

English Adverb Processing in Japanese-to-English Machine Translation

Kentaro OGURA, Satoshi SHIRAI, and Francis BOND

NTT Communication Science Laboratories
1-1 Hikari-no-oka, Yokosuka-shi, Kanagawa, 239 JAPAN
{ogura, shirai, bond}@cslab.kecl.ntt.co.jp

Abstract. When translating between different language families, the correspondence between language elements are often not straightforward. This paper analyzes Japanese expressions which correspond to English adverbs. We classify the expressions into three types which are translated differently. An adverb processing method in J-to-E machine translation using this classification is proposed.

1 Introduction

In natural language processing, the study of adverbs¹ has not developed very far to date, compared with nouns and verbs, because it was thought that adverbs do not construct the main parts of sentence meaning and have various complex grammatical functions in sentences. However, adverbs occur frequently and make important contributions to sentence meaning. Thus, the accurate processing of English adverbs is required for high quality machine translation.

From a linguistic point of view, linguists have examined adverb grammatical functions and meanings in detail [Quirk et al., 1985]. Conjuncts and disjuncts, usually called sentence adverbs, are described in detail in [Greenbaum, 1983]. Other studies by linguists include those which handle the meanings of specific adverbs, such as “*even*” [Berckmans, 1993], “*still*” and “*already*” [van der Auwera, 1993], and temporal adverb studies which handle temporal semantics in sentences [Vlach, 1993]. There are also studies of adverb position in English in general [Ernst, 1984] and positions of specific adverbs, such as “*however*” [Sugiura, 1991]. But it is difficult to apply the results of these studies to natural language processing directly because they are the knowledge for human so computers cannot understand them easily.

From the natural language processing viewpoint, few studies [Glasbey, 1993] have considered adverbs in natural language processing. [Conlon and Evans, 1992] aimed to decrease ambiguity in adverb meanings and to select words during generation by applying information about adverb semantics and syntax from linguistic studies to an adverb lexicon. These are studies of the multiplicity of adverb meanings. A method for determining where adverbs should be placed in English sentences in Japanese-to-English machine translation has been proposed [Ogura et al., 1994]. The method is based on adverb grammatical functions (subjuncts, adjuncts, disjuncts and conjuncts) and meanings (process, space, time etc.), preferred positions in sentences (initial, medial, end, pre, post), and priorities between adverbs with the same preferred position.

There are few studies of differences in expression between Japanese and English for adverbial meaning. [Ogura et al., 1995] showed that only about 55% of Japanese adverbs were translated into English adverbs in translation from Japanese to English by human translators. On the other hand, only about 17% of English adverbs that appeared in the human translation

¹ In this paper, idiomatic adverbial phrases, such as “on purpose” are considered to be adverbs.

were translated from Japanese adverbs in the original. This shows clearly the difference between Japanese and English representations for adverbial meaning and the difficulty of adverb processing in machine translation.

Adverb processing in Japanese-to-English machine translation is very complicated. Thus, in this paper, we focus our attention on the problem of differences in expressions between Japanese and English adverbial meaning from the viewpoint of English adverbs. When translating between different language families, correspondence between language elements is not straightforward. The tendency is especially prominent in translation of adverbial expressions. So first we examine in detail examples in which Japanese expressions correspond to English adverbs (Section 2). We classify the examples into three types, which are translated differently.

In Section 3, an adverb processing method in Japanese-to-English machine translation using the correspondence types and functional differences of adverb expressions is proposed. The content of Japanese dictionaries which are used to determine Japanese composition of words, Japanese-to-English word transfer dictionaries which are used to transfer by word units, and Japanese-to-English structure transfer dictionaries which are used to transfer by predicate unit, are also presented. Section 4 concludes the paper.

2 Correspondence between Japanese expressions and English adverbs

We examined Japanese expressions which correspond to English adverbs in Japanese-to-English translations made by professional human translators. 1,000 sentences in newspaper articles of industrial and economic domains and their translations were investigated from the viewpoint of correspondence between Japanese expressions and English adverbs. Examining adverb frequency in English showed that adverbs appeared 585 times in 1,000 sentences, that is, one adverb appeared in roughly every two sentences on average. Table 1 shows the result.

We classified the examples into three types, depending on how they can be translated. In Type 1, Japanese expressions can be directly translated into English adverbs, as opposed to Type 2 and Type 3. In Type 1, the Japanese expressions can be transferred by word unit. In Type 2, English adverbs do not correspond directly to Japanese adverbial expressions but rather English idiomatic expressions with adverbs. Type 2 can be also transferred by word (predicate) unit. Type 3 is more complicated and very difficult to translate. The sentence structures were changed because of the different ways of expression or thinking in Japanese and English. As a result, some expressions (nouns, verbs, adjectives etc.) were translated into English adverbs indirectly. Type 3 must be transferred by larger units than the word.

Type 1

Type 1 was the most common type. Type 1 consists of Japanese expressions which have adverbial function and correspondence to English adverbs. 55.9% of English adverbs were translation of Type 1 Japanese expressions. The table shows that Japanese expressions which have adverbial functions are expressed mostly by adverbs, but they are also expressed in other ways. English adverbs which were translated from Japanese adverbs were only 17% of all English adverbs. In Japanese, an adverbial typically takes the form of an adverb or a noun followed by some particle or continuative form of a predicate (1-1).

The most typical case of Type 1 is the correspondence between Japanese adverbs and English adverbs. For example, the Japanese adverb “*saranai*” was translated into the English adverb

Type	Japanese expressions	Correspondence between J/E	Freq.	%
1-1	Adverbs	<i>sarani</i> “even”, <i>shōrai</i> “in the future”	99	16.9
	Adjectival nouns + <i>ni</i>	<i>ōhaba-ni</i> “greatly”	38	6.5
	Nouns + <i>ni</i> or <i>de</i>	<i>chūshin-ni</i> “especially”, <i>kinyūmen-de</i> “financially”	8	1.4
	Verbs (continuative form)	<i>hikitsuduki</i> “continuously”	4	0.7
	Adjectives (continuative form)	<i>subayaku</i> “quickly”	3	0.5
1-2	Numerical prefixes	<i>yaku-</i> “about”	30	5.1
	Quantitative noun	<i>teido</i> “about”	7	1.2
1-3	Modal particles	<i>-dake</i> “only”, <i>-mo</i> “also”	62	10.6
	Modal expressions	<i>sō-da</i> “probably”	14	2.4
1-4	Conjuncts	<i>mata</i> “also”	62	10.6
	Sub total		55.9%	
2	Predicate corresponding to English verbal idiom	<i>secchi-suru</i> “set <u>up</u> ”, <i>okonau</i> “carry <u>out</u> ”	54	9.2
			9.2%	
3	Nouns	<i>chokuei</i> “directly”	68	11.6
	Verbs (attributive form)	<i>isogu</i> “rapidly”	24	4.1
	Adjectival nouns + <i>na</i>	<i>tokushu-na</i> “specially”	12	2.1
	Affixes	<i>shin-</i> “newly”	8	1.4
	Adjectives (attributive form)	<i>chikai</i> “almost”	6	1.0
	Uninflected adjectives	<i>ōkina</i> “greatly”	1	0.2
	“None” or “Paraphrase”		86	14.7
	Sub total		34.9%	
Total			585	100.0

Table 1. Japanese expressions corresponding to English Adverbs

“even”. Adjectival nouns followed by the particle “-*ni*”, continuative forms² of verbs and adjectives, and some nouns followed by particle “-*ni*” or “-*de*”, which have adverbial functions as a whole, were translated into English adverbs. Such Japanese expressions are translated as case elements in English.

Numerical prefixes, such as “*yaku-*” (about) and some quantitative nouns, such as “*teido*” (about) were translated into English intensifiers (1-2).

The Japanese adverbial particle “-*mo*” was translated as “also” a modal adverb (subjunct) in English. Japanese idiomatic expressions for modals were translated into disjuncts (1-3).

Many Japanese conjunctions³ were also translated into English conjuncts (1-4).

Type 2

9.2% of the correspondences were Type 2, where an adverb appears in a larger English expression that is stored in the lexicon. The typical pattern of Type 2 was the pattern that Japanese verbs were translated into English idiomatic verb-particle constructions.

A example⁴ of Type 2 is as follows:

² Continuative forms modify predicates.

³ There were many other Japanese conjunctive expressions, for example, noun “*ippō*” (on the other hand), “*ue-ni*” (besides).

⁴ Examples are given with the (romanized) Japanese original (Jpn), a gloss (Gloss) and the human translation (Eng). Japanese particles are glossed as follows: TOP for “-*wa*” which marks the topic, OBJ for “-*o*” which marks the object, INS for “-*de*” which marks the instrument, LOC for “-*de*” which marks the location, LOC for “-*ni*” which marks the location, and SUBJ for “-*ga*” which marks the

- (1) Jpn: *sono-kaisha-wa tōkyō-de honkaku-saiyō-shiken-o okunatta*
 Gloss: The company TOP Tokyo LOC practical tests **carried out**
 Eng: The company **carried out** practical tests in Tokyo

The Japanese verb “*okonatta*” was translated into “carried **out**” (verb + adverb particle). In this examination, all of the Type 2 examples in the corpus were this type. Another possible kind of Type 2 is “*shinkon*” (a **newly** married couple); this type is quite rare.

Type 3

Type 3 appeared 34.9% of the time in this examination, over a third of the time. Therefore, it is essential to treat this type correctly in Japanese-to-English machine translation. Most direct transfer machine translation systems cannot deal with this type.

The most typical pattern of Type 3 consisted of nouns, attributive forms⁵ of verbs, adjectival nouns followed by particle “*na*”, attributive form of adjectives, or uninflected adjectives which function as modifiers, which were translated into English adverbs. When the sentence structures were changed and the modificants were changed to predicates (verbs or adjectives), the modifiers were accordingly changed to adverbs. An example is shown below.

- (2) Jpn: *Sanyō-denki-wa seisan-o kogaisha-ni zenmen*
ikan suru
 Gloss: Sanyo Electric TOP production OBJ subsidiary company LOC **whole surface**
 transfer do
 Eng: Sanyo Electric **completely** transfers control of production to its subsidiary company

In this example, the light verb “*suru*”, together with its object the verbal noun “*ikan*” was translated into the English action verb “transfer”. Accompanying this process, the attributive noun “*zenmen*” which was the modifier of the Japanese action noun “*ikan*”, was translated into the English adverb “completely” which modifies the action verb “transfer”.

“None” and “Paraphrase” are also included in Type 3. These cases are the most difficult to translate. These cases could not easily be described by rules, because the conditions under which the paraphrase occurs are not clear. “None” is when the English translation has an adverb but there is no element corresponding to it in Japanese. Time adverbs such as “now”, “just” and “still”, or intensifiers such as “extremely”, “highly” are sometimes added in the English translation. We judged that in some of these cases, translations without the “paraphrase” or “addition of adverbs” were perfectly acceptable, so there is no real need to generate them.

3 Processing of adverbial expressions

Figure 1 shows a process of adverbial expressions in Japanese-to-English machine translation (ALT-J/E) [Ikehara et al., 1991]. It shows the clear separation of the processing which treats expressions in which the correspondence between the language elements are not straightforward from ordinary word-to-word processing. This processing is necessary to deal with differences in the ways that different languages are used to express concepts.

The main flow of the process is morphological analysis, dependency analysis, Japanese syntactic and semantic analysis (which involves adverbial element analysis and modal analysis and conjunct analysis), general transfer and adverb generation. Type 1 and Type 2 adverbial expressions are processed in the main flow. They are comparatively easy to handle.

subject. Many Japanese particles have multiple functions depending on the verb. The same function can be expressed by different particles.

⁵ Attributive forms modify nouns.

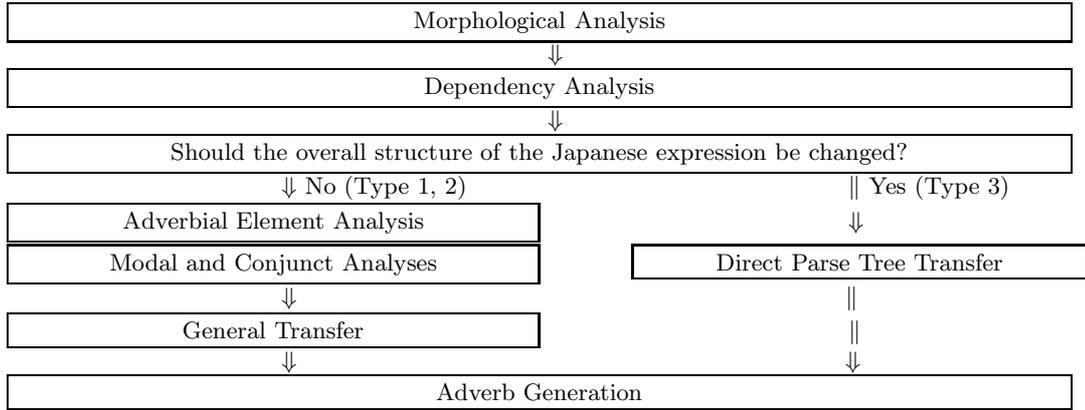


Figure 1. Processing of Adverbial Expressions

In morphological analysis, Japanese sentences are divided into words by using the Japanese word dictionary and the word-conjunction table, and morphological information of the words are examined. Examples from the Japanese word dictionary used to analyze Japanese adverbial expressions are shown below.

Entry (Stem)	Dictionary Information			Adverbial expression
	part of speech	particle	conjugation	
<i>totemo</i> (very)	adverb	—	—	<i>totemo</i>
<i>ōhaba</i> (large, great)	adjectival noun	<i>-na, -ni</i>	—	<i>ōhabani</i>
<i>hikitsuzu</i> (continue)	verb	—	<i>ka-gyō-godan</i>	<i>hikitsuzuki</i>
<i>haya</i> (fast)	adjective	—	adjective	<i>hayaku</i>

Figure 2. Japanese word dictionaries

These entries have a particle slot which shows connectable particles, and a conjugation slot which shows how to be conjugated. Using this information, Japanese expressions which have adverbial functions, such as, adjectival nouns followed by the particle “*-ni*”, and continuative forms of verbs and adjectives, etc., can be analyzed.

In dependency analysis, phrases are constructed from words, and modifier/modificant relationships between phrases are examined.

In the Japanese structure analysis, noun phrases, adverb phrases etc. are determined and parse tree structures are made from phrases. Modal and conjunct information is also analyzed. Then, the Japanese structures are transferred to the English structures by using the Japanese-to-English structure and word transfer dictionaries.

An example rule from the Japanese-to-English structure transfer dictionaries is shown below. It shows correspondence between Japanese and English phrases as a unit sentence.

Figure 4 shows an example from the Japanese-to-English word transfer dictionaries. In this case, *teinei-ni* is translated as **carefully** if it appears with a Japanese verb which means “see”, “hear” or “write”, otherwise as **politely**.

The final process is adverb generation, which generates from English structure (semantic structure) to English sentences.

IF ; “N1-*ga* N2-*o jikkō-suru*”
 J-Verb = “*jikkō-suru*”
 N1 = [Head:Agent or Machine Particle: “*ga*”]
 N2 = [Head:Abstract Particle: “*o*”]
THEN ; “N1 carry out N2”
 E-Verb = “carry”
 Particle = “out”
 Subj = N1
 Obj = N2

Figure 3. Japanese-to-English structure transfer dictionaries

IF “*teinei-ni*” <adjectival noun *-ni*>
THEN “politely” :condition DEFAULT
 “carefully” :condition [Verb heading clause: Seeing, Hearing or Writing]

Figure 4. Japanese-to-English word transfer dictionaries

Direct correspondence between Japanese structures and English structures, such as between Japanese adverbial phrases and English adverbs, and between Japanese predicates and English predicates, can be processed in the main flow.

Type 1 can be classified into four types by Japanese word compositions and functions of adverbial expression. Type 1-1 is Japanese adverbial expressions which work mainly as case elements, such as adverbs and adjectival nouns followed by particle “*-ni*” etc. Type 1-2 is numerical prefixes and some quantitative nouns which work as downtoners. These two types are transferred by word unit using Japanese-to-English word transfer dictionaries. Type 1-3 is modal adverbial expressions. This type is transferred using modal transfer rules. Type 1-4 is conjunctive adverbial expressions. This type is transferred using conjunct transfer rules.

Examples of rules for adverb processing which involve modal transfer rules, and conjunct transfer rules are shown Figure 5 and Figure 6.

IF “*darō*” <auxiliary verb>,
 “*dearō*” <auxiliary verb + auxiliary verb> or
 “*to-omowareru*” <particle + verb + auxiliary verb>
THEN “probably”

Figure 5. Modal transfer rules

Type 2 are mainly transferred by predicate unit using Japanese-to-English structure transfer dictionaries. In Figure 3, the Japanese verb *jikkō-suru* corresponds to the English verbal idiom “carry out”.

But Type 3 correspondences are complicated and the predicate unit and sometimes a Japanese transfer unit is different from an English transfer unit: for example, Japanese nouns to English verbs, Japanese adjectives to English adverbs, Japanese adverbs to English adjectives, Japanese verbs or adjectives to English nouns, Japanese clause to phrase etc.

Type 3 cannot be transferred in conventional direct transfer (word-to-word) method. We propose an adverb transfer method which uses direct parse tree transfer [Matsuo et al., 1995] for Type 3. Direct parse tree transfer provides a flexible framework for translation where source language units are different from target language units.

IF “shikaru-nochi-ni” <conjunctive noun -ni> or “shikaru-nochi” <conjunctive noun> or
“shitagatte” <conjunction> or “tsumari” <conjunction>
THEN “then”

Figure 6. Conjunction transfer rules

In direct parse tree transfer, input is a dependency analysis result and output is an English structure. This process calls analysis modules in Japanese structure analysis. An example⁶ of a Japanese sentence which has a structure that must be changed when it is translated to English is shown below.

(3) Jpn: *watashi-wa shōsai-na kentō-o okonau*
Gloss: I TOP detail examination OBJ do
Eng1: I do a detailed examination
Eng2: I examine in detail

When translating a Japanese light verb such as “*okonau*” (do) or “*suru*” (do) which has an action noun as an object, it is sometimes preferable to use the action noun as a verb. In this example, *kentō-o okonau* (do examination) is translated as “examine”. According to this, the Japanese expression “*shōsai-na*” <adjectival noun -na> (detail) which modifies “*kentō-o*” must be translated to English adverb “in detail”.

The direct parse tree transfer rule used in translating in (3) is shown below.

IF ; B-ga DC-o *okonau*
(A “*okonau*”
(B noun-phrase :particle (“*wa*” or “*ga*”))
(C noun-phrase :particle “*o*” (D adjective-phrase)))
THEN ; B C D
((SUBJ B)
(VP C :VERB)
(ADVP D :ADV))

Figure 7. Outline of direct parse tree transfer rule (adjective to adverb)

4 Conclusion

From the viewpoint analyzing Japanese expressions which correspond to English adverbs, we examined correspondence between Japanese and English. From the result of the examination, we classified the correspondence into three types. The adverb translation method according to the three types was proposed.

Almost all of Type 1 and Type 2 can be translated using the proposed method and half of Type 3 can be translated using it. We estimate that at least 80% of Japanese expressions which are translated into English adverbs can be translated by the proposed method. Currently not all kinds of direct parse tree transfer have been implemented, although the framework for the direct parse tree transfer has been implemented and rules have been provided for cases which were not so complicated. We are currently expanding the number and scope of the rules.

⁶ Eng1 shows the translation without direct parse tree transfer by ALT-J/E. Eng2 shows the translation with it.

In this paper, we proposed a method for adverb processing from the viewpoint of Japanese expressions which correspond to English adverbs. The approach from Japanese adverbs which are translated variously in English must also be studied.

Acknowledgment

The authors wish to thank Satoru Ikehara of Tottori University, Yoshihiro Matsuo and other members of the NTT Machine Translation Research Group for their valuable discussion. They also thank Roland Sussex of Queensland University for his many constructive suggestions, and the members of NTT Advanced Technology for data collection.

References

- [Berckmans, 1993] Paul Berckmans. 1993. The quantifier theory of even. *Linguistics and Philosophy*, 16:589–611.
- [Conlon and Evans, 1992] Sumali Pin-Ngern Conlon and Martha Evans. 1992. Can computers handle adverbs? In *14th International Conference on Computational Linguistics (COLING-92)*, pages 1192–1196, Nantes, France.
- [Ernst, 1984] Thomas Boyden Ernst. 1984. *Towards an integrated theory of adverb position in English*. Ph.D. thesis, Indiana University, Bloomington, Indiana, USA, April.
- [Glasbey, 1993] Sheila Glasbey. 1993. A computational treatment of sentence-final ‘then’. In *Sixth Conference of the European Chapter of the Association for Computational Linguistics (EACL-93)*, pages 158–167, Utrecht, Holland.
- [Greenbaum, 1983] Sidney Greenbaum. 1983. *Studies in English adverbial usage*. Kenkyusha Ltd. (Japanese Translation).
- [Ikehara et al., 1991] Satoru Ikehara, Satoshi Shirai, Akio Yokoo, and Hiromi Nakaiwa. 1991. Toward an MT system without pre-editing – effects of new methods in **ALT-J/E**-. In *Third Machine Translation Summit: MT Summit III*, pages 101–106, Washington DC. (<http://xxx.lanl.gov/abs/cmp-1g/9510008>).
- [Matsuo et al., 1995] Yoshihiro Matsuo, Satoshi Shirai, and Satoru Ikehara. 1995. Changing syntactic classes in transfer-based machine translation. In *Natural Language Processing Pacific Rim Symposium '95: NLPRS-95*, pages 432–437, Seoul.
- [Ogura et al., 1994] Kentaro Ogura, Francis Bond, and Satoru Ikehara. 1994. English adverb generation in Japanese-to-English machine translation. In *Fourth Conference on Applied Natural Language Processing: ANLP-94*, pages 190–191, Stuttgart, October.
- [Ogura et al., 1995] Kentaro Ogura, Satoshi Shirai, and Satoru Ikehara. 1995. Translation of adverbs in Japanese to English machine translation. *IEICE Technical Report, NLC*, 94(575):1–8. NLC-94-44, (in Japanese).
- [Quirk et al., 1985] Randolph Quirk, Sidney Greenbaum, Geoffrey Leech, and Jan Svartvik. 1985. *A Comprehensive Grammar of the English Language*. Longman, London.
- [Sugiura, 1991] Masatoshi Sugiura. 1991. The Distribution Environment of the Connective “however” and the Principle of Its Position — Based on the LOB Corpus —. *Journal of Language and Culture Chubu University Junior College*, 2:47–63.
- [van der Auwera, 1993] Johan van der Auwera. 1993. ‘Already’ and ‘Still’: Beyond Duality. *Linguistics and Philosophy*, 16:613–653.
- [Vlach, 1993] Frank Vlach. 1993. Temporal adverbials, tenses and the perfect. *Linguistics and Philosophy*, 16:231–283.