

## **Dialect Research at the National Institute for the Japanese Language<sup>1</sup>**

Takuichiro ONISHI

### **1. Dialect research and the National Institute for Japanese Language**

Since its foundation in 1948, Japanese dialects have been considered to be one of the fields of study at the National Institute for Japanese Language (henceforth NIJLA), and they continue to be the subjects of research through to the present day.

The goals of the dialect research carried out at the NIJLA can be summarized as follows: (1) the explication of the dialects of every region; (2) the clarification of the nation-wide dialectal distribution; (3) the explanation of the process by which standard Japanese arose; (4) the understanding of the influence of standard Japanese on the various dialects, and the actual penetration of standard Japanese in all regions; (5) the presentation of Japanese language dialect materials for educational purposes. In this paper I will introduce that of our research which focuses on (1), (2), and (3).

As you may well know, the Japanese language has many dialects. If one broadly classifies the dialects of Japan, one comes up with the regional classification shown in Figure 1, and there are indeed also further sub-divisions. It would be impossible for the limited number of researchers at the NIJLA to go out into the field and investigate each and every one of these dialects ourselves, and consequently the NIJLA has undertaken the investigation and analysis of the dialects through a system of regional researchers who work in cooperation with the Institute. This cooperative research with non-NIJLA researchers is a characteristic of dialect research at our Institute.

### **2. Descriptive Research**

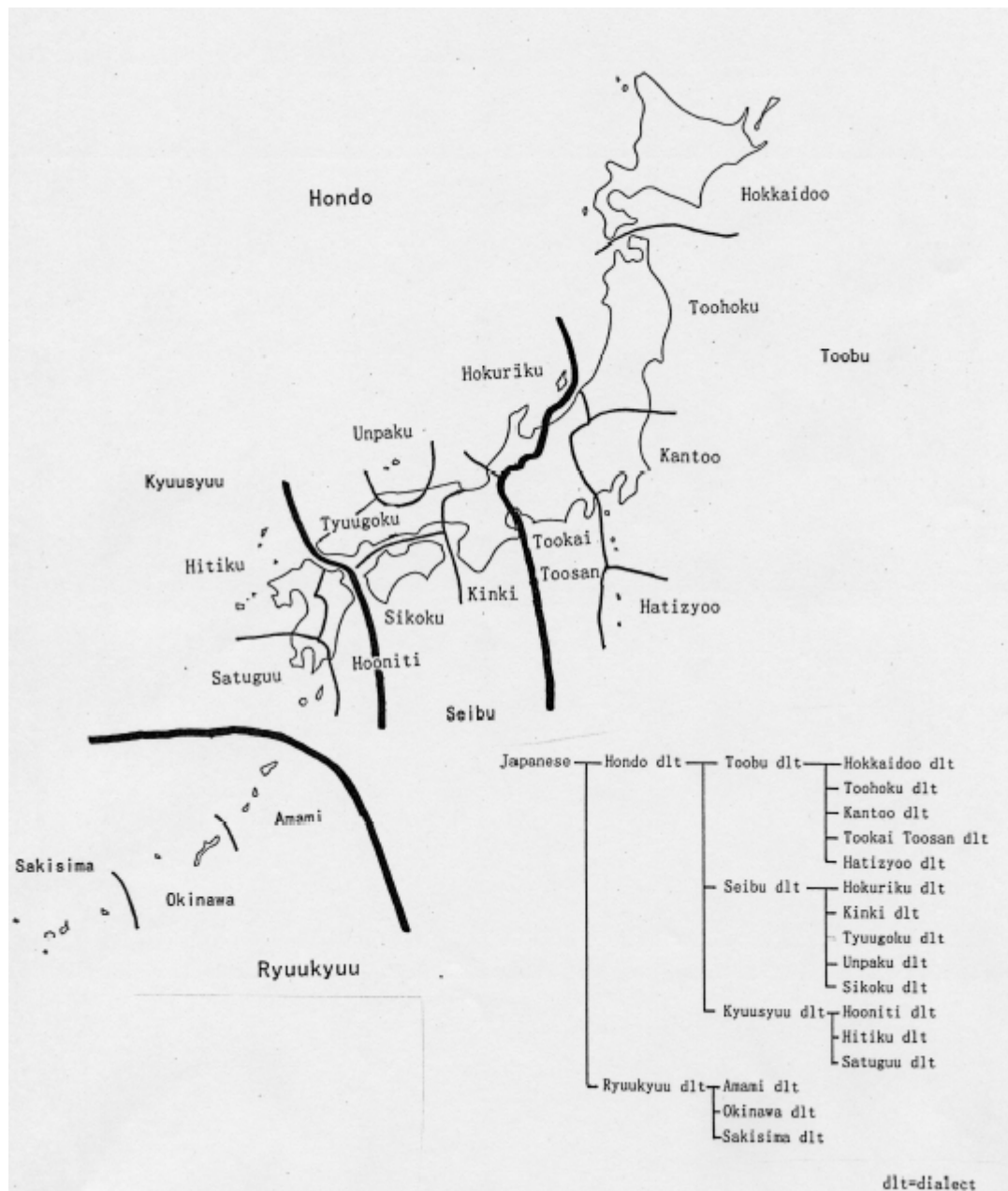
In order to grasp the entirety of the situation as it pertained to the Japanese dialects, the NIJLA, in its early stages, commissioned regional researchers to describe and report on the dialects of each region. This produced quite an amount of data. Twelve locations were chosen from this, a further three locations were added by NIJLA investigators, and the Institute produced a one-volume report of these fifteen locations, 'Descriptive Studies of Japanese Dialects' (1959). It was organized so that one could glean a general

---

<sup>1</sup> I would like to thank Peter Hendriks of the Australian National University for help with my English.

## Dialect research at the National Institute for the Japanese Language

understanding of the situation in each area as it pertained to phonology and syntax (primarily conjugation).



**Figure 1: Regional classification of Japanese dialects**

### 3. Research on the Ryukyuan dialects

As can be seen from Figure 1, when one classifies the Japanese dialects, it is the Ryukyuan dialects which are first to be differentiated from those of the mainland. Up until about 120 years ago, the Ryukyus were a kingdom and an independent nation. Partly because of this independence, the language of this region is sometimes generally known as Ryukyuan. If one regards them from an historical linguistic point of view, it is clear that the Ryukyuan dialects belong to the same lineage as the mainland dialects, but if one looks at them synchronically they differ greatly from the mainland dialects.

Furthermore, the Ryukyuan region is made up of numerous islands, and so there exists many dialects within Ryukyuan, each of which differs markedly from the others. The one which had attained the position of the standard in Ryukyuan was that of Shuri, the capital of the Kingdom of the Ryukyus. Shuri is in the Okinawa dialect region in Figure 1, and now forms a part of Naha, the central region of Okinawa Prefecture. 'A Dictionary of the Ryukyuan Language' (1963) was compiled in order to describe the dialect of Shuri. This dictionary is based on a manuscript compiled by Shimabukuro Seibin, a scholar of literature and the arts who was born in Shuri. It recorded about 15,000 words of the dialect. The NIJLA further organised this manuscript by writing out the items in phonetic transcription, adding accent marks, as well as appending notes on the phonology and the grammar.

The 'Descriptive Studies of Japanese Dialects' mentioned above did not include the Ryukyuan region. Consequently, with the publication of this dictionary the descriptive studies of the Japanese dialects carried out by the NIJLA can be seen to have attained a nation-wide coverage.

### 4. Linguistic Atlas of Japan

The dialects of Japan can be classified as shown in Figure 1, but this does not necessarily mean that the distribution of individual words will follow this classification. Each word has its own history, and can be transmitted across dialect boundaries. This is the same situation as pertains for the dialects of the languages of Europe. In order to show how various words are distributed, one needs to plot them on a map.

By producing a linguistic atlas, one can expect geolinguistic answers to questions such as what sort of geographical background the standard language has, or, what sort of history each individual word has. The 'Linguistic Atlas of Japan', six vols., (1966-1974) was produced to answer such questions. It covered 2,400 survey sites, and for each location it had, in principle, one male born before 1903 acting as consultant. There are 285 questions in the survey, many of which deal with vocabulary items. From 1955 a two-year preparatory survey was conducted, and then the whole nation was surveyed, with the cooperation of regional researchers, for the eight years from

1957 until 1964.

The results of the survey were transcribed on note cards -- one item per card -- and reported to the NIJLA by the investigator in each region. The NIJLA then organised and classified these cards, and drew the maps on the basis of this classification. The maps use at most eight colours, and use symbols to indicate the dialect words forms at the various locations. Figure 2 is an example of such a map. There are 300 maps in all in the atlas, and the actual size of each is 53.5 cm in height by 73 cm in width.

The 'Linguistic Atlas of Japan' came to take the lead in geolinguistic research in Japan. In particular, cartographical methodology such as the replacement of word forms with symbols, and the use of symbols designed to make distributions easy to identify had really not been hitherto recognised in Japanese dialectology. The Belgian-born priest W. A. Grootaers (dec. August 1999), who collaborated in the compilation of the atlas, played a major role in the introduction of such methods. Further, the publication of the atlas served to stimulate somewhat of a boom in geolinguistics, with the carrying out of distinctive surveys throughout the country, and the compilation of numerous linguistic maps.

Finally, the 'Linguistic Atlas of Japan' was subsequently reprinted from 1981 through to 1985, in a reduced format of 35.2 cm in height and 48.4 cm in width, with errors that were in the first edition being corrected.

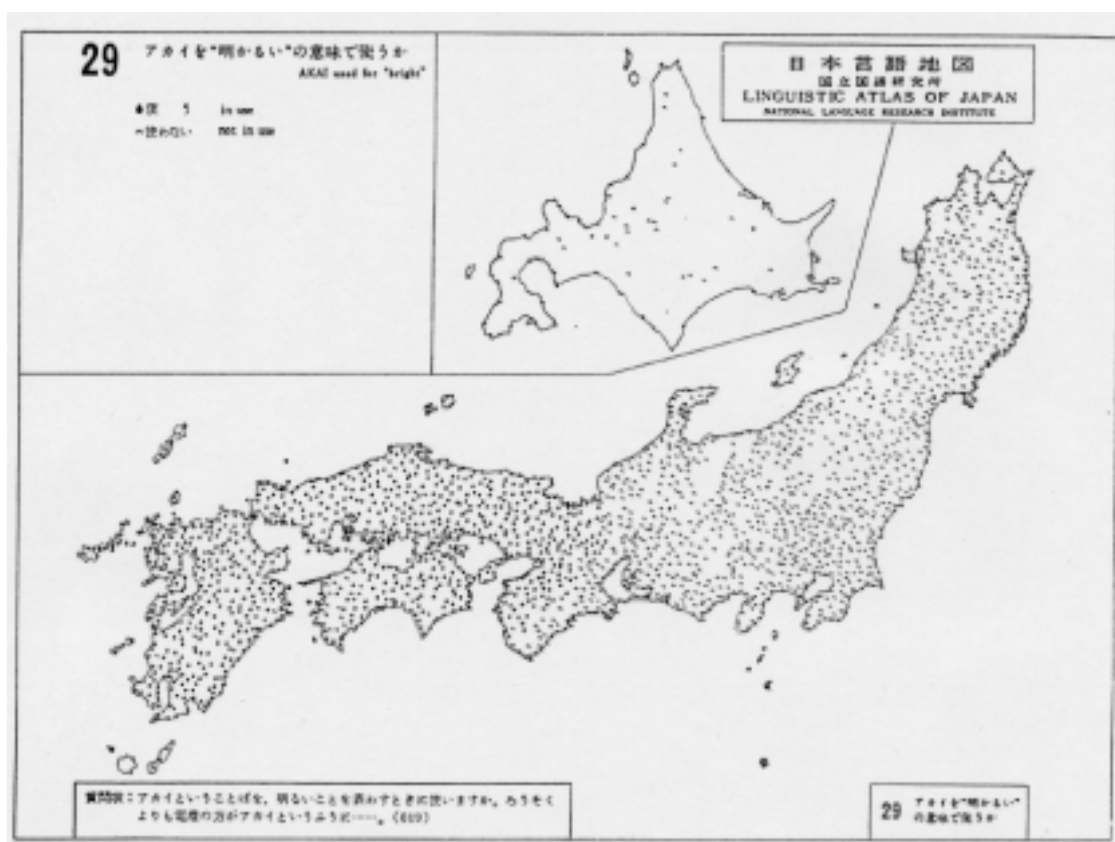


Figure 2: A map from the *Linguistic Atlas of Japan*

## **5. Conversational Data**

The production of descriptive reports and linguistic atlases furthered the understanding of the Japanese dialects. However, there remained the problem that these works did not shed any light on how dialects were really being used in actual conversation, and this led to the publication of the ten volume 'Texts of Tape-Recorded Conversation in Japanese Dialects', (1978-1987). For this work, free conversations in free settings were recorded around the nation with the cooperation of regional researchers, and were published on cassette tapes together with their transcripts. This work is being used in particular to provide data for the analysis of dialectal sentence structure and of the relationship between setting and discourse, but as a result of the conversations being on tape, can also be used for phonetic analysis.

## **6. Grammar Atlas of Japanese Dialects**

The 'Linguistic Atlas of Japan' led the field of geolinguistics in Japan, but was not the first linguistic map to be drawn in this country. In the Meiji period, a Language Research Commission was established within the Ministry of Education, and in 1903 carried out a linguistic survey by the correspondence method. The results of this survey were published as the 'Phonetic Dialect Atlas' (29 maps) (1905), and as the 'Grammatical Dialect Atlas' (37 maps) (1906). Both are atlases whose maps are multi-coloured, and which use the method of having distribution regions coloured in. Subsequent to this, a second survey was carried out in 1908, but the data collected remained unpublished and ultimately were destroyed by fire in the Great Kanto Earthquake of 1923.

The purpose of the Language Research Commission's national survey and the mapping of the results obtained was that it be used as a reference for the establishment of standard Japanese. However, questions remain as to whether it adequately filled that role. While one may surmise that it did play an indirect role, it is thought that the improvement and spread of schools and the ever-increasing spread of the media displayed the greater power, as far as the establishment of the standard language was concerned. Of course, if one regards this work from a purely scholarly point of view one recognises considerable achievements such as its documentation of the border line between the eastern and the western dialects of Japan, and its enumeration of various items in the central language whose older forms continued to exist in the dialects. However, this project was not one whose primary purpose was purely scholarly, nor did it satisfactorily achieve its stated aims. From this standpoint it is difficult to say that the national survey carried out by the Language Research Commission was very successful.

For a long time subsequent to this, the 'Grammatical Dialect Atlas' was the sole

existing nation-wide collection of data on the distribution of Japanese dialectal grammatical features, as the 'Linguistic Atlas of Japan' contained but very few maps as far as grammar was concerned. This was because from very early on the goals of its survey were centred on vocabulary. There were voices at the time which expressed the opinion that the survey should be centred on grammar, but in the end these views were not adopted.

Consequently, if one wanted to look at the distributions of dialectal grammatical items, one had no choice but to rely on the 'Grammatical Dialect Atlas'. However, not only was the number of items covered small, but the maps were also of the type which had the regions of distribution coloured in, and it was thus difficult to use for detailed scholarly research. Furthermore, it had also been pointed out that there were problems of mismatch between the actual report and the information as presented in the maps. On top of that, the fact that as the years progressed it became increasingly difficult to obtain the maps only increased the problems.

In this state of affairs, in 1976 the NIJLA began planning a nation-wide survey which would put the focus on grammar, and in 1977-1978 carried out a preliminary investigation. This was followed by the actual survey in 1979-1982 of 807 localities around the nation, investigating 267 items with the cooperation of the regional researchers. The consultants were males born before 1923, in principle one from each location. The way that the data were presented to the NIJLA by the regional researchers was the same as it had been for the 'Linguistic Atlas of Japan'. The survey was followed by an investigation of how the maps were to be edited and the data arranged, and in 1989 Volume One of the 'Grammar Atlas of Japanese Dialects' was published. Volume One concerned itself with case markers. An example of a map is shown in Figure 3, though the size of the actual map is 35 cm high by 48 cm wide, and a 35 cm by 25 cm sized summary of the relations between the word forms and the symbols is attached to its right hand side, so the whole thing becomes rectangular in shape. As was the case for the 'Linguistic Atlas of Japan', there is a symbol representing the word form at each locality surveyed, and at most six colours have been used in the drawing.

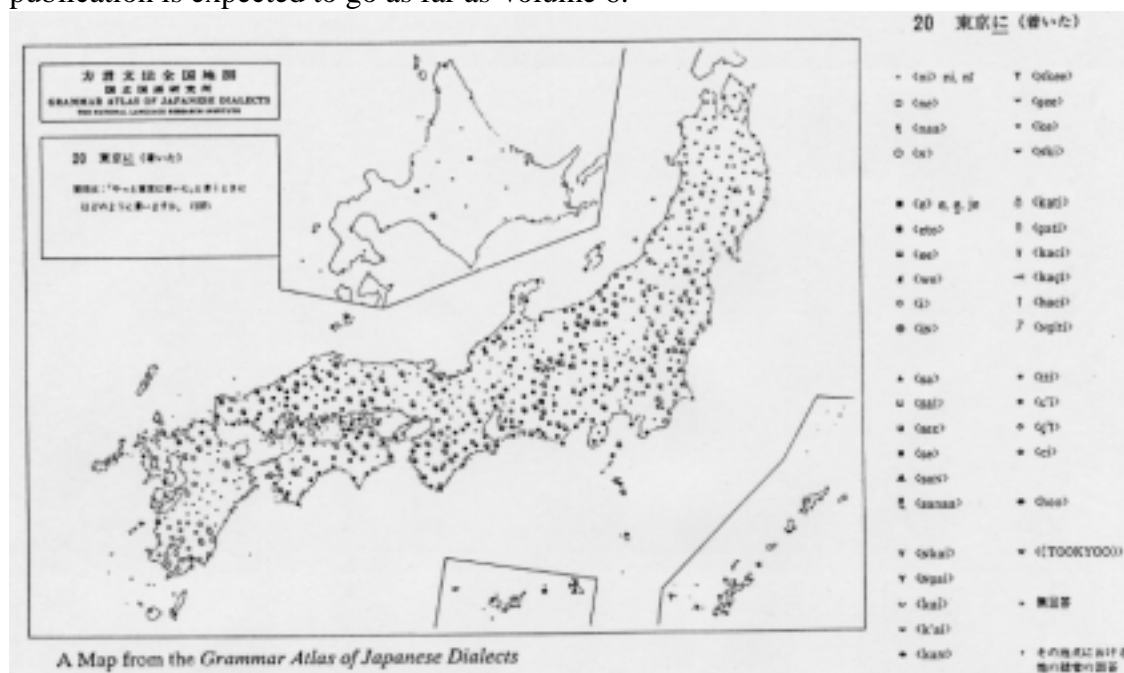
The number of localities surveyed in the 'Grammar Atlas of Japanese Dialects' has been limited to about one third of the number of the 'Linguistic Atlas of Japan'. This is due to the consideration that there would be fewer regional differences in grammatical phenomena when compared to vocabulary. On the other hand, detailed rules were established for the editing process concerning the adoption or non-adoption of word forms, the way in which word forms were to be organised, and how symbols were to be assigned. This was due to the realisation of the fact that in editing the 'Linguistic Atlas of Japan', the NIJLA was apt to rely overly on the arbitrary objectives of whomever was responsible. At the same time, a dialect material supplement has been appended to each collection of maps, enabling one to check what the survey data used in the compilation actually are. Furthermore, the data in the dialect material supplements have been prepared in an electronic format, and in principle are being

distributed freely.

The format of the maps and the process of compilation mentioned above are based on an editing policy which aims to produce data in the form of a collection of maps which, due to the objective presentation of the geographical distribution of the word forms for each locality, will robustly stand up to various analyses. In the adoption of word forms based on definite criteria, in the organisation of word forms through a set procedure, and through the clear indication of how the symbols are assigned, we have adopted the scientific method whereby the process of making the maps is reproducible. Consequently, the maps themselves do not represent a final historical interpretation, but are nothing more than objective indications of what kind of word forms are distributed where. The question of what sort of analysis is to be based on them is left to the reader. The reader may select or rearrange word forms as necessary. To be the source of basic data for this kind of work is the role to which the 'Grammar Atlas of Japanese Dialects' aspires.

Despite the fact that these were the goals on the basis of which the 'Grammar Atlas of Japanese Dialects' began publication, some criticism was levelled at it with respect to its removing from its central task the historical interpretation which was the aim of the 'Linguistic Atlas of Japan'. However, subsequent to this, understanding of the objectives of this collection of maps has deepened, and, while they have mostly been produced by the scholars involved in the compilation of the atlas, numerous analyses have appeared in print.

Volume 2 of the atlas, dealing with the conjugations of verbs, appeared in 1991. Volume 3 deals with the conjugations of verbs and adjectives and appeared in 1994, and Volume 4, which covers modals and further expressions, appeared in 1999. In all publication is expected to go as far as Volume 6.



**Figure 3:** A map from the Grammar Atlas of Japanese Dialects

## 7. Future Directions for Dialect Research

For both the 'Linguistic Atlas of Japan' and the 'Grammar Atlas of Japanese Dialects' the production of the maps themselves is carried out by hand. Basically, both atlases share the same production process whereby a map will be made by hand using coloured stamps, and this map will then be used as the basis for a further map made afresh on film at the printer's for the purpose of printing.

This procedure is a consistent method of operation, which, from the point of view of producing printed linguistic atlases, is well established and non-problematic. However, if one moves away from atlases as printed objects, one will find that it is not appropriate for situations, for example, where one would use the data from the atlas to produce a totally new map from a specific point of view and then publish this map in an academic paper. Similarly, the current methods of atlas production would not be suitable for the temporary production of a simplified map based on the data from the atlas in order to use this map in a specific educational setting. Furthermore, even as far as the production of the atlases is concerned, the current procedure for drawing the initial maps is dependent on a specific technician skilled in the art of manual cartography, and the fact that this technician is ageing is becoming a very real problem.

The solution for these sorts of problems would seem to lie in the development of a way to produce maps which does not depend upon a particular technical skill, or more concretely, in the development of a method of producing maps by using a personal computer. In particular, the data for the 'Grammar Atlas of Japanese Dialects', as I mentioned above, were converted into an electronic format before compilation into the atlas and would thus be relatively easy to manipulate with a computer.

At present, we are developing a plug-in program for a commercially available graphics software package, and are running tests on the computer generation of maps by having a text-type database read into the program. Specifically we are using the drawing software called Illustrator. As it is a drawing program it has the advantages that it will deal with enlargement or reduction of maps, and further that it easily converts maps to a browser-readable format.

In the future it will become possible, by connecting to the internet, to present Japanese dialect maps to the public not just in Japan, but throughout the world. Furthermore, if the plug-in program itself is made public, it will become possible for numerous people to draw linguistic maps without being reliant on having a specific technical skill, and this will be of great benefit both for research and for teaching.

Almost twenty years have passed since the surveys were completed for the 'Grammar Atlas of Japanese Dialects' which we are currently compiling. As we progress with the analysis of the data, the aspect in which we feel the greatest insufficiency is the amount of items surveyed. We repeatedly come across situations where we cannot obtain the information that we wish to know due to it not being among the survey items.

In previous nation-wide surveys the production of the dialect maps was always kept in mind. For this reason it was often the case that great attention was paid to the



## Dialect research at the National Institute for the Japanese Language

inclusion of items which were anticipated would display the clearest distribution possible. However, maps are ultimately no more than just one possible database output format. What is important is making sure that the database itself is complete.

I am sure that this holds throughout the world, but even in Japan too the finer regional dialect distinctions are disappearing. Traditional dialect data are disappearing upon the deaths of those who are in possession of them. We must record traditional dialect data covering a wide range of items as soon as possible. To this end we are aiming at the establishment of a complete database to be obtained through a new national survey, and as a first step the design of standard survey items is currently in progress.

We then wish to aim at having this database be as widely shared as possible, as is also the case for the linguistic atlas. In both cases, the trend for the future will undoubtedly be that these dialect data assets will not be monopolised by the NIJLA, but will be made available for anyone to consult freely and to analyse.

ONISHI Takuichiro  
National Institute for Japanese Language  
3-9-14 Nishigaoka, Kitaku, Tokyo 115-8620 Japan  
takoni@kokken.go.jp