic'	1.5 so we look at rubric 2 which says 'students can communicate direct and simple exchange of information on familiar topic'

Table 8. The Discrepancies between Native and Non-Native Speakers' Evaluations of Speaking and Listening Skills in Advanced Level

	Students' Self-Evaluation	Native Speakers' Evaluation
Listening speech speed rating	3.7 so look at rubric 4 which says 'I can follow close to the standard native speaker'	3.7 so look at rubric 4 which says 'students can follow close to the standard native speaker'
Speaking communication rating	3.5 so look at rubric 4 which says 'I can communicate on major abstract topic'	2.8 so look at rubric 3 which says 'students can express personal opinions and exchange information on topics that are familiar and of personal interest'

with Ab Initio and Intermediate levels (Table 8).

4.4 RQ4 (Did Students Increase Their Interests to Visit Japan by Using Sail App?)

Non-language questions include a) whether they increase their interest to visit Japan and continue studying Japanese by using Sail App, and b) whether they think they increased intercultural understanding including their own culture as well as other culture.

Unfortunately, only 2 students (one Ab Initio and one Advanced) completed their two self-evaluations and these results provided contrasting results. No students evaluated interest questions below rating 3, which indicates that students showed strong interest to visit Japan than language skill (speaking and listening).

Ab Initio student's cultural understanding remained unchanged before and after Sail App use except for the interest to visit Japan increased [from 'I want to go to Japan' (rating 3) to 'I want to go to Japan now' (rating 4).

On the other hand, the Advanced student's answer included a mixture of increased cultural understanding and decreased interest to go to Japan. The student's interest to visit Japan decreased from 'I want to go to Japan now and am preparing to go (rating 5) to 'I want to go to Japan' (rating 3). The interest to study Japanese also decreased from 'I will study harder to study Japanese combining with Sail App (rating 5) to 'I will study Japanese harder than before (rating 4).

However, this Advanced student's culture understanding through Sail App was enhanced as follows:

- (1) Understanding of other culture from 'recognise every culture has its own merits' to 'think beyond culture, no cultural boundaries';
- (2) Understanding of own culture from '80-89%' to '100%' and;
- (3) Obtained knowledge on Japanese society and culture from 'Knowledge has increased by participating in this research' to 'Knowledge has increased very much and I intend to increase further'.

This student's contrasting results may be considered as one student's case, which should not to be generalised due

to small sample size.

5 CONCLUSION, LIMITATION AND FUTERE STUDY

Reviewing RQs enables to conclude this paper. RQ1 asked whether students were able to continue speaking app learning within three months. The results showed that it was very difficult for students to keep using the app on regular basis for even for three months. Only 1 student out of 14 completed all 6 evaluations (both 2 self-evaluations and 4 native evaluations). Two students did not participate with the evaluations at all.

RQ2 asked if the students' speaking and listening skills were enhanced through Sail App. Among the 3 students completed native speaker's evaluations, 2 students' listening and speaking skills remained unchanged. One student's speaking and listening skills were deteriorated compared with when he started.

RQ3 asked if there are any perception gaps on their speaking and listening skills between the native Japanese speakers and students. The discrepancies of native and students' perception of the speaking and listening skills were quite large in the Ab Initio and Intermediate levels. The majority of undergraduate students tended to self-evaluate their speaking and listening skills higher compared to those from Japanese native speakers.

RQ4 asks if students increased their interest to visit Japan by using Sail App. Two students provided results, but one student's answer was contradicting - cultural understanding has enhanced while interest to visit Japan has decreased. These contrasting results may be considered as individual cases and should not to be generalised due to a very small sample size.

From the above results, it is clear that the main limitation of this study is the small sample size which may be resulted from the timing to conduct this study (the summer break). The completion rate was 21% and affected reliability and the generalisability of the conclusions drawn from the results. Nonetheless, these results provided information for a specific sample population. For future study, similar studies using large samples with longitudinal study may bring interesting outcome. Although enhanced listening and

speaking skills were not observed in this three-month study, longitudinal study may present different outcome.

6 IMPLICATIONS AND RECOMMENDATIONS

This study suggests that the evaluation perception discrepancies exist between native Japanese speakers and the undergraduate students who learn Japanese. As their level of Japanese language goes higher, the discrepancy became smaller. In other words, undergraduate students' perceptions of their language skills became closer to the native speakers' evaluation as their language level goes up. This result implies that if students wish to know their true language level from the native speakers' standard, students are recommended to lower their self-evaluation ratings which is likely to match with what the native speakers evaluate. However, there are some cases where Japanese beginner learners may face reality by speaking with native language speakers and lose their confidence which could be the reason for the majority of students not continuing the Sail App learning. In order to keep the beginner's confidence, it is recommended for beginner students to practice speaking with their peers or similar level of non-native language learners.

Lastly, the current academic movement places more emphasis on scholarship activities for language lecturers in addition to teaching responsibilities. It is hope that this paper may give insight to language lectures that collaborative language research with a business sector is also one of possible scholarship activities.

Acknowledgements

Not applicable.

Conflicts of Interest

The authors declared no conflict of interest.

Author Contribution

Winch J wrote, reviewed and amended this paper. Asabe A and Ono S contributed [Appendix](#) section and study design respectively, both read and approved this manuscript submission.

Abbreviation List

AULC, Association of University Language Communities
CEFR, Common European Framework of Reference for Languages

CLT, Communicative Language Teaching

IWLP, Institution-Wide Language Programme

JF Standard, JF Standards for Japanese Language Education

JF, Japan Foundation

JLPT, Japanese Language Proficiency Test

RQs, Research Questions

References

[1] Japan Foundation 50th Anniversary, Accessed 2023.

Available at: [\[Web\]](#)

- [2] Association of University Language Communities-UCML. AULC-UMLC survey of Institution-Wide Language Provision in universities in the UK (2019-2020), Accessed 2023. Available at: [\[Web\]](#)
- [3] Association of University Language Communities-UCML. AULC-UMLC survey of Institution-Wide Language Provision in universities in the UK (2018-2019), Accessed 2023. Available at: [\[Web\]](#)
- [4] North B, Ortega A, Sheehan S. Equals core inventory for general english. *Brit Coun*, Accessed 2023. Available at: [\[Web\]](#)
- [5] Lämsäalmi R. Conquering the world with "cutting-edge curricula": global citizens learning East Asian languages: 2012 "Van Schools tot Scriptie" Colloquium. Leiden, The Netherlands, 2012.
- [6] Hatasa Y, Watanabe T. Japanese as a Second Language Assessment in Japan: Current Issues and Future Directions. *Lang Assess Q*, 2017; 14: 192-212. [\[DOI\]](#)
- [7] Bučar CS, Ryu H, Škof NM et al. The CEFR and teaching Japanese as a foreign language. *Linguistica*, 2014; 54: 455-469. [\[DOI\]](#)
- [8] Ijichi N. The classroom and beyond: Creating a learning environment to support learners of Japanese at CEFR levels A2. 2 towards B1. *Lang Learn High Educ*, 2016; 6: 443-451. [\[DOI\]](#)
- [9] Mainichi Newspaper. Culture agency to develop Japanese language ability index based on int'l model (08/04/2019). Accessed 2023. Available at: [\[Web\]](#)
- [10] Bjorn S, Lee H. Accountability in higher education: Global perspectives on trust and power. 1st ed, Routledge: Abingdon, UK, 2011. [\[DOI\]](#)
- [11] Curran JM, Rosen DE. Student attitudes toward college courses: an examination of influences and intentions. *J Market Educ*, 2006; 28: 135-148. [\[DOI\]](#)
- [12] Ching G. A literature review on the student evaluation of teaching: An examination of the search, experience, and credence qualities of SET. *High Educ Eval Dev*, 2018; 12: 63-84. [\[DOI\]](#)
- [13] Bembenutty H. Teaching effectiveness, course evaluation, and academic performance: The role of academic delay of gratification. *J Adv Acad*, 2009; 20: 326-355. [\[DOI\]](#)
- [14] Bembenutty H, Karabenick SA. Academic delay of gratification. *Learn Individ Differ*, 1998; 10: 329-346. [\[DOI\]](#)
- [15] Figueras N. The impact of the CEFR. *Elt J*, 2012; 66: 477-485. [\[DOI\]](#)
- [16] Little D. The Common European Framework of Reference for Languages: Perspectives on the making of supranational language education policy. *Mod Lang J*, 2007; 91: 645-655. [\[DOI\]](#)
- [17] Alderson JC. The CEFR and the need for more research. *Mod Lang J*, 2007; 91: 695-663. [\[DOI\]](#)
- [18] Harsch C, Rupp AA. Designing and scaling level-specific writing tasks in alignment with the CEFR: A test-centered approach. *Lang Assess Q*, 2011; 8: 1-33. [\[DOI\]](#)
- [19] Deygers B, Zeidler B, Vilcu D et al. One framework to unite them all? Use of the CEFR in European university entrance

- policies. *Lang Assess Q*, 2018; 15: 3-15. [DOI]
- [20] Weir CJ. Limitations of the Common European Framework for developing comparable examinations and tests. *Lang Test*, 2005; 22: 281-300. [DOI]
- [21] Hu G. Potential cultural resistance to pedagogical imports: the case of communicative language teaching in China. *Lang Cult Curric*, 2002; 15: 93-105. [DOI]
- [22] Masashi N. The development of the CEFR-J: Where we are, where we are going, New Perspectives for Foreign Language Teaching in Higher Education: Exploring the Possibilities of Application of CEFR. Tokyo: Tokyo University of Foreign Studies, 2012: 10-116.
- [23] Markel S. The CEFR and English education in Japan. *J Policy Stud*, 2018; 56: 33-38.
- [24] Won Y. Common European framework of reference for language (CEFR) and test of proficiency in Korean (TOPIK). *Int J Area Stud*, 2016; 11: 39-58. [DOI]
- [25] Mori Y, Mori J. Review of recent research (2000-2010) on learning and instruction with specific reference to L2 Japanese. *Lang Teaching*, 2011; 44: 447-484. [DOI]