

# A Translation Add-in for Microsoft Office for Easy-to-use Machine Translation in Business Scenes

Service Innovation Department Takashi Shimizu Yuki Chidiwa Takaya Ono

To achieve communications that transcend the language barrier, NTT DOCOMO provides “Hanashite Hon’yaku,” “Mail Hon’yaku,” “Utsushite Hon’yaku,” “JSpeak” and “Tegaki Hon’yaku” consumer-oriented translation services. NTT DOCOMO also provides “Hanashite Hon’yaku for Biz” as a translation service aiding businesses in the support they provide to overseas visitors to Japan. The services have been designed for use in all aspects of everyday life and travel situations.

However, as globalization of business continues, opportunities and demands to exchange business documents and e-mail created in languages other than one’s own native language are increasing.

To translate such documents, mostly (1) the services of a translator are employed, or (2) a Web-based machine translation service is used.

- (1) is time consuming and involves monetary costs.
- (2) does not take much time and is low-cost, but at least involves uploading files from a browser or copying and pasting the text to be translated into a text box in a browser. Thus, operating both a word processor application and a browser can be an involved task.

In addition, documents in business scenes are characterized as follows.

- (a) With application forms or slides etc., not only the text but the layout may also have meaning.
- (b) Creating e-mail may entail frequent additional corrections to text and so forth.
- (c) The wide range of business types and occupations may require specific styles and terms used for translation.

Hence, machine translation in business must (a) not affect layout, (b) have functions to aid text creation and (c) have excellent translation accuracy<sup>\*1</sup> matched to the type of business or occupation.

In light of these circumstances, to achieve easy-to-use machine translation of business documents and e-mail in foreign languages, NTT DOCOMO and Mirai Translate, Inc. jointly studied an add-in<sup>\*2</sup> for Microsoft® Office<sup>®\*3</sup>, a well-known software package widely used for business document and e-mail creation. This add-in was developed and is sold by Mirai Translate [1].

Combining the add-in and machine translation software, we have included operations required for

©2017 NTT DOCOMO, INC.

Copies of articles may be reproduced only for personal, noncommercial use, provided that the name NTT DOCOMO Technical Journal, the name(s) of the author(s), the title and date of the article appear in the copies.

\*1 Translation accuracy: In this article, “translation accuracy” refers to the combination of degree of detail conveyed in the translation, the fluency of the translation, and the degree of adequacy of the translation to the target field (style, suitability of specialist terminology etc.).

\*2 Add-in: A program to add a function to an existing software application.

\*3 Microsoft® Office®: A trademark or registered trademark of Microsoft Corp. in the United States and other countries.

translation in Microsoft Office applications to shorten the time it takes to acquire a translation.

In this article, we describe the add-in functions we have developed, and methods of improving the accuracy of translation.

1) Add-in Functions

Figure 1 shows how the add-in is included in Microsoft® PowerPoint®\*4.

As shown in fig. 1, when the add-in is installed, a tab called “Mirai Translation” containing buttons necessary for translation is added to the ribbon. Here, the user can specify the source and target

languages for translation (Original and Translation) and a translation model, discussed later.

After specifying the languages and translation model, the user can specify translation targets from three patterns - entire document, selected page only, or selected statements only. With translation of entire document or selected page only, the results of translation are output to a file separate from the original.

Figure 2 describes the flow of translation

\*4 Microsoft® PowerPoint®: A trademark or registered trademark of Microsoft Corp. in the United States and other countries.

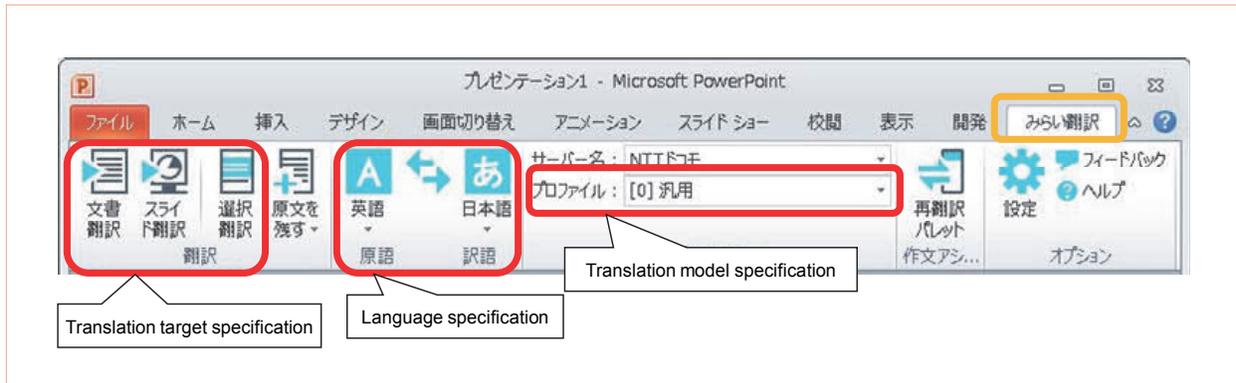


Figure 1 Add-in deployment image

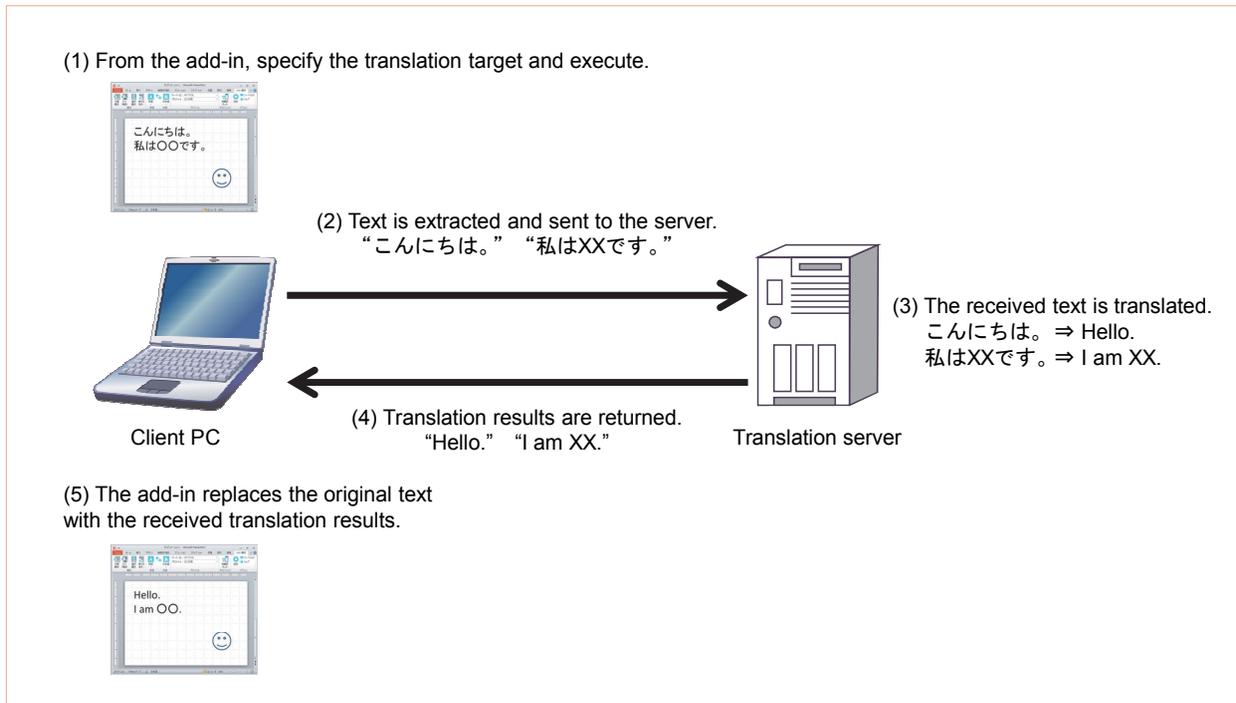


Figure 2 Translation process flow

processing with this add-in. When the target for translation is specified, processing begins (Fig. 2 (1)). The add-in extracts the text in the document and sends it to a translation server (Fig. 2 (2)). The translation server translates it with the specified languages and the selected translation model (Fig. 2 (3)), and returns the results of translation to the add-in (Fig. 2 (4)). When the add-in receives the translation results, it replaces the original text with the translation results (Fig. 2 (5)). This processing enables translation while maintaining formats such as complex document layouts and bold text etc. Also, it is possible to keep the original text with selected statements only as the target for translation.

As well as providing machine translation for the text entered, the add-in also has assistance functions for creating e-mail (in Microsoft® Outlook®\*5) that list example statements preregistered in the add-in that are similar to the text entered. If the list contains a useful expression, the user can write by selecting the example statement, without using the machine translation (Figure 3).

## 2) Improving Translation Accuracy

Because this add-in is designed for business documents and e-mail, we made improvements to specialize in these areas.

- (1) One is leveraging the aforementioned translation model. A translation model gives probabilities to combinations of words and their translation, which is adjusted according to the type of statements. For example, it is possible to improve the accuracy of translation compared to generic translation by selecting a translation model that handles unique expressions used in specialized documents such as patents. In addition, it is also possible to further tune translation models to suit an individual company using bilingual text owned by the company and the usage log of the add-in.

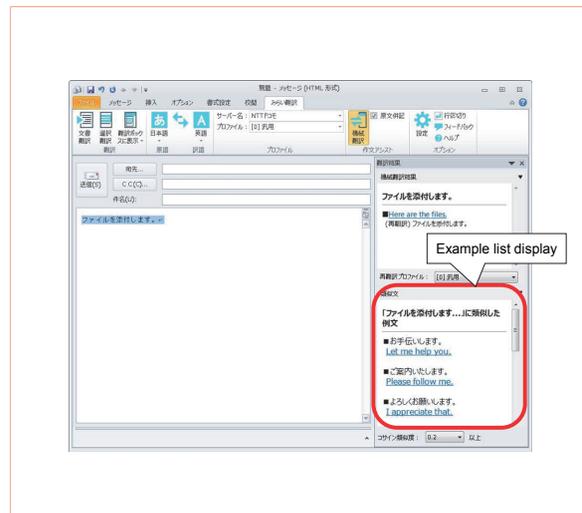


Figure 3 Assistance for creating statements in Outlook

- (2) Other improvements are the use of dictionaries in which product and service names, and industry-specific terms etc. are registered, and the registration and use of fixed phrases if such wording exists. Since such registered terms and phrases include many business and industry-specific terms and phrases, there is potential for unknown words\*6 or mistranslations occurring, which we have attempted to rectify with registration in dictionaries and lists of fixed phrases.

This article has described the functions of a Microsoft Office translation add-in designed for business documents and e-mail, and improvements to translation accuracy. Moving forward, we will take initiatives to further improve convenience and translation accuracy.

## REFERENCE

- [1] Mirai Translate: “Product introduction | Mirai Translate | Innovative machine translation,” (in Japanese). <https://miraitranslate.com/products/office>

\*5 Microsoft® Outlook®: A trademark or registered trademark of Microsoft Corp. in the United States and other countries.

\*6 Unknown word: A word not in a translation model or dictionary. These are often proper nouns.