NICT NEWS
National Institute of Information and Communications Technology
No.4
2023
Vol.500

FEATURE

Aiming for a World of Mutual Understanding with AI

500th Anniversary Issue

Demonstration

Miss MURAYAMA Kirari "Plays in English" with NICT's AI Simultaneous Interpretation and VoiceTra®

CONTENTS

NICT NEWS 2023 No.4 Vol.500
National Institute of Information and Communications Technology

FEATURE

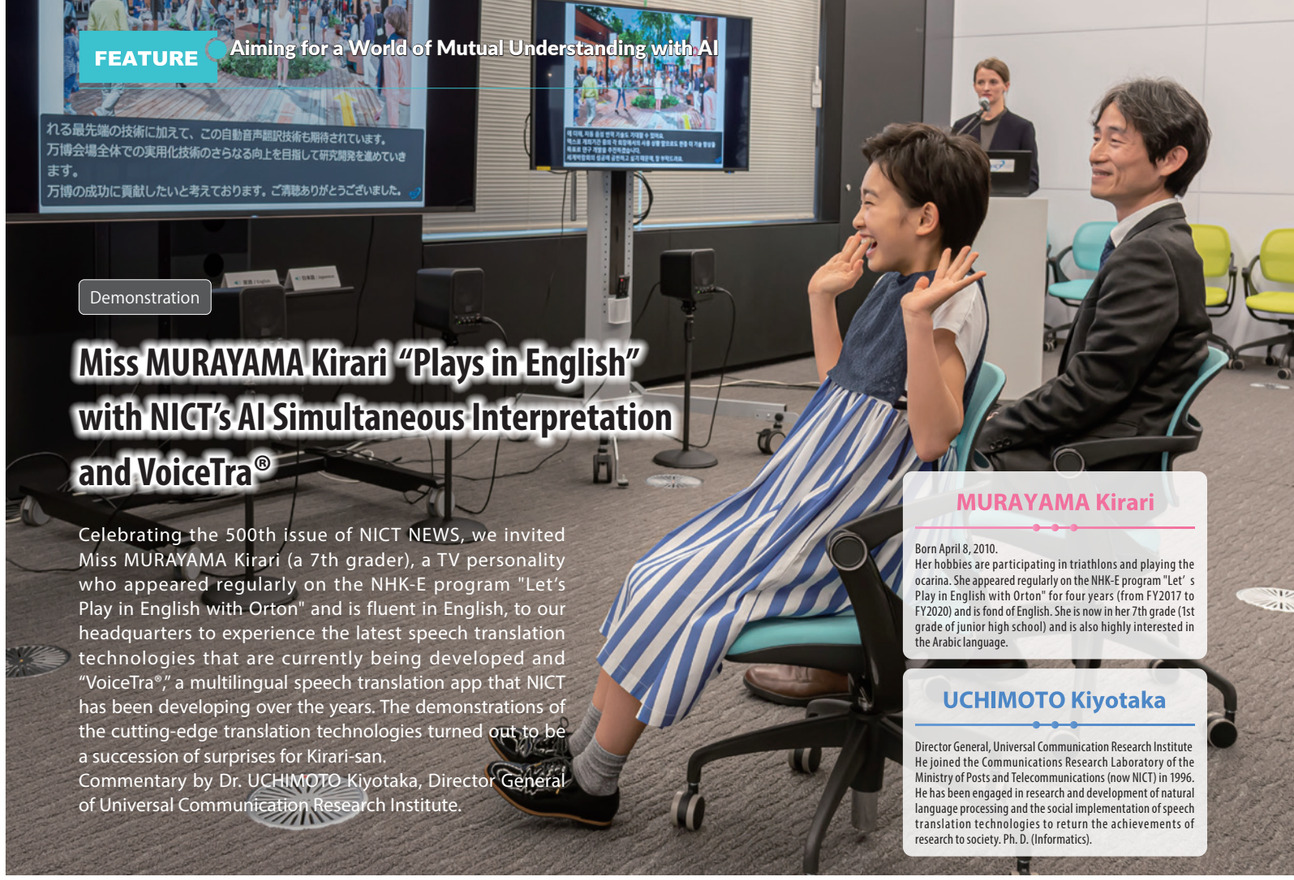
Aiming for a World of Mutual Understanding with AI

YouTube videos introducing our research, "NICT Station," released!

In June 2023, YouTube videos "NICT Station" which aims to have people understand research outcomes of NICT was released.

In the videos, researchers and NICT's official character "N" provide explanations on the following nine themes to a reporter who visited NICT. They explain the content and social impact of individual research themes in a joyful and easy-to-understand manner. The narrator is KAMISHIRAISHI Mone, a talented young actress, who also narrated in a video titled "A Future with N." Please enjoy the videos that will make you feel impatient for the future.

https://www2.nict.go.jp/publicity/nict_station/ (in Japanese)



Demonstration

Miss MURAYAMA Kirari "Plays in English" with NICT's AI Simultaneous Interpretation and VoiceTra®

Celebrating the 500th issue of NICT NEWS, we invited Miss MURAYAMA Kirari (a 7th grader), a TV personality who appeared regularly on the NHK-E program "Let's Play in English with Orton" and is fluent in English, to our headquarters to experience the latest speech translation technologies that are currently being developed and "VoiceTra®," a multilingual speech translation app that NICT has been developing over the years. The demonstrations of the cutting-edge translation technologies turned out to be a succession of surprises for Kirari-san.
Commentary by Dr. UCHIMOTO Kiyotaka, Director General of Universal Communication Research Institute.

MURAYAMA Kirari

Born April 8, 2010.
Her hobbies are participating in triathlons and playing the ocarina. She appeared regularly on the NHK-E program "Let's Play in English with Orton" for four years (from FY2017 to FY2020) and is fond of English. She is now in her 7th grade (1st grade of junior high school) and is also highly interested in the Arabic language.

UCHIMOTO Kiyotaka

Director General, Universal Communication Research Institute
He joined the Communications Research Laboratory of the Ministry of Posts and Telecommunications (now NICT) in 1996. He has been engaged in research and development of natural language processing and the social implementation of speech translation technologies to return the achievements of research to society. Ph. D. (Informatics).

VoiceTra®

UCHIMOTO Welcome to our institute, Kirari-san. The Universal Communication Research Institute promotes research and development of technologies to realize universal communication where everyone can understand each other with ease. It is research to support all kinds of human activities using AI.

Automatic translation technology, which I want to introduce to you now, is one form of such technology, and it allows anyone to easily overcome the language barrier when communicating with someone who speaks a different language. First, I would like you to try out VoiceTra®, a multilingual speech translation app that we developed. After that, we would like you to experience the AI simultaneous interpretation technology that can be used on occasions like international conferences.

KIRARI Thank you. I'm very excited.

UCHIMOTO VoiceTra® is a multilingual translation app that can be used free of charge simply by downloading the app on your smartphone. It supports both iOS and Android OS. Open the app, tap the microphone button on the screen, and speak in Japanese—your speech will be automatically translated into English and you will be able to hear the audio output. Please say something.

KIRARI "Konnichiwa, watashi ha MURAYAMA Kirari desu."

VoiceTra *"Hello! I am Kirari Murayama."*

KIRARI Wow! We can hear the English translation!

UCHIMOTO How about a longer sentence?

KIRARI "Asukara atsuku narisounanode, necchuushou niha kiwo tsukete kudasaine."

VoiceTra *"It will be hot from tomorrow, so please be careful of heatstroke."*

KIRARI It's impressive that even quite complicated dialogue can be translated instantly. Both the translated English and the back translation into Japanese are shown on the screen. Using this, I can probably have a carefree conversation with people from overseas.

UCHIMOTO Now, let's try a conversation that is closer to a real-life situation using VoiceTra®. You have just arrived at the airport in Los Angeles. You want to ask someone for a sightseeing recommendation. Just at the right moment, a kind-looking person passes by. It's Ananda.

KIRARI "Sumimasen. Kanko ni ikitainodesuga, Los Angeles no yumeina bashowo oshiete itadakemasenka."

VoiceTra *"Excuse me, I'd like to go sightseeing. Could you tell me some famous places in Los Angeles?"*

ANANDA "Nice to meet you. I recommend Griffith Observatory. Do you like stars?"

VoiceTra *"Konnichiwa, Griffith Temmond-*

Demonstration

Miss MURAYAMA Kirari "Plays in English" with NICT's AI Simultaneous Interpretation and VoiceTra®

ai e ikukotowo osusumeshimasu. Hoshi wa sukidesuka?"

KIRARI "Hai, namae ni hoshi to iu mojiga haitteirunode daisuki desu."

VoiceTra *"Yes, I love stars because my name has the word star in it."*

UCHIMOTO You can carry a conversation in such way.

KIRARI I'm surprised that the app speaks both English and Japanese so fluently. It's really good that the app accurately translates what I want to say and conveys it to the person I'm talking to. It's so nice that you can talk with people around the world normally using Japanese. I think having conversations with people in other countries will help to expand your knowledge and views.

AI Simultaneous Translation -Presentation-

UCHIMOTO Tapping the translation button like VoiceTra® may hinder smooth communication when it comes to high-paced conversations or long dialogues, since there will be a time lag. Thus, we are currently developing a technology called "AI simultaneous interpretation." Now, please try this. It can be used for occasions like international conferences where people from different countries speak in different languages. This would allow the audience to hear others speak in their own languages.

Let's ask Ananda, a native English speaker, to give a speech in English. Her speech will be instantly translated into Japanese and you can hear the audio output as well.

This system is being developed with the assumption that it will be used at EXPO 2025, Osaka, Kansai, Japan, under the theme of "People's Living Lab."

KIRARI The speech is being constantly translated just like a human simultaneous interpreter is speaking. How does it judge where to cut within a speech?

UCHIMOTO AI automatically identifies the segments in a sentence and translates in chunks that contain meaning. This was achieved by training the AI with experiences of human simultaneous interpreters.

KIRARI So that's why it's translated into such beautiful and proper Japanese!

UCHIMOTO Since the translation is generated constantly, you can understand what the person is saying as if you are actually listening to their speech. Many international conferences and keynote speeches by speakers from overseas are scheduled to be held at Expo 2025 in Osaka, so we hope this will help eliminate the language barriers in such occasions. I think this is something that has never been done in the history of the Expo.

Another goal of this translation technology is to keep the users unaware of the advanced technology used. We want to commoditize the technology to such an extent that people will use it without even thinking about it, just like we do with a pencil or an eraser.

Miss Kirari experiences remote meeting using NICT's AI simultaneous interpretation technology

KIRARI Technology as accessible as a pencil! Awesome!

UCHIMOTO There are so many languages in the world, but Japanese and English is actually one of the most difficult combinations for translation. For instance, the word order is completely different between the two languages. In English, the verb comes after the subject, but in Japanese, the verb comes at the end of the sentence. You cannot understand a Japanese sentence until you hear it until the end. Therefore, if Japanese-English translation—which is very difficult—becomes possible, the technology can be also applied to the translation of other languages.

Japanese is merely one of the many languages in the world, so it tends to be less prioritized in services overseas. However, as Japanese citizens, we would like to set Japanese as the basis. NICT has the advantage of possessing a vast amount of translation data that the institute has been accumulating for many years, and most of it is in Japanese.

AI Simultaneous Interpretation -Teleconference-

UCHIMOTO Next, let's try a teleconference with someone supposedly far away in a foreign country using AI simultaneous interpretation technology. It is useful not only for companies that communicate frequently with people abroad but also for private communication with friends overseas who speak different languages, so it can broaden your horizons.

Now, let's suppose you are visiting the U.S. in the near future and have a chat with Ananda, who lives in Los Angeles (but in reality is sitting in the next room).

KIRARI "America de zehi itta houga yoi tokorowa arimasuka?"

Translated Speech in English *"Are there any places you'd like me to visit in America?"*

ANANDA "Well, America is so big, so there are many, but you should go to the Grand Canyon in Arizona and the rainforest in Washington State."

Translated Speech in Japanese *"So ne, America wa totemo hiroi kara takusan arukedo Arizona-shuno Grand Canyon to Washington-shuno nettaiurin niwa ikubekidayo."*

KIRARI "Grand Canyon ha kiitakotoga arimasu. Zehi ittemitaidesu."

Translated Speech in English *"I've heard of the Grand Canyon. I'd love to go there."*

ANANDA "Yes, it's a very beautiful landscape, wide open space. It's not a crowded place like Tokyo. All you can see are stars and nature."

Translated Speech in Japanese *"Eh, asokowa totemo kireidesushi, hiroku aita kukan desu. Tokyo no yona konzatsushita basho denaku, mirukotoga dekirunowa hoshi to shizen dakedesu."*

KIRARI "Watshiwa hoshi mo shizen mo daisuki nanode ikunoga tanoshimidesu."

Translated Speech in English *"I like stars and nature, so I'm looking forward to going there."*

UCHIMOTO Conversation with someone who speaks a different language will go on like this. How was it?


Miss Kirari experiences multilingual communication using VoiceTra®

KIRARI There are some words that I know, but there are also parts I that couldn't make out, because native speakers talk fast and fluently, so it is really helpful that those parts are translated.

Future of Translation Technology

UCHIMOTO AI translation technology has already come this far. We would like to further increase the accuracy of translation and make it available for ordinary use. Kirari-san, how did you like communicating in a foreign language using our translation technology today?

KIRARI Usually, I don't have the courage to talk with people from overseas even if I wanted to. I'm often worried that I won't be able to clearly express what I want to, so I hesitate to speak. But if this kind of translation technology is available, I am sure that I can convey anything I want to. If I can talk with local people in English, I might be able to gain knowledge that I can't get from textbooks, so I think this is great.

UCHIMOTO Certainly. You can ask about various things you do not know, which you can understand in your native language; this technology may also be used to expand your knowledge. It also allows you to communicate in languages other than English too.

KIRARI I would like to have conversations using the translation app because I don't have any knowledge of or experience in foreign languages other than English.

UCHIMOTO It takes a huge amount of time to learn a new language, so I think it is better to let the machine do the work for certain cases.

The translation technology that we are developing is learning from so-called "accumulated experience," which is a collection of many case examples of parallel translation data. Whether it is English or another language, we collect a lot of experience on how sentences are translated.

The goal is to achieve technology that can tell you what you really want to know. Also, in business settings, there may be cases where negotiation is necessary. Therefore, we would like to develop a translation technology that can also cover social and cultural backgrounds so that people can understand each other in an even more natural manner.

KIRARI Yes, it would be even better if the translation could read the room and nuances of the situation.

UCHIMOTO In order to make AI translation system more sophisticated, a vast amount of parallel translation data is necessary. We would like to collect as much of these data as possible while being careful to prevent important intellectual properties from leaking to markets abroad, and develop technologies original to Japan.

KIRARI I'm looking forward to the future of translation technology. It's so exciting to hear that such advanced technology will be commonly available at EXPO 2025 in Osaka. Thank you very much for everything today.

Oh, my manager has downloaded VoiceTra® on her smartphone, so I will continue using it!

Development of Speech Communication Technologies

KAWAI Hisashi

Director of Advanced Speech Technology Laboratory, Advanced Speech Translation Research and Development Promotion Center (ASTERC), Universal Communication Research Institute

He received his Ph.D. in Engineering in 1989. He joined the Kokusai Denshin Denwa Co. Ltd. (currently KDDI Corporation) in April 1989. He has been engaged in research on speech synthesis and speech recognition. From 2000 to 2004, he worked for the Advanced Telecommunications Research Institute International (ATR). He joined NICT in October 2014.

In order to achieve the Global Communication Plan 2025 (GCP2025), which was announced by the Ministry of Internal Affairs and Communications in March 2020, the Advanced Speech Technology Laboratory has been working on three challenges, 1) R&D of speech recognition, language identification, and speaker recognition technologies for automatically converting speech of lectures and business/international conferences into texts in the primary languages, such as Japanese, English, and Chinese, with practical precision, 2) R&D of technology for synthesizing human-level speech to smoothly convey the translation results, etc. and technology for controlling voice quality, such as speech rate, while curbing deterioration in naturalness to the extent possible, and 3) expansion of the supported languages for speech recognition and text-to-speech regarding daily conversations.

Speech Recognition of Lectures and Meetings

In the process of enhancing accuracy in speech recognition, the creation of large-scale speech corpora (data accumulating a large amount of speech and text of the content thereof) for training speech recognition models while simultaneously improving modeling methods is critical. Speech in lectures or meetings are less clear as speakers' levels of consciousness for articulation are lower than in the case of carefully speaking sentence by sentence to smartphones using speech translation applications, and the distribution of acoustic features is also apt to vary. When trying to recognize such speech using an acoustic model trained on clear speech, the speech feature quantity deviates from the identification boundary and the accuracy in speech recognition deteriorates significantly. In order to deal with this problem, we trained acoustic models by adding a corpus of less clear speech at a level representing around 30% of the total to the existing corpus of clear speech (2,000 hours for the Japanese, English, Chinese, and Korean language, and 1,000 hours for other languages) to make the identification boundary adapt to such unclear speech. Materials for the acoustic corpus were mainly collected from mock lectures and meetings in order to avoid the issues of confidentiality, protection of personal information, copyrights, etc., but they also include real speech from open lectures and meetings within NICT, and the acoustic corpus is being used for training speech recognition models and evaluating the accuracy.

By FY2022, we created a speech corpus of around 3,000 hours in total mainly in Japanese, English, Chinese, and Vietnamese and achieved recognition with practical precision in Japanese and recognition with accuracy at an almost practical level in English. Generally, speech recognition accuracy is measured by the word error rate, but the rates fluctuate significantly depending on the content of the speech data for testing. Accordingly, the accuracy is described using the terms shown in Table 1. Furthermore, we improved the algorithms and acoustic models and halved the time lag from the input of speech into the speech recognition engine to the output of the results, which had been 4.1 seconds. While making these steady efforts for improvement, we are working on studies for practical realization of end-to-end speech recognition in which a string of characters is directly output by inputting speech signals into a neural network.

In language identification, when acoustic features differ between speech data to be used for identification model training and real speech data which is the target of identification, the problem of domain dependency arises, and the identification accuracy deteriorates significantly. It is thus difficult to secure identification accuracy at a practical level mainly for languages for which real speech data for training are hard to collect. As possible causes, over-fitting to characteristics of surrounding noise when recording and frequency response characteristics of microphones were suspected; therefore, we devised the Transducer-based language embedding method which allows modeling of linguistic features (features of a chain of phonemes, the minimum units of spoken language), in addition to acoustic features (Figure 1). When we evaluated the language identification accuracy by applying this method to training data that recorded a mock conversation and real data for testing, the identification error rate for a two-second speech almost halved from 17.7%.

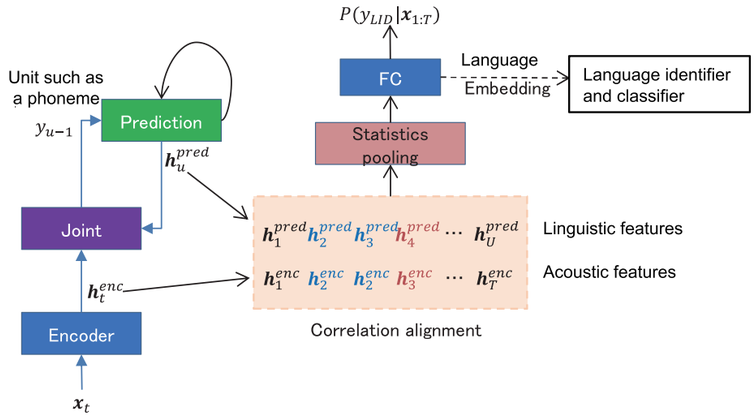


Figure 1 Transducer-based language embedding to identify languages in consideration of both acoustic features and linguistic features

Table 1 Classification of speech recognition accuracy

S	Human level	The results of speech recognition can be read and understood without any problem.
A	Practical level	The results of speech recognition contain minor errors but can be read and understood sufficiently.
B	Almost practical level	The results of speech recognition contain errors but can be read and understood to some extent.
C	Experimental level	The results of speech recognition contain many errors and are difficult to understand.
D	Prototype level	The results of speech recognition contain many errors and are extremely difficult to understand.

Human Voice Text-to-Speech

As for text-to-speech, we are also working on the development of speech corpora for model training and the improvement of modeling methods. In principle, the speech corpora contain speech in individual languages by pairs of one male and one female who are voice actors or other occupational speakers. We adopt neural networks for the text-to-speech models and are conducting R&D while balancing sound quality and speed improvement. We focus on using general-purpose CPUs to overcome the economic obstacle for commercial implementation which may arise if text-to-speech services require expensive GPGPUs.

One of the neural networks that we are working on has the structure as shown in Figure 2. The computation speed is RTF = 0.4 (Real Time Factor; the ratio of the computation speed to the length of speech time). The sound quality is MOS = 4.3 (Mean Opinion Score; subjective evaluation value on a scale of five), a value close to around MOS = 4.5 of a human voice.

When applying text-to-speech as the output of automatic simultaneous interpretation services, the synthesized speech is sometimes adjusted to play faster than normal speed to shorten the delay between the indication of the translated text and the completion of the synthesized speech. Stretching or shortening the whole speech depending on the elasticity of phonemes is a well-known method. When we applied this to the aforementioned neural network, the results of subjective evaluation showed that the sound quality does not deteriorate significantly within the stretching rate between 0.75-1.25 times the original speed.

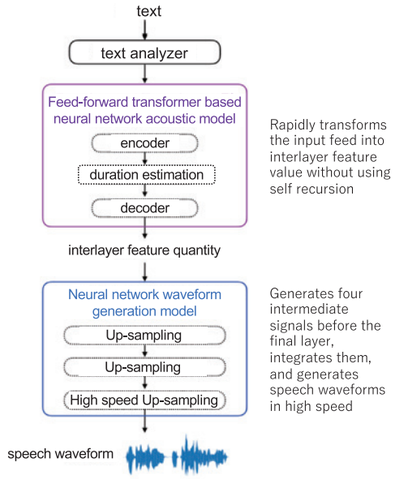


Figure 2 Neural network that can synthesize high quality speech at a high speed

Speech Recognition and Text-to-Speech for Daily Conversations

As for the 15 primary languages stated in the GCP 2025, we have mostly completed our missions for speech recognition and text-to-speech, and have published the results via VoiceTra®, a speech translation app released as a field experiment, and have commenced the granting of commercial licenses.

Under the supplementary budget for FY2021, it was decided to newly support Russian, Arabic, Hindi, Italian, and German. We created speech corpora and developed models for these languages in FY2022, and with a few exceptions, we released them on VoiceTra® by the end of the same fiscal year. Additionally, in the same year, there was an urgent need to support speech recognition and text-to-speech for Ukrainian in consideration of the Ukrainian refugees. We intensively put efforts into the procurement of the speech corpus and model development, and as a result, we were able to support Ukrainian on VoiceTra® by August 2022. It turns out that models were not the only ones that were trained--the fact that we were able to respond to a new language and complete social implementation in such short period of time, is the outcome of our long-term efforts on multilingualization.

Future Prospects

The GCP 2025 aims to achieve social implementation of automatic simultaneous interpretation technology, taking the opportunity of the EXPO 2025 Osaka, Kansai, Japan as a showcase. The EXPO will be held from April 13, 2025, and there is less than two years remaining as of the time of authoring this article. We will arduously make further efforts to improve the accuracy and speed of speech recognition, as well as the quality and speed of text-to-speech in this short stretch of time.

Automatic Simultaneous Interpretation Technology

Technology that automatically interprets spoken words without delay

UTIYAMA Masao

Director of Advanced Translation Technology Laboratory, Advanced Speech Translation Research and Development Promotion Center (ASTERC), Universal Communication Research Institute

After completing graduate school, he joined the Communications Research Laboratory (currently NICT) in 1999. He has been engaging in research on natural language processing, especially machine translation. Ph.D. (Engineering)

Speech translation has become quite familiar thanks to the emergence of various applications. However, at present, sequential translation to translate speech sequentially after a speaker finishes speaking is main stream. However, NICT has been conducting R&D on simultaneous interpretation technology to have an application simultaneously translate speech while a speaker is speaking. This technology makes it possible for a speaker to continue speaking without being aware of the intervention of an interpreter and also significantly reduces the time lag in translation.

Background

Foreign visitors had decreased in number due to COVID-19 but are increasing again. Additionally, foreign residents are also increasing. Accordingly, it is expected that opportunities for communications with foreign nationals will increase nationwide, including in local areas, and that people will face a language barrier in various scenes at work and in daily living, such as in administrative procedures or on occasions of receiving medical services, using transportation services, or enjoying sightseeing. Toward the World Exposition (Osaka Kansai EXPO) scheduled to be held in 2025, dramatic development of automatic translation technology is expected to realize the multilingual simultaneous interpretation system, which enables conversations with people around the world on a real-time basis. The technology is expected to be utilized at pavilions and lecture venues. Under such circumstances, NICT established an R&D consortium together with other companies,[*1] and has been conducting R&D on automatic simultaneous interpretation technology under entrustment of the Ministry of Internal Affairs and Communications for a program, "Research and Development of Advanced Multilingual Translation Technology,"[*2] in its project, "R&D Project for Information and Communications Technology (JPMI00316)."

Outline of Automatic Simultaneous Interpretation Technology

Figure 1 shows the basic form of simultaneous interpretation by computer. [i] First, a speech that was input is changed into text using a speech recognition module and the text is seamlessly input into the next module, a divider. [ii] The text is divided by the divider module into chains of parts of words (chunks as translation units) that are suited to translation as had been learned in advance. [iii] Those chunks are converted into one of multiple languages by an automatic translation module. [iv] They are output in sequence as speech by a speech synthesis module.

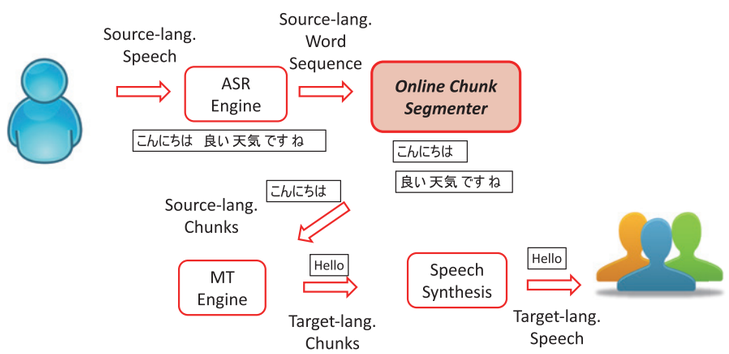


Figure 1 Outline of automatic simultaneous interpretation technology (a significant difference from sequential translation is the existence of a chunking module)

Table 1 Japanese sentences divided into chunks and the results of their automatic translation into English

おはようございます	Good morning.
F 誌 記者 の カーティスタッカー と申します	I'm Curtis Tucker, a reporter for F Magazine
プレゼン の 中 で	In the presentation.
この 製品 の 開発 に 七 年 かかった ことに 言及 されて いました	It was mentioned that it took 7 years to develop this product.
特に ご苦労 なさった 点 は 何 だった と思わ れます か	What do you think was the most difficult point?
設置 手順 を 簡単 に する ため	In order to simplify the installation procedure.
持ち運び できる 大きさ に する こと が	You should make it portable.
開発 当初 から 主な 目標 の ひとつ でした	It has been one of the main goals since the beginning of development.

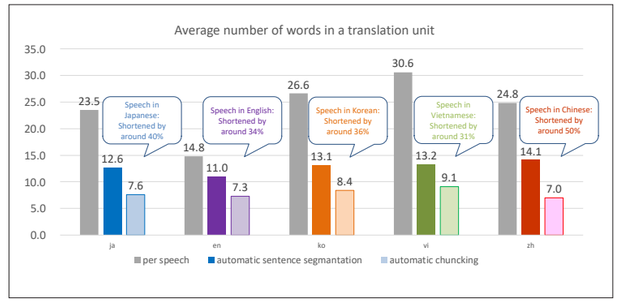


Figure 2 Average number of words in a translation unit for multiple languages

Differences between simultaneous interpretation and sequential translation are as follows.

- In the case of sequential translation, translation starts after a speaker finished speaking, while in the case of automatic simultaneous interpretation, the system needs to decide on the timing of automatically starting interpretation.
- In the case of sequential translation, one to several sentences are translated at once, while in the case of automatic simultaneous interpretation, it is necessary to do interpretation in a shorter unit (chunk) to reduce the time lag.
- In the case of sequential translation, speech as a whole must be translated without omission as it is, while in the case of automatic simultaneous interpretation, it is sometimes required to sort out the content of speech in order to promote listeners' understanding and reduce the time lag.

Additionally, for both sequential translation and simultaneous interpretation: which need to do translation or interpretation in multiple languages, it is necessary to apply technologies for dealing with those differences mentioned above for multiple languages.

Technology to Intermittently Divide Chains of Input Words

In simultaneous interpretation, a speaker's speech is recognized in succession as a chain of words, and dividing this chain of words into chunks is the first step. For this purpose, we conducted research and developed a dividing module through machine learning technology using a neural network. In the machine learning, we (1) created training data: which designate where to divide an input sentence, in multiple languages, and (2) created a dividing module that simulates the former. Table 1 shows an example in which a chain of words in Japanese is divided into chunks and they are automatically translated into English. Figure 2 shows the number of words per speech, the number of words by automatic sentence segmentation, and the number of words by automatic chunking. From this, it is found that the number of words in a translation unit decreases significantly only by dividing speech into sentences. For example, in the case of Japanese in Figure 2, the number of words until a speaker finishes speaking is around 23, but when dividing speech into sentences, translation can be started earlier when the number of words is around 13. Furthermore, when dividing speech into chunks, the length of a translation unit becomes shorter than in the case of dividing speech into sentences and the length is shortened by 30% to 50% depending on languages. In this manner, by doing translation by units of chunks without delay, smooth communications can be promoted.

Sorting out of the Content of Speech

In the case of simultaneous interpretation, technology to shorten sentences in speech is sometimes required. We also realized this technology through machine learning using a neural network. Through the use of the technology to shorten sentences, it is possible to plainly communicate the content of speech. Furthermore, the results of speech recognition do not contain punctuation marks, and therefore, we also conducted R&D on technology to insert punctuation marks. The following examples are a sentence that is a result of speech recognition, a sentence wherein punctuation marks are automatically inserted, a shortened sentence, and a translated sentence.

○ Example Japanese to English

(Output of speech recognition) 先だって 発信 し た 通り この 度 当社 製品 を より 広く 紹介 する ため 三か月 後 の 展示 会 に 出展 する こと に なり ま し た

(With punctuation marks inserted) 先だって発信した通り、この度当社製品をより広く紹介するため、三か月後の展示会に出展することになりました。

(Shortened sentence) 当社の製品をより広く紹介するために、3か月後に展示会に出展することになりました。

(Output of translation) In order to introduce our products to a wider audience, in three months' time, we will be exhibiting at an exhibition.

○ Example English to Japanese

(Output of speech recognition) under what kind of criteria and when was the seismic capacity evaluation conducted in kawanishi city

(With punctuation marks inserted) Under what kind of criteria, and when was the seismic capacity evaluation conducted in kawanishi city.

(Shortened sentence) What kind of criteria and when was the seismic assessment in Kawanishi City?

(Output of translation) 川西市での耐震診断はどのような基準でいつでしたか?

Future Prospects

At present, automatic simultaneous interpretation technology is being developed rapidly. Please expect that the multilingual interpretation system that enables real-time conversations with people all over the world will have been made available by the time of the 2025 Osaka Kansai EXPO.

*1 https://www.toppan.co.jp/news/2020/08/newsrelease_200828_2.html (In Japanese)
*2 https://www.soumu.go.jp/main_content/000700925.pdf (In Japanese)

Data-driven Intelligent Communication Technology

OHTAKE Kiyonori

Director General, Data-driven Intelligent System Research Center, Universal Communication Research Institute

After finishing a doctoral course in 2001, he joined ATR spoken language communication laboratory. He joined NICT in 2006. He has been engaged in research projects on natural language processing and spoken language processing. Dr. of Engineering

A multimodal dialogue system for elderly care (MICSUS)[*1] is one of the data-driven intelligent communication technologies researched and developed by the NICT Universal Communication Research Institute (DIRECT). During FY2022, large-scale MICSUS usability testing was conducted across Japan from Hokkaido in the north to Shikoku and Kyushu in the south. This article reports the final results of this testing and gives updates on our MICSUS R&D.

Background

During NICT's fourth medium-to-long-term plan implementation period, DIRECT has researched and developed a technology capable of performing in-depth semantic analysis of text information available primarily on the internet (i.e., web-based information). We also worked to put this technology into practical use. During the subsequent fifth medium-to-long-term plan period which began in FY2021, we have continued to develop the web-based information analysis technology into a verbally interactive technology able to engage in friendly, face-to-face conversation with its users for the purpose of providing them with helpful web-based information and collecting specific information from them. We collectively call these technologies "data-driven intelligent communication technologies." MICSUS is a data-driven intelligent communication technology developed in collaboration with KDDI Corporation, NEC Solution Innovators, Ltd. and the Japan Research Institute, Ltd. This project was funded by the Japanese Cabinet Office's Cross-ministerial Strategic Innovation Program (SIP) during the program's second period (FY2018–FY2022).

MICSUS Objectives

Japanese society has been rapidly aging in recent years due to declining birth rates and economic difficulties. This is causing a serious social issue: the lack of an adequate workforce to provide nursing care to senior citizens. While demand for nursing care is growing, the amount of financial support and the number of laborers needed to sustain services are falling short. Although it is estimated that 253,000 nursing care workers will be needed by 2025, this demand is unlikely to be met. Japan faces a shortage of 377,000 workers.[*2] Traditionally, care managers—a key player in nursing care services—evaluate the health of elder care recipients through in-person interviews. MICSUS is designed to perform some of these care managers' tasks, thereby reducing their workload. In addition, MICSUS is intended to offer elderly people more frequent communication—a vital activity in maintaining health—through casual conversation with them using web-based information.

Figure 1 MICSUS usability testing in action in Hidaka, Kochi Prefecture

For details about the structural and technological aspects of MICSUS, please refer to the NICT Journal article published in 2022 (Vol. 68, No. 2, pp. 141–149)[*3]. In this report, we mainly focus on the results of large-scale MICSUS usability testing carried out in FY2022 which are not included in the previous article.

Large-scale MICSUS Usability Testing with Elderly People

Large-scale usability testing was conducted across Japan from June 2022 to January 2023. 179 elderly people participated. A total of 927 conversations consisting of 26,704 speech exchanges between MICSUS and elderly subjects (who were responsible for 12,885 of the remarks) were recorded during the testing period. The cumulative duration of these conversations was approximately 95 hours and individual conversations lasted approximately six minutes on average. A subject spoke approximately 14 times on average during a single conversation. These indicate that elderly MICSUS users conversed for significantly longer periods of time than elderly people who used other related technologies (i.e., smart speakers). An elderly lady in Hidaka, Kochi Prefecture, one of the usability testing participants, used MICSUS daily for 15 days. A video showing her conversing with MICSUS can be viewed at https://www.youtube.com/watch?v=cuIHPYEqOKc&t=75s (in Japanese) (Figure 1). We evaluated the nearly 13,000 remarks made by elderly people recorded during the nationwide experiments—one of the most extensive sample sizes for this type of research ever conducted in Japan. We asked three third-party evaluators to examine the legitimacy of each remark by reviewing the entire conversation in which it appeared. When their opinions were divided, they took a vote. For example, they determined whether a subject responded with "yes," "no" or "I don't know" in answering MICSUS' question, "Are you eating three meals daily?" When MICSUS's semantic interpretation of its users' responses was compared with that of the human evaluators, its interpretation accuracy using its deep learning capability was found to be 93.5%. In addition, 92.9% of MICSUS' casual conversation responses to its elderly users and its communication in relation to other services were judged to be appropriate. During casual conversations, 91.8% of MICSUS' 1,346 responses made using web-based information were found to be appropriate. Furthermore, 25% of MICSUS' answers made elderly users smile and more than half of elderly subjects found conversing with MICSUS interesting. These results suggest MICSUS' potential effectiveness in offering elderly people more frequent communication. We believe that these results would be difficult even for humans to achieve and indicate MICSUS' ability to hold high-quality conversations. We also compared the time it takes for care managers to check the health of elderly people with and without the assistance of MICSUS. We found that the use of MICSUS can reduce the amount of time care managers spend on health checks by 69% or more. In addition, we conducted a survey of 120 elderly people who participated in the MICSUS experiments. The average rating of the overall MICSUS experience was 4.2 on a scale of 1 to 5, with 5 being the best experience. When we interviewed elderly users, their comments included, "Conversation with MICSUS was very enjoyable and I was even tempted to joke around with it," and "Talking to MICSUS encouraged me to use my mind."

Future Prospects

We may consider integrating a large-scale generative AI-based language model, such as OpenAI's ChatGPT, into MICSUS' casual communication function in the future. However, generative AIs are known to sometimes produce false responses when answering questions (i.e., they may lie). Our assumption is that most MICSUS users are elderly people with relatively low levels of information literacy. To ensure the credibility of the information MICSUS provides, we may program it to orally inform users of the URLs of websites from which it derives information. We will also encourage MICSUS to frequently warn its users to use web-based information carefully, as the risk of obtaining false information from the internet cannot be completely eliminated. We are currently against incorporating a large-scale generative AI-based language model into MICSUS because this would make its use significantly more costly due to the extensive resources needed to run such models. During the usability testing, MICSUS held casual conversations using a medium-sized web-based information database prepared in advance. This scheme proved effective in enabling MICSUS to hold proper conversations without performing intensive deep learning tasks. Cost efficiency is an important consideration, since elderly people will be MICSUS' primary users. In addition to the MICSUS project, DIRECT is researching and developing a generative AI-based dialogue system capable of formulating hypotheses using the huge amount of information available on the internet and providing them to its users along with the hypotheses' factual bases. Our goal is to make this system capable of facilitating brainstorming by users to address challenging issues.

---

*1 MICSUS: Multimodal Interactive Care SUpport System

*2 Ministry of Health, Labor and Welfare, Press Release, "Demand/supply Estimation of elderly care workforce" https://www.mhlw.go.jp/stf/houdou/0000088998.html (in Japanese)

*3 NICT "Journal of NICT". Vol.68, No.2 https://www.nict.go.jp/publication/shuppan/kihou-journal/houkoku68-2_HTML/2022U-03-05.pdf (in Japanese)

Smart Data Analytics Technology

Predictive analysis and behavior support through sensing data collaboration

ZETTSU Koji

Director General, Big Data Integration Research Center, Universal Communication Research Institute

After completing graduate school, he joined NICT in 2005. He has been engaging in research on data mining, machine learning, and databases. Doctor of Informatics

A smarter, more sustainable society could be developed through integrated use of various types of sensory data. This approach to data utilization could enhance comprehension of complex real-world events and allow optimal actions to be taken in specific circumstances. We have been developing a smart data analytics technology capable of integrating different types of sensory data and making it available for prediction and analytical purposes. Our specific R&D projects include: 1) the development of cross-data analysis techniques which can be used to collect various types of real-world sensory data and identify, analyze and estimate correlations between events across different disciplines; 2) development of a decentralized, federated AI technology for constructing prediction models designed to help people comprehend and respond appropriately to real-world events using data obtained anonymously from different individuals; and 3) construction of an xData (cross-data) platform into which the foregoing technologies will be incorporated. In addition, we are working with our collaborators to tackle various issues using the xData platform.

Background

Global climate change and growing population disparities between urban and rural areas in recent years have prompted the advent of complex urban issues related to the environment, traffic, natural disasters and energy. To cope with these issues, the Japanese government formulated a comprehensive data strategy with the goal of making society smarter and more sustainable. This strategy states that the key to creating new value is twofold: integrated use of a wide variety of data and construction of platforms to analyze integrated data for the purpose of creating and offering helpful services. A cyber-physical system (CPS) combines various types of data from physical space, analyzes it using AI techniques to comprehend real events appropriately and suggests proper actions people can take under specific circumstances. These CPS functions can be used to meet comprehensive data strategy objectives. Our research has therefore focused on developing a smart data analytics technology as a CPS component.

Cross-data Analysis Techniques

We have been researching and developing a multimodal AI technology capable of performing cross-data analysis by identifying and estimating difficult-to-detect correlations between different types of data. This technology is able to discover temporal and spatial associative patterns which may exist in data collected from various physical space sources using a convolutional recurrent neural network (CRNN), a graph neural network, etc. Discovered associative patterns have been used to predict short-term changes in environmental quality (e.g., using an air quality index (AQI)) in localized areas. We have also developed an MM sensing technology capable of predicting traffic hazards (e.g., traffic-related air pollution and dangerous driving conditions) and AQI in surrounding areas for drivers using image log data captured by portable and fixed cameras. It supports drivers by predicting traffic pollution and driving risks (near-misses, etc.) using dashboard cameras, etc. (Figure 1). To improve the prediction performance, we have developed a cross-modal attention model that can accurately predict correlations at the event-relationship level as well as at the conventional feature space level by employing multiplex attention (a machine learning mechanism that dynamically identifies key data) to object elements extracted from different types of data. The model received a high score in international benchmarking*. Furthermore, we have been researching zero-shot/few-shot learning methods which can be used to make prediction models adapt to new situations even when only very small amounts of relevant training data are available.

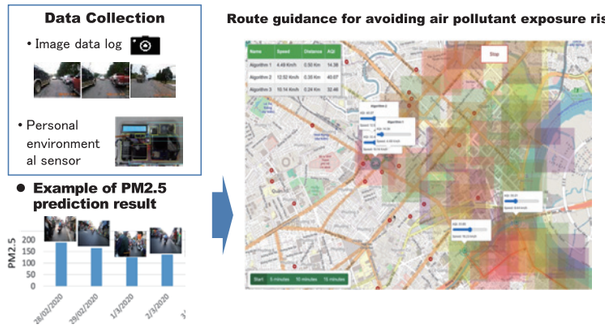


Figure 1 Support for traffic pollution control in an ASEAN region

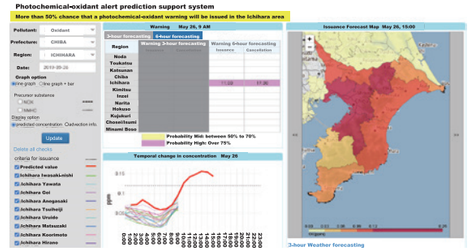


Figure 2 Early warning of photochemical-oxidant pollution based on short-term environmental quality prediction (collaboration with environmental monitoring service providers)

Distributed Federated AI Technology

Centralized AIs—which use cloud platforms to perform all necessary tasks from data collection to machine learning—have traditionally been the most widely used AI resources. However, this form of AI is becoming increasingly problematic because of growing global concern about privacy and domestic data protection. To address this issue, we have been researching and developing a federated learning technology that enables distributed learning by multimodal AI models while maintaining data in individual edge environments.

Since data collected on events at the edge will vary in terms of area, duration, frequency and density, we are developing methods to aggregate models to account for these variations and improve the performance of federated learning. Moreover, since available computing resources change dynamically in the edge environment, we are developing a distributed learning method that can flexibly respond to these changes.

xData Platform

We are constructing an xData platform which will allow us to develop smart applications using cross-data analysis technology. This platform stores hundreds of terabytes of event data extracted from sensory data in various fields, such as environmental, transportation and health data. The platform is equipped with a number of application programming interfaces (APIs) to facilitate various tasks—data collection and integration, correlation analysis and prediction using data mining, multimodal AI, etc., generating maps to graphically represent analytical results, issuing alerts and searching for optimal routes. We are also developing xData Edge, a tool that will provide users with more secure analysis and prediction based on their data using federated learning and other techniques. We will also equip the xData Platform with information assets that package data models, prediction models, processing programs, sample data, development manuals, etc., for each field in which cross-data analysis is applied. The functions and information assets of the xData platform are also deployed in the Data Centric Cloud Service (DCCS) of the NICT Testbed to promote application development and verification experiments as a platform for data and service collaboration and application development.

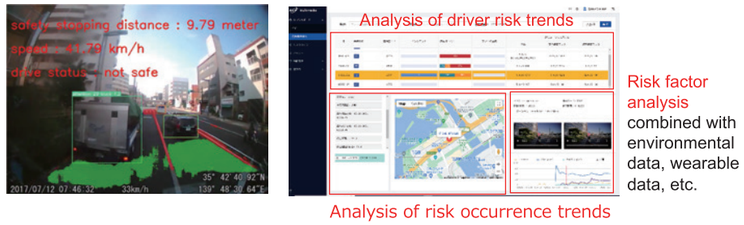


Figure 3 Prediction and analysis of driving risk events using MM sensing (collaboration with transport companies)

Practical Application Examples

NICT's short-term environmental quality prediction information asset is able to analyze a combination of environmental and meteorological data collected from various sources and perform short-term, fine-grained environmental quality prediction. Using the information asset, an environmental monitoring company is developing an application to provide early warnings of high levels of photochemical oxidants (Figure 2)—the only air pollutant in Japan whose concentration continues to exceed environmental quality standards. The company is using DCCS to customize the application to meet the specific needs of different municipalities by collecting additional data, adjusting training data, optimizing prediction processes and adding post-processing functions in accordance with municipalities' workflows.

NICT's MM sensing information asset detects the trajectory and speed of surrounding objects from dashboard camera data, etc. and estimates safe areas where collisions can be avoided while extracting potential events leading to driving risks and classifying events in detail according to the type of driving risk, occurrence status, etc. A transport company has developed an operation management application using the MM sensing information asset (Figure 3). The application has been used to analyze risk trends for individual drivers and analyze risk factors in combination with environmental data and driver biometric data collected from wearable devices, etc.

Future Prospects

Based on the xData platform, we designed cyberspace architecture for data distribution, storage, analysis and prediction in Beyond 5G/6G and published it in NICT's Beyond 5G/6G white papers, etc. In addition, we will contribute to the realization of CPS in the beyond-5G/6G era through research and development on the connection between cyberspace and physical space, edge AI for behavioral support and digital twin collaboration using an orchestrator.

* ACM Multimedia Grand Challenge on Detecting Cheapfakes

History of Changes up to NICT NEWS 500th Issue

Press Office, Public Relations Department

This issue of NICT NEWS is the 500th issue counting from the first issue of its predecessor, Radio Research Laboratory (CRL) News. Here, we would like to look back on the history over nearly half a century.

The first issue of Radio Research Laboratory News, a PR magazine of the Radio Research Laboratory under the jurisdiction of the Ministry of Posts and Telecommunications of the time, was published in April 1976 (Photo 1). 1976 was a year of political turbulence during which former Prime Minister TANAKA Kakuei was arrested due to the Lockheed scandal. The first issue contained an article concerning acoustic influences of lightning discharge and introduced an experiment of underwater laserscope.

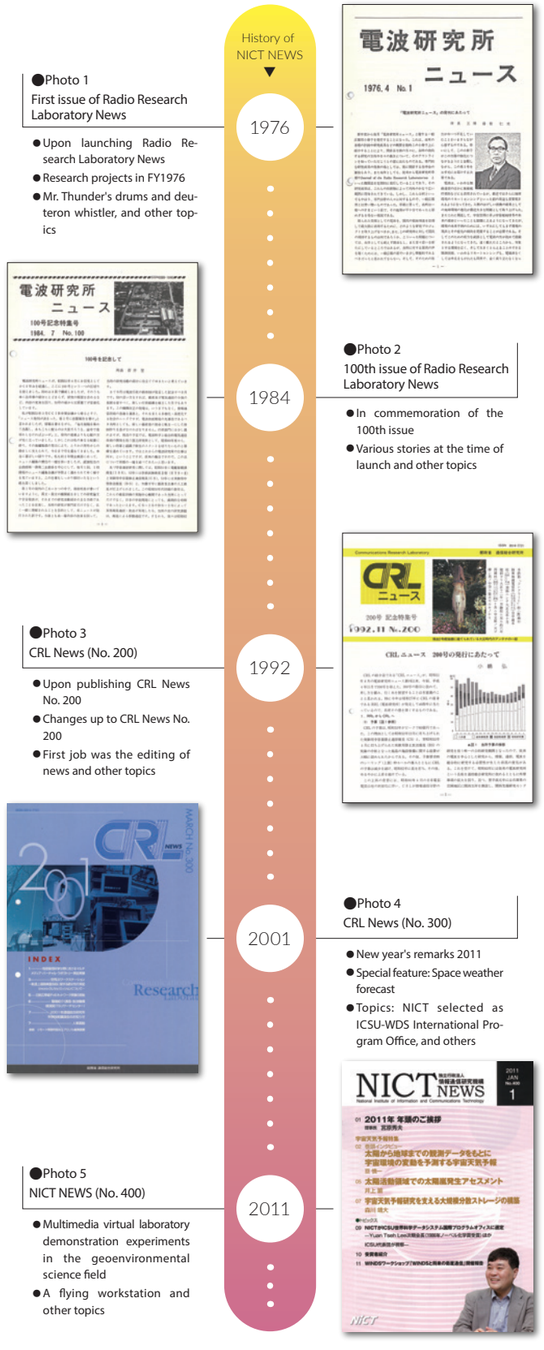
It was July 1984 that the 100th issue of the Radio Research Laboratory News was published (Photo 2). During this year, satellite broadcasting was commenced in Japan. In the Olympic Games held in Los Angels in July, YAMASHITA Yasuhiro, who is currently the JOC president, won the gold medal in Judo (open-weight category), and Carl Lewis of the United States won the gold medals in the men's 100 meters and in the long jump.

The 200th issue was published as CRL News of the Communications Research Laboratory, the Ministry of Posts and Telecommunications, in November 1992 (Photo 3). This issue published immediately after the bursting of the bubble economy already warned about young people's lack of interest in studying science and engineering as a serious problem.

The 300th issue was published in March 2001 in color mode, while changing size from the conventional B5 size to A4 size (Photo 4). It introduced demonstration experiments in which multiple research institutes connected with a network observed the global environment as virtual laboratories. During this period of time, the Ministry of Posts and Telecommunications became the Ministry of Internal Affairs and Communications through the central government reform, and the 300th issue stated that CRL would start afresh as an incorporated administrative agency from the new fiscal year starting from coming April.

The 400th issue was the New Year's edition for 2011 of NICT NEWS (Photo 5). This issue featured a space weather forecast, which continues to be one of the important missions of NICT. On March 11 of this year, the Great East Japan Earthquake occurred and caused tremendous human and economic damage, which is still fresh in our minds.

During this half century, social circumstances have changed and the forms and names of the organization have also changed. However, the mission of NICT NEWS to communicate the latest research outcomes in an easy-to-understand manner will never change. We hope that you continue to enjoy reading NICT NEWS.



Multiple Sound Spot Synthesis System: a Next-generation Technology Capable of Simultaneously Synthesizing Different Sounds in Different Zones

OKAMOTO Takuma
Senior Researcher, Advanced Speech Technology Laboratory, Advanced Speech Translation Research and Development Promotion Center (ASTERC), Universal Communication Research Institute

Biography
1981 Born in Numazu City, Shizuoka Prefecture
2004 Earned a bachelor's degree in Electronics and Mechanical Engineering from Tohoku University
2006 Completed a master's program in System Information Sciences at Tohoku University's Graduate School of Information Sciences
2009 Earned a Ph.D. in System Information Sciences from Tohoku University's Graduate School of Information Sciences
2009–2012 Worked as a postdoctoral researcher at Tohoku University
2012 Joined NICT
2020 Current Position
2003 Appointed as a Visiting Associate Professor, Tohoku University

Awards
2012 32nd Awaya Prize Young Researcher Award
2018 57th Sato Prize Paper Award
2022 9th Society Activity Contribution Award

Q&As

Q What is good things to be a researcher?

A As a researcher, I'm thrilled when my research proposals/manuscripts are accepted. I would be very honored if my research publications could impact the research community around the world and if my research achievements could benefit the public. Aside from work, I enjoy having drinks with other researchers and collaborators.

Q What are you interested in other than research?

A Jogging 10 km is my daily routine. Research ideas often come to me while running. I'm very happy that I can finally go out for drinks with my colleagues and friends again after a three-year hiatus due to the COVID-19 pandemic.

Q How do you spend your time on holidays?

A I jog and take my child to soccer practice in the morning. My family and I then go shopping in the afternoon. I often join my neighbors as they play with their children in a nearby park. I also enjoy eating out and going on overnight trips with my family.

When voices or audio recordings are played through loudspeakers, they are propagated in all directions. This causes interference when multiple voices and/or audio recordings are produced simultaneously, making it difficult for listeners to identify specific sources. To address this issue, we have developed a multiple sound spot synthesis system, consists of a circular array of 16 loudspeakers able to simultaneously synthesize different sounds in different, targeted areas. This allows sound transmitted to non-target areas to be cancelled out using antiphase sound signals. In addition, this system is expected to be able to simultaneously deliver different audio recordings in different directions without interference.

I previously carried out two projects—one on multilingual speech synthesis and another on sound field control—funded by the JSPS KAKENHI Grants-in-Aid. I am now integrating the results of these two projects with the goal of putting the multiple sound spot synthesis system into practical use. This project has been conducted in conjunction with PoCC launched as part of NICT's IDI. HIKITA Keita (NICT Europe Center), SUDO Miyuki (NICT General Affairs Department) and KUWAHARA Manae (currently working at an NTT facility in USA)—young NICT general staff members involved in the PoCC project—have also been giving me a help with this technology integration effort.

I've made progress achieving practical use of the multiple sound spot synthesis system through many events. These include on-site demonstration of the portable multiple sound spot synthesis system, acquisition of external funding, construction of a website designed to publicize the system, usability testing at the National Museum of Emerging Science and Innovation, exhibition and demonstration of the system during the general conference of the Institute of Electronics, Information and Communication Engineers (exhibitor: ISHII Kensuke, NICT ASTREC), exhibition and demonstration of the system during the G7 meeting in Gunma Prefecture, (exhibitor: SORA Wakana and TANIOKA Daisuke, NICT Public Relations Department) permanent exhibition at NICT headquarters (exhibitor: YASUI Yuri, NICT Public Relations Department) and invited talks. I'm pleased to have received many positive comments from the general public at these events. I'm now working to integrate the real-time multilingual speech translation system developed by ASTREC and the multiple sound spot synthesis system. This integrated system was demonstrated at a NICT Open House event in June 2023.

I will continue my efforts to put the system into practical use in collaboration with other researchers and cooperators.

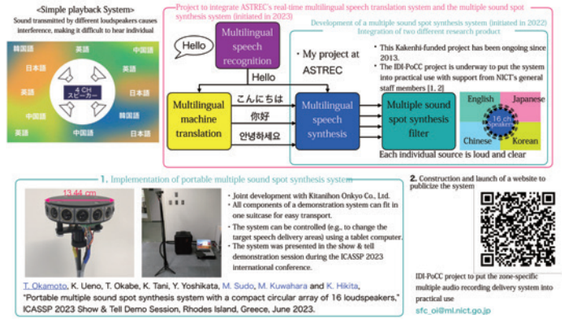


Figure Multiple Sound Spot Synthesis System

NICT researcher receives 2023 Japan Prize

NICT Principal Researcher HAGIMOTO Kazuo and his research collaborator, NAKAZAWA Masataka (Professor Emeritus, Tohoku University), received the 2023 Japan Prize. The prize ceremony was held at the Imperial Hotel Tokyo on April 13, 2023 and was attended by the Emperor and Empress of Japan.

Photo provided by The Japan Prize Foundation

- **Receiver: HAGIMOTO Kazuo** Principal Researcher, NICT (Program Director of Advanced Research in Information and Communication)
- **Area of Award: Electronics, Information and Communication**
- **Awarded for**
  Significant contributions to the development of long-distance, large-capacity optical fiber networks, mainly through the development of a laser diode pumped optical amplifier

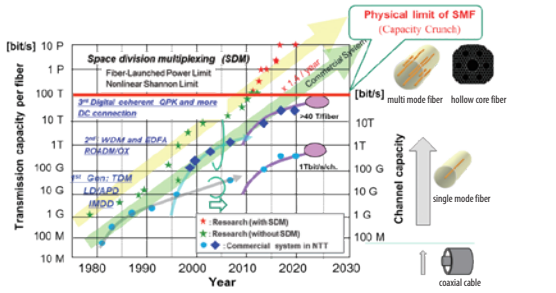
Receiver's comment Optical fiber communications technology is basically composed of optical fibers, laser diodes, photodetectors and optical amplifiers. I was involved in the development of an erbium-doped fiber amplifier (EDFA), a very compact optical amplifier designed to achieve broad-band transmission, noise reduction and high signal output without oscillation. As transmission signals enter an EDFA, the input optical signal (1.55 μm wavelength range) is coupled with excitation light (1.48 μm or 0.98 μm in wavelength) before passing through erbium-doped fibers. An EDFA is a very simple optical circuit equipped with an optical multiplexer/demultiplexer and optical isolators. Its greatest advantage is its excellent connectivity and compatibility with optical fibers through which optical signals are transmitted.

In 1988, ITU adopted the synchronous digital hierarchy (SDH) as a global standard technology for time division multiplexed transport networks. Consequently, an optical transmission rate of 2.5 Gbit/s was employed across the globe, including Japan. The intensity modulation/direct detection (IM/DD) format has since been used in the modulation and demodulation of optical signals within tiny devices, including laser diodes and avalanche photodiodes (APDs). At university, my research aimed to achieve use of the optical wavelength range in communications in addition to other wavelengths. At the time, I thought it would be impossible for laser diodes to generate 200 THz optical waves, which have a spectral purity as high as microwaves generated by microwave oscillators. I believed that the newly adopted IM/DD format could be used to reduce sensitivity to phase fluctuations, make the particle properties of light very effective and easily improve the performance of laser diodes. Based on previous cost and effect studies, it was decided that the transmission capacity of next-generation terrestrial communications should quadruple. However, achieving this target was expected to be very challenging because of the requirement that existing communications relay infrastructure—including telecommunications stations and optical fiber devices—had to be used without updating it. Quadrupling transmission capacity using the existing facilities required improving the fundamental S/N ratio to the minimum of 6 dB. Some researchers started presenting their results at scientific meetings to convince others that optical coherent communications systems could be a viable alternative to the existing system. I had been unproductively struggling to reduce thermal noise at the receiver side using a laser-diode amplifier. This motivated me to seek collaboration and outside expertise regarding the erbium-doped fibers. I began experiments in coordination with Professor Nakazawa's research team. During these efforts, modifications I had previously made to improve a laser diode amplifier led to a major breakthrough. In the subsequent optical transmission experiment—the world's first of its kind—I set a new coherent transmission record (the February 1989 Optical Fiber Communication Conference). A frequency division multiplexed EDFA with a frequency range approaching 10 THz divided into 80 channels at 100 GHz intervals has since been put into practical use.

This EDFA with digital signal processing capability is able to correct not only phase fluctuations induced by a laser diode but even transmission distortion. 40 Tbit/s signals can be relayed using only a single EDFA unit in combination with a multilevel transmission scheme. The use of EDFAs has enabled energy-efficient long-distance transmission (10,000 km across the seafloor and 1,000 km or more on the ground) with only minimal delay (figure). This is a tremendously beneficial achievement both technologically and economically.

I spent a majority of my time researching peripheral technologies that could bring out the high-performance potential of erbium and take advantage of its outstanding physical properties. This effort was made in collaboration with many other Japanese and overseas scientists and the Japanese government. When I completed the development of EDFAs in 1995, achieving an optical transmission rate of 10 Gbit/s, many people in Japan and elsewhere doubted the usefulness of such large transmission capacities. I am now glad to have created this technology.

History of R&D of ultra large-capacity optical communication systems



Fortune: Almost three generation (30 years) of history, folks has chattered "You're still on it". However I was supported by real demands.

The Maejima Hisoka Award was established in memory of Mr. Maejima, one of the founders of Japan's telecommunications industry, and to pass on and promote his spirit. The award is presented to those who have made an outstanding contribution to the advancement of the information and communications industry (including postal services) or the broadcasting industry. One group of NICT researchers received this award in FY2023.

Tsushinbunka Association 68th Maejima Award

OTOMO Akira Senior Expert, Advanced ICT Research Institute/ Director of Nano-scale Functional Assembly ICT Laboratory, Kobe Frontier Research Center, Advanced ICT Research Institute

YAMADA Toshiki Senior Researcher, Nano-scale Functional Assembly ICT Laboratory, Kobe Frontier Research Center, Advanced ICT Research Institute

KAJI Takahiro Senior Researcher, Nano-scale Functional Assembly ICT Laboratory, Kobe Frontier Research Center, Advanced ICT Research Institute

From left, KAJI Takahiro, YAMADA Toshiki, and OTOMO Akira

- **Awarded for: Research and development of organic electro-optic polymers and ultra-high-speed light control technology**
- **Date: April 6, 2023**

Receivers' Comment: We are very honored to receive the Maejima Hisoka Award. We appreciate the recognition it brings to our long efforts, including the development of organic electro-optic polymer materials and practical ultrahigh-speed light control devices. Our heartfelt thanks go out to the many people who have helped and supported our research activities. A series of fundamental technologies we have been developing for many years has finally come together to create products with potential practical applications. This has led us to launch joint research projects with many private companies. We will work with these companies to put products into practical use.

NICT NEWS 2023 No.4 Vol.500
Published by Public Relations Department, National Institute of Information and Communications Technology
Issue date: Jul. 2023 (bimonthly)
4-2-1 Nukui-Kitamachi, Koganei, Tokyo
184-8795, Japan
TEL: +81-42-327-5392 FAX: +81-42-327-7587
URL: https://www.nict.go.jp/en/
@NICT_Publicity
#NICT
ISSN 2187-4050 (Online)