

Sentence-Style Conversion of Japanese News Article for Text-to-Speech Application

Yukiko Hayashi¹ Shigeki Matsubara²

¹Graduate School of Information Science, Nagoya University, Japan

²Information Technology Center, Nagoya University, Japan

¹hayashi@el.itc.nagoya-u.ac.jp ²matubara@itc.nagoya-u.ac.jp

Abstract

This paper proposes a method of sentence-style conversion for generating spontaneous Japanese speech in a text-to-speech synthesis system. Since written language is different from spoken language linguistically, the speech by direct reading of written Japanese texts might be unnatural. The method takes a fully rule-based approach to convert the sentence style and to complement sentences, which are lacking some functional words or verbs. The method was evaluated using Japanese newspaper articles. The precision was 89.7%, and the recall was 86.7%.

1 Introduction

There are huge quantities of printed media such as newspapers, magazines, web sites, e-mail. At the same time, we are not always able to spend a great deal of time reading these vast amounts of texts, and therefore the method for efficiently surveying these texts is desired. For example, by listening speech generated from the texts while driving or walking, we can get the information efficiently.

In order to generate speech from texts automatically, we can use text-to-speech systems. Recent text-to-speech systems have been improved in quality of sound, and can produce speeches with almost human-like pauses and intonation. Between written language and spoken language, however, there exist many differences in vocabulary, phrases, and so on. This means that the speech directly generated from written-style texts might be unnatural.

This paper proposes a method of sentence-style conversion for generating spontaneous Japanese speech in a text-to-speech synthesis system. The articles in a newspaper are used as the objects of the conversion, for the reason

that we are planning to deliver spoken news generated from texts automatically.

Ohizumi et al. proposed a method of converting nominalized verbs as a part of their research on paraphrasing written language into spoken language (Ohizumi et al., 2003). They did not treat *taigen-dome*,¹ a Japanese-specific rhetoric, however, which we discuss in this paper. Kaji et al. proposed a method of paraphrasing written-language-specific vocabulary into spoken-language-specific vocabulary by using written and spoken language corpora collected from the Web (Kaji et al., 2004).

This study takes particular note of the importance of the differences between “*da-dearu*” style (written style) and “*desu-masu*” style (spoken style), and also works on “*taigen-dome*.” *Taigen-dome* sentences need to be complemented with an appropriate word. We take a fully rule-based approach to convert the sentence style and to complement *taigen-dome* sentences. The experiment was conducted on Japanese newspaper corpus, and the precision was 89.7%, the recall was 86.7%.

This paper is organized as follows: Section 2 discusses the differences between written language and spoken language. Section 3 describes a method of sentence-style conversion. Section 4 reports the experimental results.

2 Differences between Written Language and Spoken Language

In general, natural language is divided into written language and spoken language. Written language is used for text-based communications, and spoken language for speech-based communications. Table 1 shows the differences

¹ *Taigen-dome* is a rhetoric to cut off the words following a certain noun in order to end the sentence with the noun. The eliminated words, however, have to be inferable from the context. In Japanese, *taigen* means an indeclinable word including a noun and a pronoun, and *dome* means “to end,” hence *taigen-dome* approximately corresponds to “ending-with-noun.” By doing this, the writer can make an impression on the reader.

A newspaper article in written language

murayama-tomiichi-shusho-wa nento-no kisha-kaiken-de sozo-to yasashisa-no kuni-zukuri-no bijon-to daisuru shokan-o happyo-**shita**.
 kongetsu-chu-ni shusho-o kakomu gakusha-gurupu-ga happyo-suru murayama-bijon-no kihon-teki-na kangae-o shimeshi-ta **mono**.
 waga-kuni-ni fusawashi kokusai-koken-niyoru sekai-hewa-no sozo-to meiut-ta hi-gunji-bunya-no kokusai-koken-nado yottsu-no sozo-o uchidashi-te **iru**.
 shokan-wa boto-de sengo goju-shunen-no fushime-no toshi-no kyatchi-furezu-o kaikaku-kara sozo-e-to **hyogen**.

A newspaper article converted into spoken language

murayama-tomiichi-shusho-wa nento-no kisha-kaiken-de sozo-to yasashisa-no kuni-zukuri-no bijon-to daisuru shokan-o happyo-**shi-mashi-ta**.
 kongetsu-chu-ni shusho-o kakomu gakusha-gurupu-ga happyo-suru murayama-bijon-no kihon-teki-na kangae-o shimeshi-ta **mono-desu**.
 waga-kuni-ni fusawashi kokusai-koken-niyoru sekai-hewa-no sozo-to meiut-ta hi-gunji-bunya-no kokusai-koken-nado yottsu-no sozo-o uchidashi-te **i-masu**.
 shokan-wa boto-de sengo goju-shunen-no fushime-no toshi-no kyatchi-frezu-o kaikaku-kara sozo-e-to **hyogen-shi-te i-masu**.

(Prime Minister Tomiichi Murayama offered his view titled “The nation-building vision with creation and benevolence” at the first-of-the-year press interview. This view demonstrates basic policies of “Murayama Vision” which the prime minister’s brain trust will announce later this month. His view presents “the four creation” including a international contribution in the nonmilitary sector named “The creation of global peace by a international contribution suited to our country.” The view initially described the catch-phrase of this milestone year being the 50th anniversary of the end of the Pacific War as “From reform into creation.”)

Figure 1: Comparison between written language and spoken language.

Table 1: Differences between written language and spoken language.

	Written language	Spoken language
Sentence style	mainly <i>da-dearu</i> style	mainly <i>desu-masu</i> style
Length of sentence	rather long	rather short
Vocabulary	rather difficult	rather easy
Tone of the words	rather formal	rather informal
Diffuseness	low	high

between written language and spoken language (Yamamoto et al., 2003; Ohizumi et al., 2003).

In this paper, among the differences between written language and spoken language, we take up a sentence style and diffuseness. Note that we equate diffuseness with *taigen-dome* sentences. Figure 1 shows the comparison between written language and spoken language using an actual newspaper article. The differences in sentence style and the *taigen-dome* sentence are indicated by boldface.

2.1 Differences of Sentence Style

The differences of the sentence style distinguish spoken language from written language. *Desu-masu* style is important for listeners in receiving the speech naturally.

2.2 Taigen-dome sentences

Taigen-dome is a rhetoric to end the sentence with the noun as noted in the footnote in Sec-

tion 1. The following are examples:

- yoshi-wa tsugi-no **tori** (desu).
(The summary is as follows.)
- taisaku-no hitsuyosei-o **shiteki**(shi-mashi-ta).
(He indicated the need for countermeasures.)

Both of the example sentences end with a noun (“tori” and “shiteki”), and the words in parentheses (“desu” and “shi-mashi-ta”) are the eliminated words which can be inferred from the context information.

Taigen-dome sentences occur frequently in newspaper articles. They are intended to eliminate the diffuseness and condense the articles as well as to leave the readers with an impression. On the other hand, a *taigen-dome* sentence in speech is unnatural because such a sentence sounds like it was cut off halfway. In order to make it natural, we need to complement it with an appropriate word.

3 Conversion from Written Language into Spoken Language

3.1 Conversion of Sentence Style

We adapt a rule-based approach in order to convert written-style sentences into spoken-style sentences. The rules are based on word sequences and morphological information such as surface string, part of speech, and conjugation. The rules were made by analyzing 1130 sentences in articles on January 1st, 1995 in

the Mainichi newspaper text corpus. Table 2 shows the conversion rules and the respective examples.

3.2 Complement of *taigen-dome* sentences

In complementing *taigen-dome* sentences, there exist three kinds of points to be taken into account in.

(1) Type of the complementary word

As noted in Section 2.2, the complementary word of the *taigen-dome* sentence varies depending on the context. In order to determine the type of the complementary word, we focused on the subtype of the last noun of the sentence.

For example, in the sentence “yoshi-wa tsugi-no tori,” the subtype of the last noun “tori” is “general.” In the sentence “taisaku-no hitsuyosei-o shiteki,” the subtype of the last noun “shiteki” is “*sahen*.” The subtype “*sahen*” means that the noun is a verbal noun and can be followed by a verb “shi-masu” which is roughly the English equivalent of “do.”

Thus, if the last noun of a sentence is a “*sahen*” noun, the sentence is complemented by “shi-masu” or its past form “shi-mashita,” otherwise it is complemented by “desu” or its past form “deshi-ta.” In the sentence “kankoku-gawa-kara gutai-teki-na tean-ga aru yote. (South Korea will make a concrete suggestion.)”, however, although the last noun “yote” is “*sahen*” noun, the proper complement of the sentence is “desu.” Thus, even if the last noun is a “*sahen*” noun, a sentence including the word which can become the subject is complemented by “desu” or its past form.

(2) Tense

In addition to the type of the complementary word, we have to decide the proper tense of the word. In order to determine the tense, we use the particular auxiliary verb and the phrase which indicates the past form. A sentence which includes any of the words listed below is considered as the past tense.

- Auxiliary verb “*ta*” which indicates the past form.
- The phrases which indicates the past form such as “*sakunen* (last year),” “*san-nen-mae* (three years ago).”

Table 3: Experimental results.

Precision	Recall	F-measure
619/690 (89.7%)	619/714 (86.7%)	88.2

(3) Necessity of the complement

There exists a *taigen-dome* sentence which needs no complement. For example, a sentence “kanada-no ese-toshi-gun-no hitotsu richimondohiru. (Richmond Hill, one of Canadian satellite cities.)” needs no complement. If it were complemented by a word like “desu,” it would be considerably artificial. Therefore, if the subtype of the last noun in a sentence is not “*sahen*” and the sentence does not include the word which can become a subject, the sentence is considered to need no complement.

In consideration of the above three viewpoints, we made the decision tree as shown in Figure 2. The complementary word for each *taigen-dome* sentence is determined according to this decision tree.

4 Experiment

The articles on January 3rd, 1995 in the Mainichi newspaper text corpus were used as an experimental object. The articles consist of 687 sentences including 714 places to be converted. Of them, 559 are written-style (*dadearu* style) and 155 are *taigen-dome*. The rules in Table 2 were used to convert the sentence style. The decision tree in Figure 2 was used to complement the *taigen-dome* sentences. For morphological analysis, MeCab (Kudo, 2006) was used. Table 3 shows the precision and the recall of the whole experiment. The precision and recall were as follows:

$$Precision = \frac{\# \text{ of correctly converted places}}{\# \text{ of converted places}}$$

$$Recall = \frac{\# \text{ of correctly converted places}}{\# \text{ of places to be converted}}$$

The experimental results of sentence-style conversion and the results of *taigen-dome*-sentences complement are separately discussed below.

4.1 Results of Sentence-style Conversion

Among the 687 sentences in the article on January 3rd, 1995 in the Mainichi newspaper text corpus, there exist 559 places to be converted from written style into spoken style. Table 4 shows the results of sentence-style conversion. The precision was rather high. The only error

Table 2: Conversion rules and examples.

Conversion rule	Before conversion	After conversion
1. BF of a verb → CF of the verb + “masu”	risuto-ni-wa samazama-na uttae-ga narabu.	risuto-ni-wa samazama-na uttae-ga narabi-masu. (Various appeals are on the list.)
2. CF of a verb + BF of an auxiliary verb “ta” → CF of the verb + “mashi-ta”	sefu-suji-ga sanjuichi-nichi akiraka-ni shi-ta.	sefu-suji-ga sanjuichi-nichi akiraka-ni shi-mashi-ta. (Government sources showed it on 31st.)
3. IF of a verb + BF of an auxiliary verb “nai” → IF of the verb + “mase-n”	judan-wa mitsukat-te i-nai.	judan-wa mitsukat-te i-mase-n. (The bullet has yet to be found.)
4. IF of a verb + CF of an auxiliary verb “nai” + BF of an auxiliary verb “ta” → IF of the verb + “mase-n-deshi-ta”	kyoi-ni omot-ta kankesha-wa hoto-ndo i-nakat-ta.	kyoi-ni omot-ta kankesha-wa hoto-ndo i-mase-n-deshi-ta. (There were few participators who thought is as a threat.)
5. BF of an auxiliary verb “da” → “desu”	se-to kan-ga ittai-ni nat-te kokumin-ni kotae-te ika-nakere-ba nara-nai kadai- da.	se-to kan-ga ittai-ni nat-te kokumin-ni kotae-te ika-nakere-ba nara-nai kadai- desu. (It is a problem which requires a joint effort by statesman and bureaucrats.)
6. CF of an auxiliary verb “da” + BF of an auxiliary verb “ta” → “deshi-ta”	sore-wa konphuta-hyogen-no mannerizumu-o migoto-ni daha-sita mono- dat-ta.	sore-wa konphuta-hyogen-no mannerizumu-o migoto-ni daha-sita mono- deshi-ta. (It broke down the mannerism of computer art wonderfully.)
7. BF of an auxiliary verb “nai” → “ari-mase-n”	tonai-ni sorehodo-no doyo-wa nai.	tonai-ni sorehodo-no doyo-wa ari-mase-n. (The party members are not so upset.)
8. CF of an auxiliary verb “da” + BF of an auxiliary verb “aru” → “desu”	ta-no tsuiju-o yurusa-nai tokoro- de-aru.	ta-no tsuiju-o yurusa-nai tokoro- desu. (There is nothing comparable to it.)
9. CF of an auxiliary verb “da” + CF of an auxiliary “aru” + BF of an auxiliary “ta” → “deshi-ta”	dokuji-no sakufu-o uchitate-ta-no- de-at-ta.	dokuji-no sakufu-o uchitate-ta-no- deshi-ta. (He established a original style.)
10. CF of an auxiliary verb “da” + IF of an auxiliary “aru” + BF of an auxiliary verb “u” → “desho-u”	jiyu-na sekai-ga kuru-to shinji-ta-kara- de-aro-u.	jiyu-na sekai-ga kuru-to shinji-ta-kara- desho-u. (This was likely because she believed that the free world would come.)
11. IF of an auxiliary verb “da” + BF of an auxiliary “u” → “desho-u”	kesshite raku-na tatakai-de-wa-nai- darou-u.	kesshite raku-na tatakai-de-wa-nai- decho-u. (It will never be an easy fight.)
12. a conjunction “daga” or “ga” → “desuga”	daga jishitsu-shotoku-wa gekigen-shi-te iru.	desu-ga jishitsu-shotoku-wa gekigen-shi-te i-masu. (But actual real income has fallen dramatically.)
13. BF of an adjective “nai” → “ari-mase-n”	kori-to wakat-te i-te-mo hoka-ni sentakushi-ga nai.	kori-to wakat-te i-te-mo hoka-ni sentakushi-ga ari-mase-n. (They have no choice but to borrow the money from loan sharks.)
14. CF of an adjective “nai” + BF of an auxiliary verb “ta” → “ari-mase-n-deshi-ta”	kozoku-no unko-ni sisho-wa nakat-ta.	kozoku-no unko-ni sisho-wa ari-mase-n-deshi-ta. (It did not interrupt the following service.)

BF: basic form, CF: conjunctive form, IF: imperfective form

Table 4: Results of sentence-style conversion.

Precision	Recall	F measure
525/526 (99.8%)	525/559 (93.9%)	96.8

was as follows. ((a) is a original sentence, (b) is the converted sentence.)

1. (a) kantoku-no hyojo-wa ima-hitotsu **sae-nakat-ta.**

(b) kantoku-no hyojo-wa ima-hitotsu **sae-ari-mase-n-deshi-ta.**

(The coach’s expression was not quite good.)

In this sentence, the predict “sae-nakat-ta” should be converted into “sae-mase-n-deshi-ta” based on the rule 4 in Table 2. In the experiment, however, “sae” was regarded as an

Table 5: Causes of the failed conversion.

Cause	Number
sentences ending with an adjective	22
others	12

intensifying particle (which means “even”) by the morphological analyzer, despite “sae” is actually a stem of a verb “sae-ru” (which means “to feel good”). Therefore, this sentence was converted according to the rule 14 instead of the rule 4 shown in Table 2.

The recall was slightly lower than the precision. Table 5 shows the causes of failure to convert 34 places. The details are below.

Sentences ending with an adjective

If a sentence ends with an adjective, it is difficult to create rules for converting the sentence

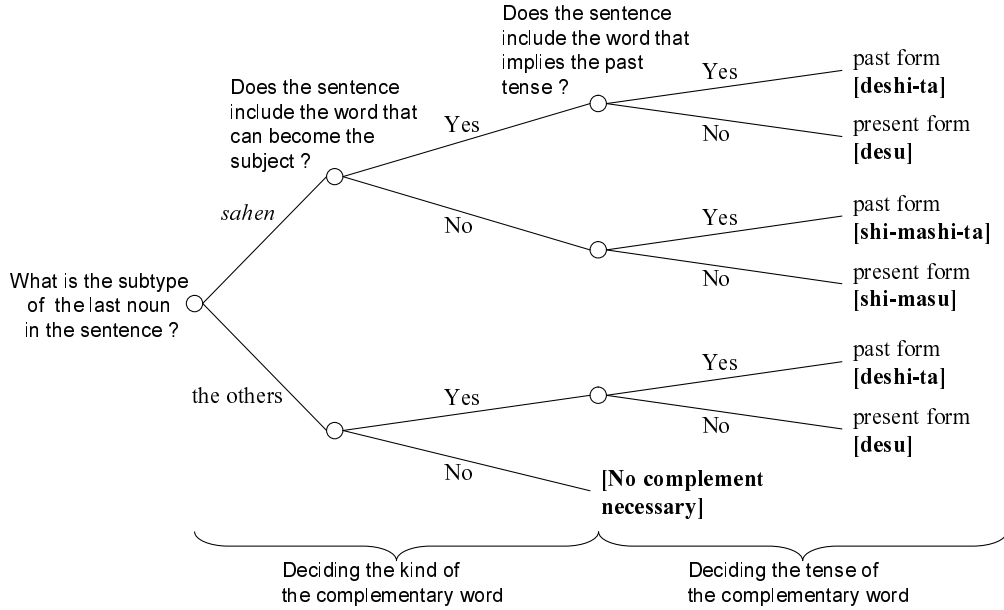


Figure 2: Decision tree for determining the complementary word of *taigen-dome* sentence.

because it requires contextual information to convert this kind of sentences properly. The following are sentences which were unable to be converted into spoken style and their proper conversion. ((a) is a original sentence, (b) is the properly converted sentence.)

2. (a) honkon-ni yat-te-kuru chugokujin-ya gaikokujin-mo **oi**.
 (b) honkon-ni yat-te-kuru chugokujin-ya gaikokujin-mo **oku-i-masu**.
 (There are a great deal of Chinese and people from other countries who come to Hong Kong.)
3. (a) mushiro kange-sa-re-te iru-to it-te-mo **yoi**.
 (b) mushiro kange-sa-re-te iru-to it-te-mo **yoi-desho**.
 (It's not an exaggeration that they are more welcomed.)

The last adjectives “oi” and “yoi” can be converted into “oi-desu” and “yoi-desu” respectively. The converted adjectives are grammatically correct, but unnatural in speech.

The others

The sentences ending with a word “shita” (would like to) or “rashi” (to seem), as with the sentences ending with an adjective, can be converted into “shita-desu,” “rashi-desu” respectively, however, they are unnatural in

Table 6: Results of *taigen-dome*-sentence complement.

Precision	Recall	F-measure
94/164 (57.3%)	94/155 (60.6%)	58.9

Table 7: Causes of the faulty complement.

Cause	Number
tense	47
type of the complementary word	19
necessity of the complement	4

speech. It is difficult to create rules for converting these sentences and it is necessary to consider a method of context-sensitive conversion.

4.2 Results of *Taigen-dome*-Sentence Complement

In the articles on January 3rd, 1995 in the Mainichi newspaper text corpus, 155 of 687 sentences need to be complemented. Table 6 shows the experimental results of the *taigen-dome*-sentence complement. Table 7 shows causes of the faulty complements. The details of the causes are below. In the following examples, (a) is the sentence complemented incorrectly by our method, (b) is the correctly complemented sentence.

Tense

4. (a) sakunen sangatsu-no sutato-irai kanyusha-wa nematsu-made-ni yaku-niman-nin-**deshi-ta**.

- (b) sakunen sangatsu-no sutato-irai kanyu-sha-wa nematsu-made-ni yaku-niman-nin-**desu**.

(Since the start in last March, the number of subscribers has increased to about 20 thousand by the end of this year.)

In this example, the sentence includes the word “sakunen (last year)” which implies the past tense. Actually, however, the present tense is suitable for this sentence considering the context information.

The presence of the auxiliary verb “*ta*” or the word which implies the past tense is not sufficient to determine the tense of the sentence.

Type of the complementary word

5. (a) kaigai-no gakkō-ya byōin-ni taisuru shakai-koken-katsudo-ni-mo sefu-no kise-no ami-**desu**.
(b) kaigai-no gakkō-ya byōin-ni taisuru shakai-koken-katsudo-ni-mo sefu-no kise-no ami-**ga kakari-masu**.
(The governmental regulations are imposed on social action programs for foreign schools or hospitals.)

This indicates that the more kinds of the complementary words are required in addition to “*desu*,” “*deshi-ta*,” “*shi-masu*,” “*shi-mashi-ta*.” Another examples are as follows:

- nihon-sanke-no miyajima-wa futsuka-kan-de kyūman-nin. (ga otozure-mashi-ta)
(Miyajima, one of the three most scenic spots in Japan, was visited by 90 thousand people in two days.)
- kokoku-wa kaku-peji-no gedan-ni-kesai. (sare-masu)
(The advertisement is put on the bottom of each page.)

Necessity of the complement

6. (a) bekoku-no gasu-gaisha-de kono-ni-sannen-no aida-ni shiken-teki-ni jisshi-sarete iru shikumi (no word was complemented).
(b) bekoku-no gasu-gaisha-de kono-ni-sannen-no aida-ni shiken-teki-ni jisshi-sarete iru shikumi-**desu**.
(It is a scheme which American gas companies have introduced on a trial basis in these years.)

In this example, the sentence does not include the word which can become subject.

Therefore, according to the decision tree in Figure 2, the sentence was not complemented even though it needs to be complemented by “*desu*.” We need to review the complement rules which determine the sentence needs to be complemented.

5 Concluding Remarks

In this paper, we have proposed a method of conversion from written language into spoken language for generating spontaneous Japanese speech in a text-to-speech synthesis system. We focused on sentence-style conversion and complement of *taigen-dome* sentences. For sentence-style conversion, we took a fully rule-based approach, and confirmed that only the lexical analysis could achieve the high-precision conversion by the experiment. In the experiment of the complement of *taigen-dome* sentences, the precision and the recall were not high.

In future work, the complement rule for *taigen-dome* sentences will be reconsidered. We have a plan to introduce a machine learning approach to generate the complement rule automatically using a large-scale spoken language corpus.

References

- Masako Yamamoto, Goroh Ohnishi. 2003. Hanashi-kotoba to Kaki-kotoba no Sogo-kanke — Nihongo Kyoiku no Tameni —. Language and Culture (Bulletin of Institute for Language Education, Aichi University), No. 8, pp. 73-90. (in Japanese)
- Toshiki Ohizumi, Nobuhiro Kaji, Daisuke Kawahara, Masashi Okamoto, Sadao Kurohashi, and Toyooki Nishida. 2003. Kaki-kotoba kara Hanashi-kotoba eno Henkan. In *Proceedings of The 9th Annual Meeting of The Association for Natural Language Processing (NLP-2003)*, pp. 93-96. (in Japanese)
- Nobuhiro Kaji, Masashi Okamoto, and Sadao Kurohashi. 2004. Paraphrasing Predicates from Written Language to Spoken Language Using the Web. In *Proceedings of the Human Language Technology Conference (HLT/NAACL-2004)*, pp. 241-248.
- Taku Kudo. MeCab: Yet Another Part-of-Speech and Morphological Analyzer.
<http://mecab.sourceforge.jp/>.