Technology Reports

| Japanese | Training | AI |

Special Articles on AI Supporting a Prosperous and Diverse Society

"Japanese Language Training AI" Supporting Japanese Conversation Training for Foreigners

Communication Device Development Department   Shin Oguri   Misa Tanaka[†]

The number of foreign workers in Japan has been increasing annually reaching 1.28 million in 2017. However, in conventional Japanese language education for foreigners, a gap has existed between the acquisition of "correct" reading/writing, pronunciation, and use of expressions on the one hand and practical Japanese conversation in real business situations on the other hand. Foreign workers are bewildered by this gap, which may hinder their work duties and even isolate them in the workplace. "Japanese Language Training AI" is a Japanese conversation training support service that was developed to solve these problems.

1. Introduction

In Japan, the number of foreign workers has been increasing annually reaching 1.28 million in 2017 [1]. Nevertheless, a shortage of human resources is still a problem in a variety of industries, so a bill was passed to revise the Immigration Control Act [2]. Enacted on April 1, 2019, this revision expands the range for which foreign workers having certain specialties and skills can be accepted, so the need for Japanese language education is expected to grow for both foreign workers and the companies accepting them.

As a result of conducting interviews with foreign workers, we found that there were some who had studied Japanese in their home countries before coming to Japan, passed the Japanese-Language Proficiency Test, and acquired a certain level of

†   Currently Solution Service Department

Japanese. However, due to the gap between Japanese studied at a Japanese language school or in textbooks and Japanese used in the workplace, they could not communicate well, which hindered their work and left them feeling isolated in the workplace. This situation led some to even return to their home countries.

To eliminate this gap between conventional Japanese language education and conversation in real business situations, NTT DOCOMO developed Japanese conversation training support service "Japanese Language Training AI" (hereinafter referred to as "JLT") as a departure from conventional language teaching materials centered about the memorization of example sentences. This service features a function that enables the user to freely create conversation that he or she would actually like to speak and practice with. It also judges whether that conversation is made up of appropriate "words and expressions" and offers advice as well (**Figure 1**).

This service was achieved by forming a cross-organizational joint team composed of NTT DOCOMO R&D departments and corporate sales and marketing departments and developed as a "TOPGUN" project that aims to solve social and business issues. To solve the problems in conventional Japanese language education that cannot necessarily be said to be practical, this service has undergone hypothesis testing through verification experiments and its app has been improved. This article describes the JLT service and its development as a TOPGUN project.
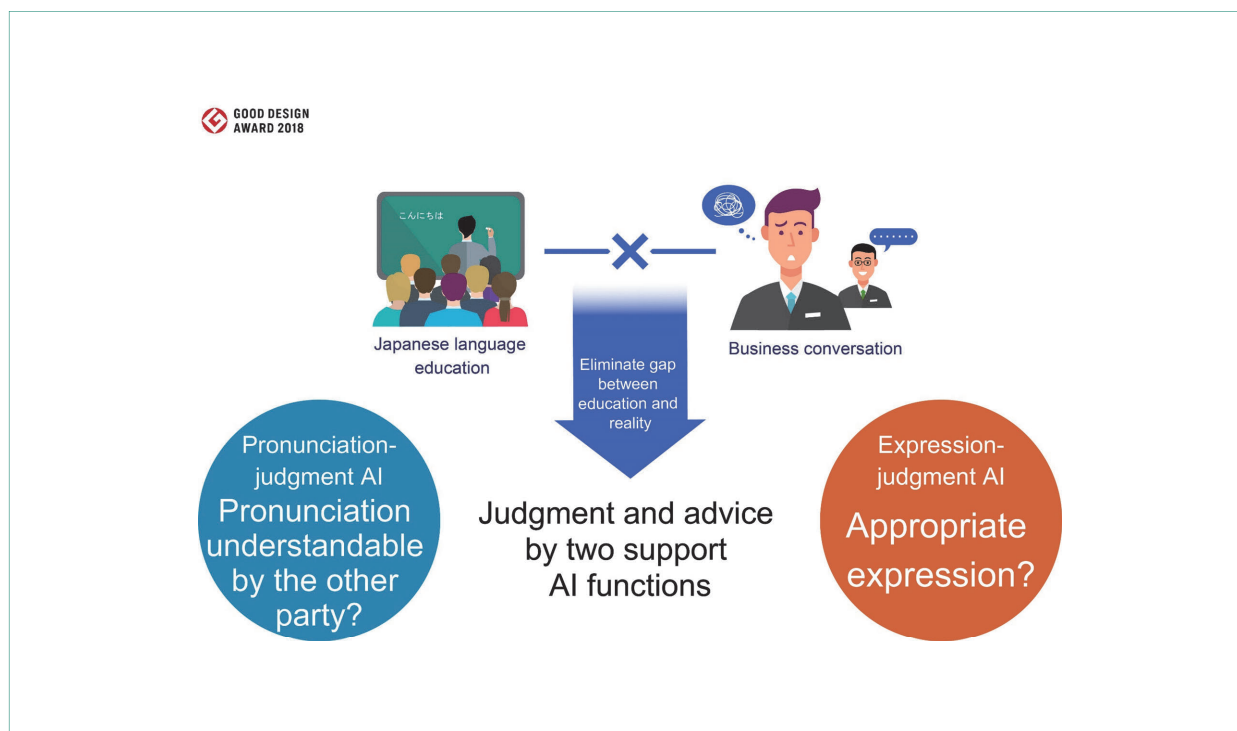

Figure 1   JLT features

2. JLT Overview

In contrast to conventional language teaching materials centered about the memorization of example sentences, JLT features NTT DOCOMO-developed AI functions (pronunciation judgment, expression judgment) that enable the user to freely create conversations that he or she would actually like to speak and to learn practical Japanese that can be understood by a Japanese native speaker.

The JLT service also provides training content useful in actual business situations for various fields and applications (dining, lodging, IT, retail sales, caregiving, and job-hunting activities) (**Figure 2**). Furthermore, to enable Japanese language training specific to the work of individual companies, NTT DOCOMO can provide a customer with customized training content.

The following describes the pronunciation-judgment function, expression-judgment function, and conversation-creation function.

2.1 Pronunciation-judgment AI

Pronunciation-judgment AI is a function that asks the user to read an example sentence in Japanese and judges not whether the result is correct Japanese pronunciation but whether it's pronunciation that can be understood by a Japanese native speaker. It also offers advice on improving pronunciation. For example, if the user mistakenly says "nimotsu wo *omochi* itashimasu" ("please let me carry your baggage") as "nimotsu wo *omachi* itashimasu," a Japanese native speaker would still understand the meaning. While a conventional Japanese conversation training support service would treat this as a mistake, the JLT service would judge this to be "GOOD" since it's a statement that could be understood while advising the user that "*omochi*" is correct (**Figure 3**).
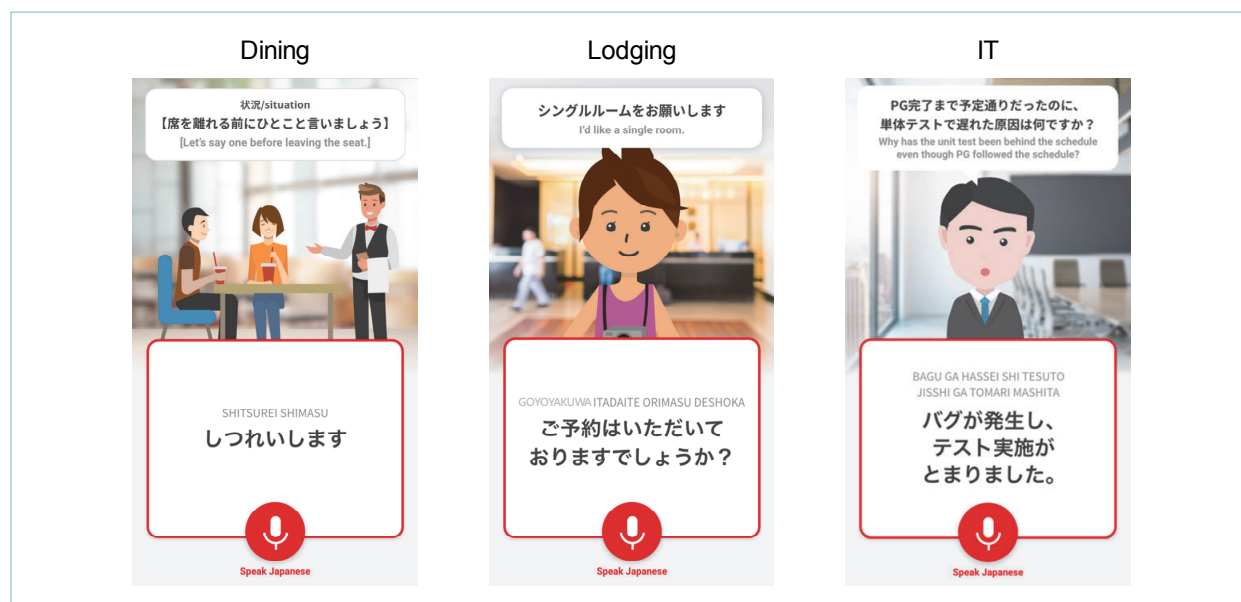

Figure 2　Training content by industry and application

2.2 Expression-judgment AI

Expression-judgment AI asks the trainee to speak in Japanese a phrase presented in the trainee's native language (English or Vietnamese at present) and judges whether the expression or wording used would be understandable to a Japanese native speaker while offering advice if needed.

For example, if the trainee renders the English sentence "If you take this bus, you can get to the station." as "kono basu de eki ni ikemasu," the function would judge it to be "GOOD" since the meaning is understandable but would advise the trainee that "kono basu ni noreba, eki ni tsukimasu" is a better choice of words (**Figure 4**). Another feature
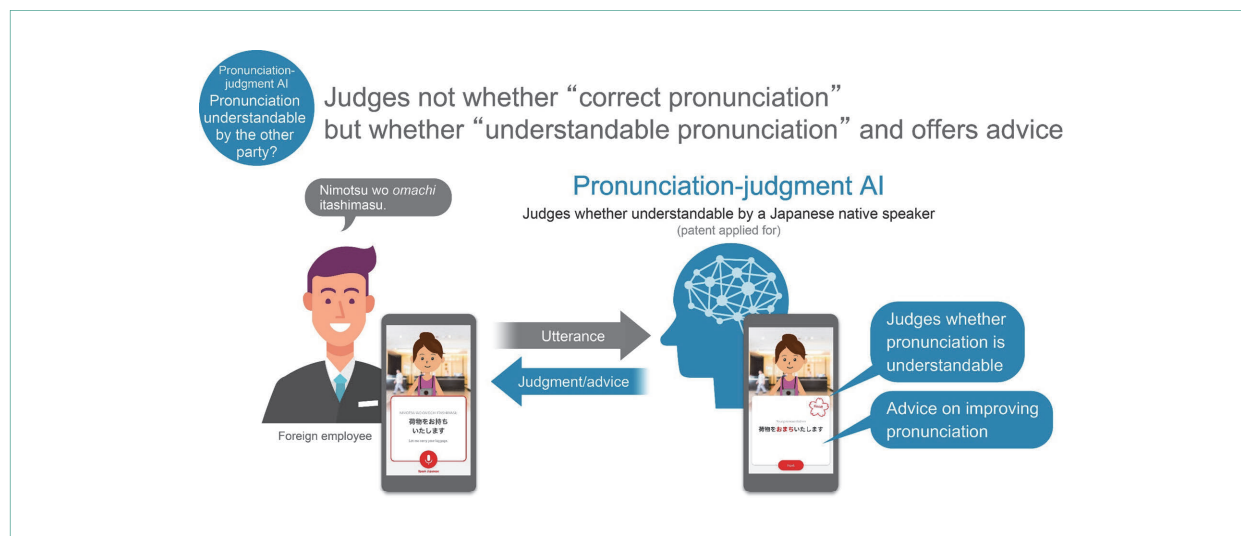

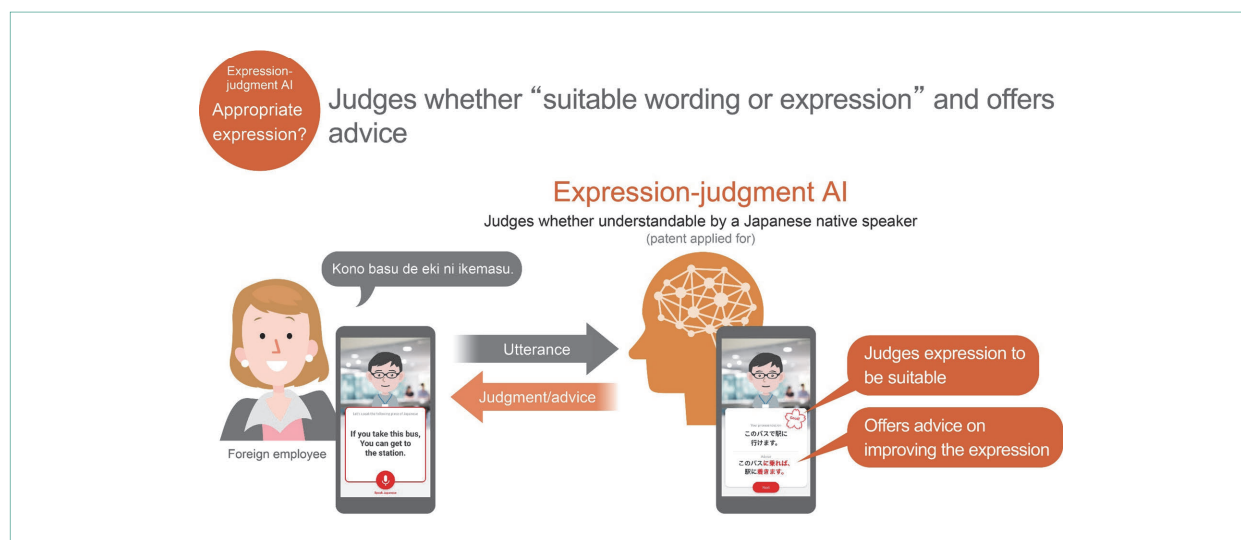Figure 3　Pronunciation-judgment function

Figure 4　Expression-judgment function

of this function is that it gives a good evaluation to even a different expression such as "kono basu de eki ni ikemasu yo" as long as it expresses the correct meaning.

2.3 Conversation-creation Function

The conversation-creation function enables the user to create original training content by inputting a practice phrase in the user's native language (English or Vietnamese) into the user's smartphone by speech or text. The JLT service therefore supports not only training with preinstalled training content but also user needs in the manner of "I would like to try saying this too in such a scenario."

3. JLT Configuration and Technology

The configuration of JLT is shown in **Figure 5**.

In JLT, pronunciation-judgment processing is performed on the terminal app while expression-judgment processing is performed on a server. We here describe the AI technology behind the pronunciation-judgment function.

The key feature of pronunciation-judgment AI is that it runs on the terminal app. Although there are systems that construct an acoustic model from the user's (foreigner's) spoken Japanese and perform associated processing on the server side, the cost of creating such an acoustic model and constructing a server for that purpose has been an issue. The pronunciation-judgment AI library[*1] that we have developed uses the text of a practice sentence and the text of the speech recognition result to judge pronunciation. Specifically, the function partitions the speech recognition result into individual words, removes the vowels in those
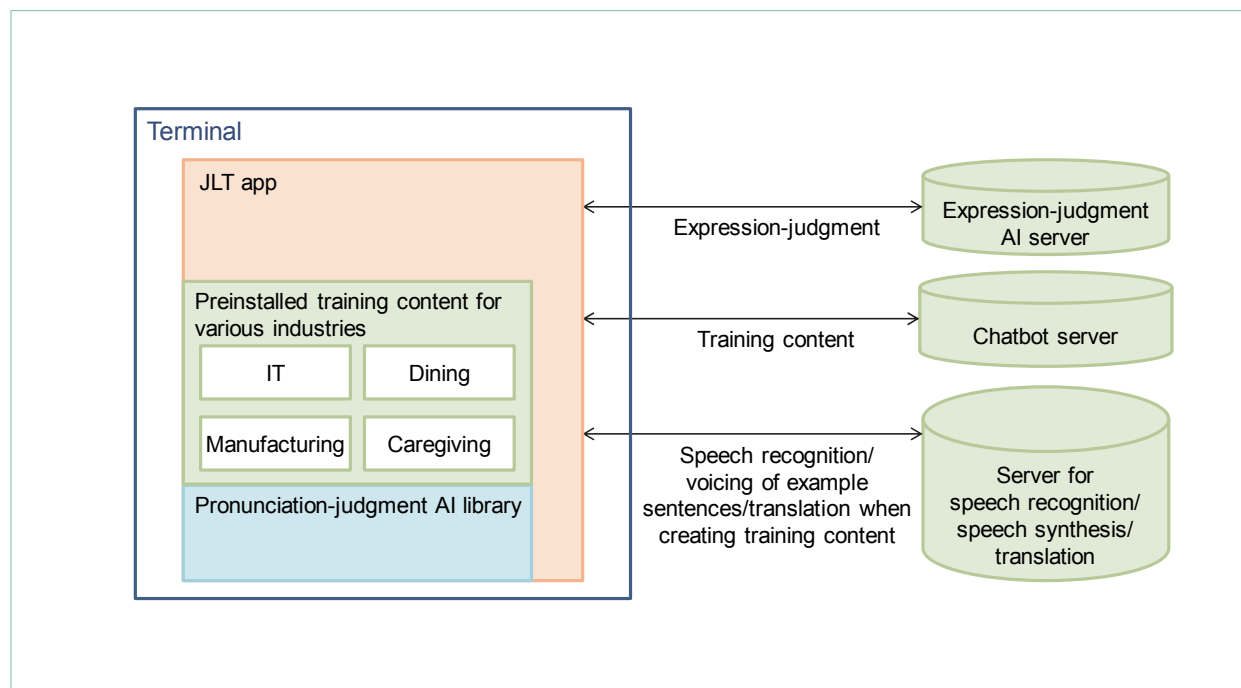

Figure 5　Function configuration

-----

words and substitutes consonant groups taking pronunciation similarity into account, and calculates the degree of similarity with respect to the combinations of all the words included in the text of the correct sentence and all the words included in the text of the speech recognition result (**Figure 6**).

The JLT service also implements a function for removing fillers before performing judgment. This makes it possible to appropriately judge the pronunciation of an utterance mingled with fillers sounds that would intrinsically be understood by a Japanese native speaker. Japanese spoken by a foreigner may include fillers (such as "ah"), repetitions, etc. that act as noise. Consequently, the results of judgment may be low even if the utterance is understandable to a Japanese native speaker. For example, given "kono basu de eki ni ikemasu" as the correct sentence, the user may utter "kono basu de eki ni *ah* ikemasu" so that the speech

recognition result would be exactly that. That is to say, if simply comparing the correct sentence with the speech recognition result of the user's utterance, the latter would turn out to be inappropriate with respect to the former due to the frequent use of "ah" in speaking resulting in a judgment of "error."

However, JLT performs judgment after removing fillers and repetitions so that the user's utterance in this case would be judged to be appropriate with respect to the correct sentence.

Pronunciation-judgment AI is not limited to judging the Japanese spoken by foreigners—it can also be applied to judging the English spoken by Japanese.

In this regard, a Japanese person skillful in English (corresponding to a TOEIC®*2 score of 800) and a Japanese person weak in English (corresponding to a TOEIC score of 400) were each asked to utter
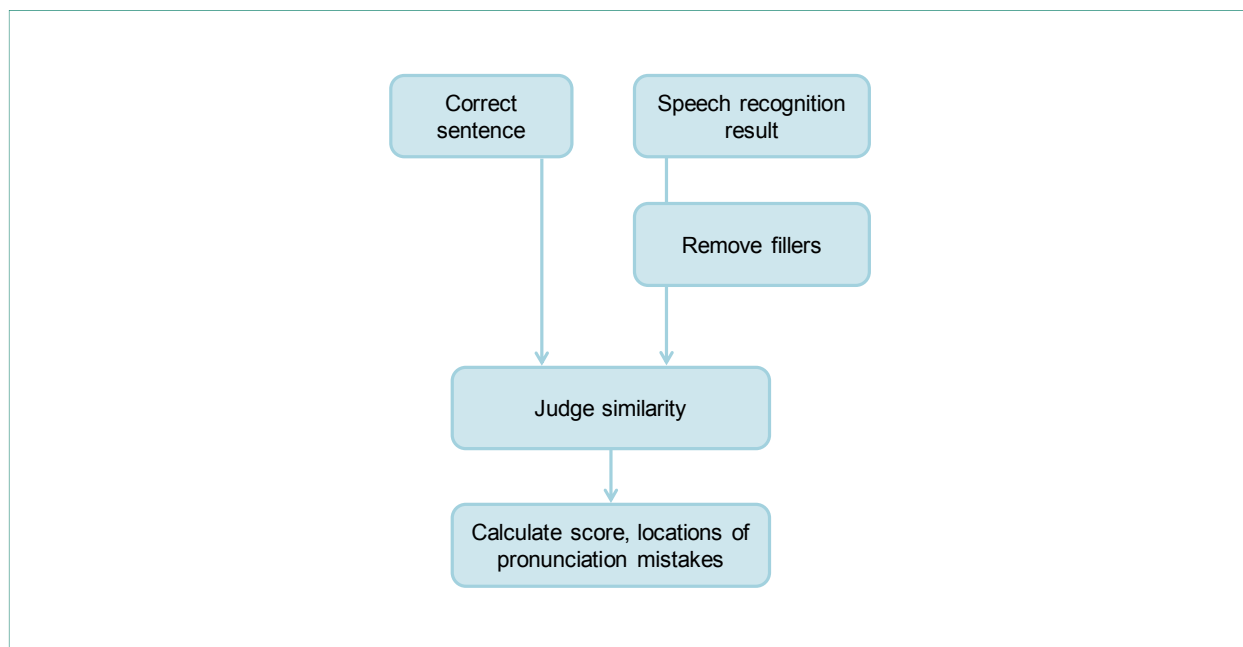

Figure 6   Pronunciation-judgment processing

---

150 example sentences in English. **Figure 7** shows the results of judging the pronunciation of those speakers by an English native speaker, another company's pronunciation-judgment system, and our pronunciation-judgment AI.

In judging whether pronunciation was understandable, our pronunciation-judgment AI demonstrated a performance approximately 5% higher with respect to utterances by the Japanese person skillful in English and approximately 16% higher

(a) Utterances of Japanese skillful in English

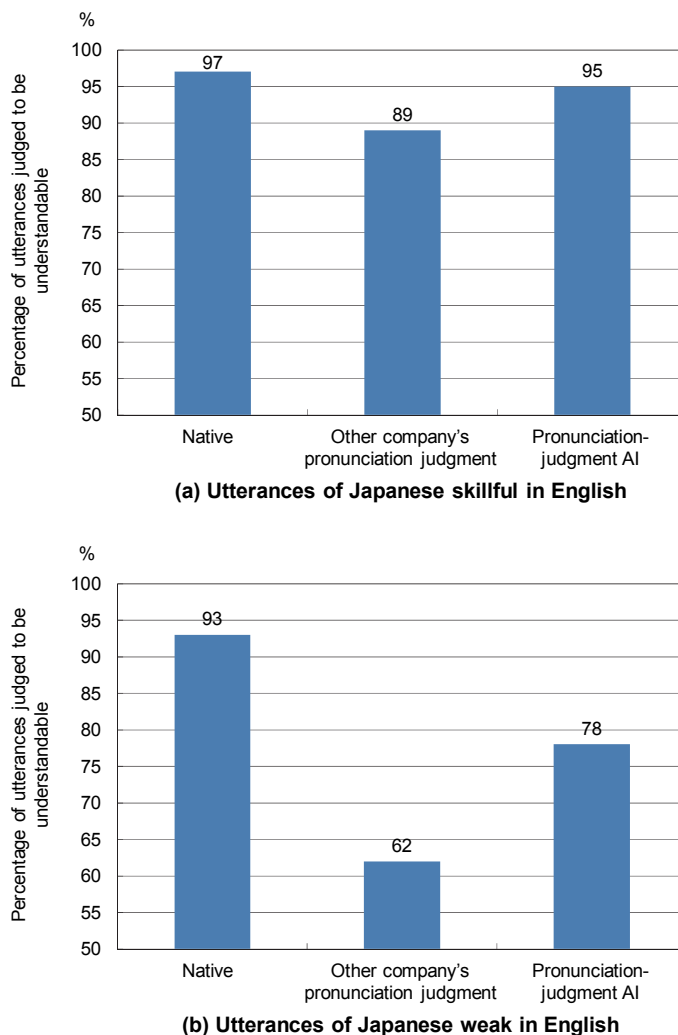**(b) Utterances of Japanese weak in English**

Figure 7   Performance evaluation of pronunciation-judgment AI

with respect to utterances by the Japanese person weak in English compared with the other company's pronunciation-judgment system. These results show that this technology is effective in judging whether pronunciation is understandable.

4. Verification Experiment with FPT Japan Holdings

To test the training support effect of the current version of JLT, we have been conducting verification experiments as a NTT DOCOMO TOPGUN project [3]. In this article, we introduce the verification experiment that we are conducting with FPT Japan Holdings Co., Ltd., which is the Japanese arm of FPT Software, the largest IT company in Vietnam.

While technical competence is, of course, essential, FPT Japan Holdings recognizes that communication in Japanese is also vitally important in getting customers in the Japanese market to entrust their work to another company with peace of mind. For this reason, the company is focusing its efforts on language acquisition by its Vietnamese employees by inviting a Japanese language lecturer every weekend and holding Japanese conversation classes free of charge for Vietnamese engineers living in Japan. There is also a plan to open an "FPT Japanese Language School" in Tokyo sometime in the future. Among these initiatives at FPT supporting Japanese language education, we have begun a verification experiment to explore the possibility of using JLT.

In the experiment, ten Vietnamese system engineers working at FPT Software Japan Co., Ltd., a subsidiary of FPT Japan Holdings, have been using

JLT equipped with training content oriented to the IT industry.

Comments such as those below have been received on JLT.

- The many items of content having different degrees of difficulty and designed for various conditions enable training that can be tailored to individual employees. We expect communication between customers and employees to be vitalized as a result.
- I feel an improvement in my conversational ability since I have to think about expressions and words on my own.

5. Conclusion

This article presented an overview of "Japanese Language Training AI," explained pronunciation-judgment AI technology, and described a verification experiment conducted with FPT. After receiving a 2018 Good Design Award, JLT is expected to develop even further from here on [4] [5]. We are currently providing JLT to companies undertaking Japanese language education for foreign staff and technical interns, companies helping foreigners with living in Japan, organizations that support international students, etc. We are also promoting tests to assess the JLT training effect in foreign staff education, interview practice for international students, and other applications [6] [7]. Looking to the future, we plan to provide multilingual support so that foreigners studying Japanese overseas can make good use of the JLT service.

REFERENCES
[1]   Cabinet Office: "On a Foreign Work Force," Feb. 2018

(In Japanese).
https://www5.cao.go.jp/keizai-shimon/kaigi/minutes/2018/0220/shiryo_04.pdf

[2] Ministry of Justice: "Bill to Revise Immigration Control and Refugee Recognition Act and Ministry of Justice Establishment Act," (In Japanese).
http://www.moj.go.jp/nyuukokukanri/kouhou/nyuukokukanri05_00010.html

[3] NTT DOCOMO: "TOPGUN Corporate Site—Japanese Language Training AI," (In Japanese).
https://www.nttdocomo.co.jp/biz/special/topgun/story06.html

[4] Japan Institute of Design Promotion: "List of Recipients of 2018 Good Design Award—Japanese Language Training AI."
https://www.g-mark.org/award/describe/47048?token=st2v3GuNJW&locate=en

[5] "NTT DOCOMO Receives 2018 Good Design Award," NTT DOCOMO Technical Journal, Vol.27, No.2, p.54, Jul. 2019 (In Japanese).

[6] NTT DOCOMO News Release: "(Notice) Trial Provision Begins of Japanese Conversation Training Service for Foreigners "Japanese Language Training AI"—AI judges and advises Japanese-specific "pronunciation" and "expressions"—," Oct. 2018 (In Japanese).
https://www.nttdocomo.co.jp/info/news_release/2018/10/03_00.html

[7] NTT DOCOMO News Release: "Verification Experiment of Japanese Conversation Training Service for Foreigners "Japanese Language Training AI" Begins with Provision to Companies—Tests training effectiveness when used in education of foreign staff—," Apr. 2019 (In Japanese).
https://www.nttdocomo.co.jp/binary/pdf/info/news_release/topics_190418_00.pdf#page=1