

# SPOKEN-STYLE EXPLANATION GENERATOR FOR JAPANESE KANJI USING A TEXT-TO-SPEECH SYSTEM

*Yoshifumi Ooyama\**

*Hisako Asano\*\**

*Koji Matsuoka\*\**

\* NTT Communication Science Laboratories

\*\* NTT Information and Communication Systems Laboratories

E-mails: ooyama@nttkb.ntt.jp, {hisako, matsuoka}@nttnly.isl.ntt.jp

1-2356, Take, Yokosuka-shi, Kanagawa 238-03 JAPAN

## ABSTRACT

In this paper we describe a spoken explanation generator, **PLANET**,<sup>1</sup> for Japanese Kanji (ideograms), especially Kanji used in people's names. A number of text-to-speech systems for Kanji texts have been proposed but this is the first one that can explain Kanji characters so as to disambiguate characters from many homophone Kanji candidates. To accomplish this the generator explains the Kanji by using both internal composition and use in other words. The system has a database of over 6,000 Kanji characters that breaks down any given Kanji into its components and a text corpus of explanations for explaining two or more Kanji characters. It is capable of generating both other words that include the Kanji characters in question, and identifying information. Using these other words and the information the system makes phrases and sentences. Furthermore, this system generates natural prosodic information by classifying the pattern of semantic connections between words and phrases. The explanations are output through a natural-sounding voice synthesizer. Hearing examinations confirmed that this system achieves high accuracy in disambiguating Kanji characters from among many candidates. This system will make it possible to provide advanced and user-friendly human-computer interfaces.

## 1. INTRODUCTION

There are many Japanese text-to-speech systems that automatically convert original texts into synthesized but natural speech [1, 6]. In addition, there are a number of approaches for rewriting texts. One is to disambiguate homophones with similar meanings for a special purpose such as newspaper proofreading [2], and another is to convert on-line text into spoken text using words with many syllables or voiced sounds or polite expressions for words having the same meaning [3]. We have also developed a Japanese Kanji name input support system for Telephone/Telegraph operators (**TOPAZ**) that helps operators to input the numerous Kanji characters used in Japanese people's names [5]. This system has a Kanji information dictionary which includes Kanji components for

each Kanji character. This dictionary knowledge is used to analyze the input sentences.

The newly developed system, **PLANET**, described in this paper, uses the same information for text generation. This system can explain and speak the person's name automatically using the original readings and give additional explanation for recipients to choose between many Kanji character candidates. Unlike most English names, Japanese names have many homophones. For example, *Hiroshi*, one of the most popular Japanese given names, has over a hundred homophones. When Japanese explain their names over the telephone, they usually use two types of explanations. One is to use a well-known word, such as a famous person's name, that includes the character; the other is to describe the shape of the Kanji.

This system selects suitable words and generates explanatory phrases from information about these words and about the Kanji components. Furthermore, this system generates natural prosodic information.

In spoken explanations of Kanji, there are many unusual morphological connections. To generate natural prosodic information, the system does the following: (1) rewriting verb-like forms as noun forms for natural accentuation. (2) adding extra pauses in antonyms or sequences for audible and comprehensible information. The explanations are output through a natural-sounding voice synthesizer [1].

## 2. NAME-EXPLAINING METHODS

### 2.1. Explanation Patterns

When we transmit messages by audio signals alone, we must be careful with homophonic. In particular, when trying to tell someone about Kanji character compounds, sometimes not only the original reading but also the individual characters themselves need to be explained.

#### 1. Use general knowledge

When we explain person's names, we can often give a more effective explanation using general knowledge.

---

<sup>1</sup>spoken style exPLANation gENERaTor for kanji characters.

For example, we can say “*Futsu no // Ka'to desu* (the name *Kato* written with the usual Kanji characters)”. This kind of explanation can be used if the name is very common one and one set of Kanji characters are used far more often than any others that sound the same. For the Japanese pitch accent, we use an apostrophe to indicate the point where pitch changes from high to low and “//” to indicate the accent phrase boundary. Also, if the name uses the same Kanji as a famous person’s name, we can use this to explain the characters, e.g. “The same Kawabata as in Yasunari Kawabata (a Nobel Prize winner)”.

## 2. Identify the Kanji by its use in a well known word

This confirmation of individual characters is also used in English. We often say, “A[éi]” as in “America” to accurately transmit the character “A”. Figure 1 (a) shows the Kanji of *saku* “yester-”. There are more than ten<sup>1</sup> Kanji candidates that can be read *saku*. *saku* “yester-”, however, is unambiguous when used in the two character word *saku-jitsu* “yester-day”. Recipients can thus grasp the target Kanji, *saku* “yester-”, easily. In this case the spoken-style sentence is *saku'jitsu no // sa'ku* “‘yester-’ as in ‘yesterday’”.

## 3. Describe the composition of the Kanji

A single Kanji can be explained in terms of its components. There are very few similar cases in English, but one is “W” (i.e. “Double U”), which indicates the shape of “W”. Figure 1 (b) shows the Kanji for *anzu* “apricot” which is used by itself as a person’s name or as a part of the name *Kyôko*. This Kanji consists of two basic characters, which are *ki* “tree” and *kuchi* “mouth”. We explain this Kanji as “*ki' ni // kuchi*”. *Ni* means “and” or “plus”.

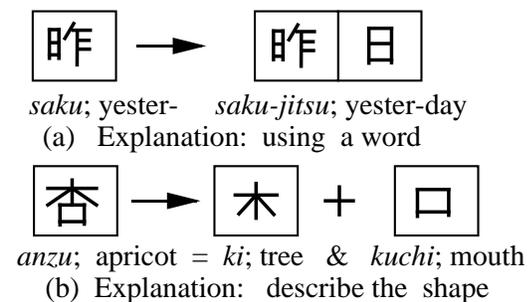


Figure 1: Kanji explanation method

## 4. Use attributes of the Kanji character itself

It also useful to use Kanji meanings for disambiguation. We can say “The *kaki* meaning persimmon” to distinguish the *kaki* which means “oyster” or “fence”. Another type of attribute is character sets. There are

many old-fashioned types of Kanji which are still used in person’s names, which have been replaced by simpler Kanji for normal words. For example, there are old and new forms of Kanji *sawa* “swamp”. In this case, we can effectively explain by saying “The old/new form of *sawa*”.

Combining these methods lead to a convenient way of explaining Kanji characters.

## 2.2. Explanation Generation

Figure 2 shows an overview of the PLANET system. Which explanation should be used to explain Kanji words that include the target Kanji are selected according to the following three criteria.

- (1) felicity : good meaning or neutral meaning words
- (2) familiarity : popular and familiar words
- (3) clarity : words that have few homonyms

Names, especially personal names, have a personality of their own. It is important to select positive-sounding words in a spoken dialog. We gathered these words from a Kanji information dictionary [5], a text corpus of about 700 personal names explained by the person in question, and a written text corpus of dialog between two people transmitting 4000 people’s names using sound only. Explanations using these words are stored in the text corpus. For example, the given name *Hiroshi*, is explained by *hakubutukan* “museum”, *yûfuku* “rich” or *hiroï* “wide”. These words satisfy the above three criteria.

**Sentence Generation** An explanation using a Kanji’s components (Fig 1 (b)) is generated by the readings of the Kanji components conjoined with *ni* “and” (a post-positional particle). We use the Japanese particle *no* “of” for genitive meanings. Additional and optional explanations can be useful, for example, to explain “\*” by saying “asterisk, star”. In this case we use the particle *de* that means apposition or *to-kaku/kaite* “which spells”.

There are also some Kanji characters that have the same components but with the components in different positions. In this case, the system uses a Kanji information dictionary<sup>2</sup> that indicates position information, so the system can expand these particles as *no-migi-ni* “on the right side of” or *no-shita-ni* “under”.

Verbs/Adjectives with dictionary entry inflection style are used to explain Kanji, but usually the entry style does not include an adjoining *ni*, so this system rewrites verb/adjective-like patterns to virtual nouns using quotation marks for

<sup>1</sup>Such as: fence, work, yester-, plan, vinegar, bloom, tear, split, search, cut, blend, squeeze, etc.

<sup>2</sup>We define about 600 Kanji components and positioning information [5]. This set of components is an expansion of the traditional Kanji radicals, usually 150 Kanji radicals are defined [4].

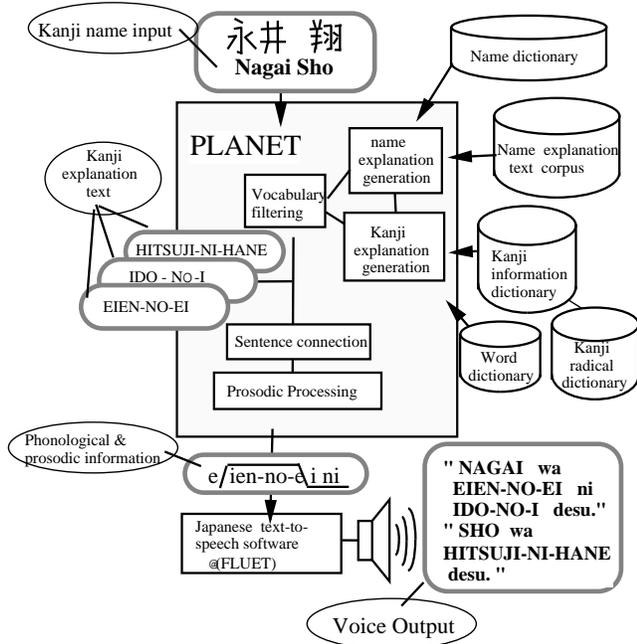


Figure 2: System configuration

better prosodic processing. A rewriting example is from “Hiroi(adjective)ni-kodomo” to “Hiroi(virtual-noun) -“ni”-“kodomo”.

**Vocabulary Filtering and Default Priority** We use filtering to select candidates for explanation. The first task is to avoid redundant explanations. For example, *sakujitsu no saku wa sakujitsu no saku* “‘Yester-’ of ‘yesterday’ is ‘yester-’ as in ‘yesterday’ ”; which is tautologous. If the explanation is in the context of explaining *sakujitu* “yesterday”, we can not use the word *sakujitu* “yesterday” except in apposition. In this case, *saku-nen no saku* “‘yester-’ of ‘yesteryear’ ” is better. Next, in order to avoid too detailed explanation, the priority sequence is 1, 2, 4, 3. with 3, the explanation using components, used only as a last resort. If there are two or more candidate explanations with the same priority, we select the shortest.

**Readings and Accent** As shown in Fig 1 (a) , to pick out a Kanji, usually information about its use in other words, ‘word information’, is used. ‘word information’ is the word readings and accent. ‘Kanji information’ depends on the word or the Kanji. If the Kanji is a common word, the popular word readings and accent can be generated by Audiotex [3]. On the other hand, if the Kanji is not a common word itself, the system decides the reading from the word readings. For example *gi* of *jo-gi* (ruler) is the same Kanji as *ki* of *ki-soku* (rule). Since voiced sounds are often changed in Japanese, a Kanji can be read *gi* or *ki* depending on the word involved. The Kanji particle accent falls just after the first mora. Other word accent rules depends on the program [3].

	Family names	Given names
Accuracy per Kanji	98.5%	96.8%
Error / all characters	59 / 3920	125 / 3880

Table 1: Accuracy of transmitted character

Recipient	A	B	C	D	E	F	G	H	I	J
Family names	7	8	6	4	4	3	9	10	5	3
Given names	14	14	14	9	7	17	17	20	8	5

Table 2: Errors for each recipient

**Pauses** This system generates natural prosodic information. To do this, it classifies the patterns of semantic connection between words in the corpus. The classification is done by using antonyms (fast vs slow; man vs woman ...) and parallel phrases/words. Japanese has pitch accents, which are stored in the dictionary. When two or more words are combined, the accent sometimes shifts to facilitate smoother speech [3]. However, if words are antonym pairs or sequences, an extra pause must be inserted between words to recognize them distinctly. For example, to confirm *sen* “thousand”, an enumeration is useful. *Sen* has many homonyms but enumerating is audible and familiar, for example *ichi, jū, hyaku, sen no sen* “thousand as in 1, 10, 100, 1000”. In this case, the system inserts pauses between numbers.

### 3. EXPERIMENT AND RESULT

We tested the system as follows:

**Method:** The subjects/recipients were ten people who usually use word processors but who had never heard the system’s voice before the examination. They heard the system read names and explain them automatically. They can have the explanations repeated, if necessary, but can use only the same explanations. They then enter the Kanji using Japanese word processors.

**System:** PLANET speaks a family name and given name and gives an explanation of each name. The examination is given for about 200 personal names per recipient.

**Target:** The target names were about 200 Japanese personal names selected randomly from a name list.

**Result:** We checked the original Kanji character and entered Kanji characters. We regarded omitted characters or redundant characters as errors.

Table 1, the average of accuracy per Kanji, shows the very high accuracy of Kanji transmission obtained. Given names usually have more candidates than family names, and this seems to be reflected in the results. Table 2 shows the error counts for each recipient. As can be seen, some recipients made three or more times as many errors as others. To avoid revising errors that are not the focus of this investigation, we

checked three or more recipients' overlapping mistakes for the same names. These errors, occurring over 22 people's names, are about 60% of all errors and seem to have identifiable causes.

After the experiment, we asked the recipients about the reasons for their errors, especially the overlapped errors. The major reasons for the errors are as follows:

- **Similar sounding words**

Words such as, *nichi* "day" and *michi* "unknown" or *ryuki* "elevation" and *yuki* "courage" have very similar sounds. These words caused some misunderstandings.

- **Homonyms used in the explanations**

For example, the word, *Ta' isho* used in the explanation, means a King, but *Taisho*, with a different accent pattern has another meaning. Some users had no faith in the accuracy of the synthesized voice, which resulted in mistakes.

- **Uncommon words or readings**

Different recipients can consider a given Kanji common or uncommon depending on their level of knowledge. The system has correct readings of Kanji components, but for some components, alternative names are suitable for a public use system. For example, people often say *ne-hen* or *o-hen* even though the 'official' readings are *shimesu-hen* and *tama-hen*.

Next, the reasons for confusion because of inaudibility are as follows:

- **Phrase parsing misinterpretation**

The reading combining the particles *Masa* and *ni* has the same reading as that of an adverb, *masa-ni* "exactly". This confused some recipients. Another example was *shin-jitai* "new form Kanji" and *shinji-tai* "want to believe" which also have the same readings.

- **Overly detailed or overly simple explanations**

Almost all Japanese know how to explain the Kanji of their own name, but there are some demerits to using this knowledge. Since the standardization for explanation is not perfect, it sometimes results in overly detailed or overly simple explanations. For example, a kanji may be broken down too far, into three components rather than two, and so on.

## 4. CONCLUSION

**PLANET**, a system for explaining Japanese Kanji using sound only, is described. It is capable of generating Kanji information to enable a hearer to disambiguate from numerous homonyms. Using a text-to-speech system, it achieves accuracy in transmitting the Kanji used in people's names to a human listener of about 96.8%-98.5% per Kanji character. Although in human-to-human communication many

misreadings can occur, we can resolve ambiguities by using other words in this case. The next step is to expand this system into a dialog system that knows about word inaudibility and user recipient models. This system will have many beneficial applications, such as editing information to or from visually or speech impaired persons, taking orders over the telephone, and providing automatic directory assistance services.

## 5. REFERENCES

1. Kazuo Hakoda, Tomohisa Hirokawa, Hajime Tsukada, Yuki Yoshida, and Hideyuki Mizuno. Japanese text-to-speech software based on wave form concatenation method. In *AVIOS*, 1995.
2. Satoru Ikehara, Tsuneo Yasuda, Katsumi Shimazaki, and Shin ichiro Takagi. Revision support system for Japanese verbal error(revise). In *NTT Electrical Communications Laboratories Technical Journal (in Japanese)*, volume 36, pages 1159-1167, 1987.
3. Koji Matsuoka, Eiji Takeishi, Hisako Asano, Ryomi Ichii, and Yoshifumi Ooyama. Natural language processing in a Japanese text-to-speech system for written-style texts. In *IEEE IVTTA*, 1996.
4. Andrew Nathaniel Nelson. *Japanesen-English Character Dictionary*. Charles E. Tuttle Company, 1974.
5. Yoshifumi Ooyama, Tsuneaki Kato, and Kenzi Imamura. Japanese Kanji name input system using spoken-style transcribed Japanese analysis. In *IEEE CAIA*, pages 2-8, 1995.
6. Yoshifumi Ooyama, Masahiro Miyazaki, and Satoru Ikehara. Natural language processing in a Japanese text-to-speech system. In *ACM Computer Science Conference*, pages 40-47, 1987.

## APPENDIX

1. **Sound examples** [SOUND A263S01.WAV] and [SOUND A263S02.WAV] show the explanation examples generated by the PLANET system, using a text-to-speech system [1, 3]. The first example is the name of a Japanese statesman, "*Kan Naoto*", and the second example is the name of a Japanese astronaut, "*Wakata Koichi*".
2. **Image data** [IMAGE A263G01.GIF] shows these Kanji and explanations for them.