Gefördert vom
bmb+f
Bundesministerium für
Bildung, Wissenschaft,
Forschung und Technologie

Verbmobil
Verbundvorhaben

# The 101 Translation Problems between Japanese and German/English

Tsutomu Fujinami and
Christine Nanz

Universität Stuttgart

Vm Report 199
July 1997

July 1997

Tsutomu Fujinami and
Christine Nanz

Institut für Maschinelle Sprachverarbeitung
Universität Stuttgart
Azenbergstraße 12
D – 70174 Stuttgart

Tel.: (0711) 121 - 1363
Fax: (0711) 121 - 1366
e-mail: {tsutomu, nanz}@ims.uni-stuttgart.de

**Gehört zum Antragsabschnitt:** 3.1 Contrastive Phenomena and Fragment Encoding

Die vorliegende Arbeit wurde im Rahmen des Verbundvorhabens Verbmobil vom Bundesministerium für Bildung, Wissenschaft, Forschung und Technologie (BMBF) unter den Förderkennzeichen 01 IV 701 N3 gefördert. Die Verantwortung für den Inhalt dieser Arbeit liegt bei dem Autor.

**Abstract**

We investigate differences between Japanese and German/English and explain characteristic phenomena to Japanese. The study helps us to realize what can be problematic when translating Japanese into German/English and vice versa.

# Contents

**1 Introduction** **2**

**2 Japanese as a language** **3**

2.1 Lexical aspects . . . . . 3

2.2 Grammatical aspects . . . . . 4

2.3 Pragmatic aspects . . . . . 7

**3 Lexical phenomena** **9**

3.1 Counters . . . . . 9

3.2 Time expressions . . . . . 16

3.3 Verbal honorifics . . . . . 20

**4 Grammatic phenomena** **22**

4.1 Suffixes to noun phrases . . . . . 22

4.2 Sentence constructions . . . . . 30

**5 Pragmatic phenomena** **42**

5.1 Viewpoint, honorifics and zero-pronouns . . . . . 42

5.2 hai/iie answer . . . . . 44

5.3 Dialogue acts . . . . . 44

**6 Conclusion** **45**

# 1 Introduction

We list up characteristic phenomena in Japanese realised when we compare the language with some other languages such as German or English. Through the study we present problems we encounter when translating Japanese into German/English and vice versa. We are first of all concerned with the translation from spoken Japanese into German or English, but we try to observe the difference in the opposite direction, too, hoping that the contrastive study will clarify characteristic points in each language.

In listing up the phenomena, we categorise them into three classes depending on the sort of information required for translation:

- Lexical phenomena: A translation is *lexical* if the operation only refers to lexical information. That is, no syntactic, semantic, or pragmatic information is necessary.

- Grammatic phenomena: A translation is *grammatic* if the operation only refers to lexical, syntactic, and semantic information. That is, no pragmatic or extra-linguistic information is necessary.

- Pragmatic phenomena: A translation is *pragmatic* if the operation refers to lexical, syntactic, semantic, and pragmatic information. That is, there is no limitation to the resources utilised in translation.

The boundary between lexical and grammatic phenomena is determined based on whether or not the source representation is composed of smaller parts. If the source representation is a lexical item, the translation is lexical. If the source representation is composed of smaller parts related by a syntactic structure, the translation is grammatic or pragmatic.

The boundary between grammatic and pragmatic phenomena is determined based on whether or not all the items of information in the source representation are encoded in the sentence. If the source representation only contains items of information encoded in the sentence, the translation is grammatic. If the source representation contains some items of information which are extracted from other places than the sentence itself, the translation is pragmatic.

The report is accordingly organised into three sections, each describing the phenomena lexical(§3), grammatic(§4), and pragmatic(§5). The definition of these classes needs further refinement, but we believe that the distinction will become clearer when we examine each phenomenon in order.

**Overview** We start with a short introduction to Japanese as a language(§2). We list up features of Japanese grammar and refer the reader to related sections in the report. We then explain characteristics to Japanese, lexically(§3), grammatically(§4), and pragmatically(§5). The last section on pragmatics largely remains to be filled in. We opt for another few reports as follow-up to make this report compact. The number of problems, 101, in the title suggests the degree of complexity we expect of in the project though we can check less than half here. Finally, we conclude the report by giving the reader what to come(§6).

Some words on the origin of this report. We started writing the report to explain issues in the translation between Japanese and English, but the goal of our project has been changed to the translation between Japanese and German while working on the draft. We managed to revise our draft according to the change, but the reader might find some inconsistencies. We however believe that the report is still useful because we primarily looked into characteristics to Japanese, most of which are still characteristic when contrasted against German.

# 2 Japanese as a language

## 2.1 Lexical aspects

### 2.1.1 No gender

There is no gender in Japanese. The feature becomes problematic when we translate Japanese sentences into German because the subject does usually not specify its gender. For example, the following sentence, *"watashi wa gakusei desu"*(I am a student) can be translated into two different sentences depending on the gender of the subject.

(1) watashi wa gakusei desu
I [TOPIC] student be
Ich bin Student/Studentin.

### 2.1.2 No future tense

There is no marking for future tense on verbs. The same sentence (1) can refer to a future event just by adding a time adverbial as below, where the verb form does not change.

(2) rainen watashi wa gakusei desu
next year I [TOPIC] student be
I will be a student next year.

The past tense is however marked explicitly with *'ta'* as shown in (3).

(3) watashi wa gakusei deshi ta
I [TOPIC] student be[PAST]
I was a student.

### 2.1.3 No determiner

There is no determiner in Japanese, therefore, no way to distinguish definite and indefinite descriptions. The distinction between singular and plural is often obscure, too. The plural form has to be either lexically determined or constructed with particular suffixes. The feature leads to considerable degree of ambiguity. A noun phrase, *'hon'*(book), can for instance be translated to either *'a book'*, *'the book'*, *'books'*, or *'the books'*.

The lack of determiners means that the way of quantifying objects or events is quite different from English or German, too. We will discuss the issue in §4.2.1.

### 2.1.4 Verb-object correlations

Different objects can require different verbs even when the action can be seen more or less invariant. For example, the verb, *'kiru'*(to wear, put on), can be specialised depending on the things to be put on, e.g., *'haoru'* for gowns, *'haku'* for pants or shoes, and *'kaburu'* for huts. Different ways of classifying actions across languages can be problematic in translation.

## 2.2 Grammatical aspects

### 2.2.1 Free word order

The word order is quite flexible. The only strict rule is that the predicate of the sentence should come to the final position. The following examples show that the positions of *'Ken ga'* and *'hon o'* are exchangeable, but *'yomimasu'* cannot be moved away from the final position.

(4) a. Ken ga hon o yomimasu
Ken[NOM] book[ACC] read
Ken reads a book/books.

b. hon o Ken ga yomimasu
book[ACC] Ken[NOM] read
Ken reads a book/books.

c. *hon o yomimasu Ken ga
book[ACC] read Ken[NOM]

### 2.2.2 Particles

The grammatical functions are indicated by a number of particles as shown above, e.g., *'ga'* and *'o'*. The particle, *'ga'*, often indicates the subject, but it can indicate the object of the action, too, as shown below. The example shows that the mapping to thematic roles from cases is not straightforward.

(5) watashi wa hon ga yomi tai
I [TOPIC] hon[ACC] read want
I want to read a book/books.

### 2.2.3 No relative pronoun

There is no relative pronoun. Relative clauses therefore have to be identified syntactically. In the following example, the clause, *"Ken ga katta"*(that Ken bought), forms a relative clause modifying on *'koohii'*(coffee). The noun phrase, *"Ken ga katta koohii"*(the coffee that Ken bought), can be thought to be formed by rasing the noun, *'koohii'*, rightwards from the sentence, *"Ken ga koohii o katta"*(Ken bought the coffee).

(6) a. watashi wa [[Ken ga katta $_{RC}$] koohii $_{NP}$] o nomimasu
I Ken [NOM] bought the coffee [ACC] drink
I drink the coffee that Ken bought.

The lack of relative pronoun can lead to ambiguities when combined with zero-pronouns. For example, the syntax of the sentence, *"Ken ga katta koohii o nomimasu"*, can be analysed in two ways as shown below. The first analysis treats *"katta"*(bought) as relative clause, assuming its subject ⟨Ken ga⟩ is suppressed. The second analysis treats *"Ken ga katta"* as relative clause, assuming the subject of *'nomimasu'*(drink), e.g., *"watashi wa"*(I), is suppressed. Each analysis results in different interpretations and translations.

(7) a. Ken ga [[⟨Ken ga⟩ katta $_{RC}$] koohii $_{NP}$] o nomimasu
Ken[NOM] ⟨Ken⟩ bought coffee [ACC] drink
Ken drinks the coffee that he bought.

b. ⟨watashi wa⟩ [[Ken ga katta $_{RC}$] koohii $_{NP}$] o nomimasu
⟨I⟩ Ken [NOM] bought the coffee [ACC] drink
I drink the coffee that Ken bought.

### 2.2.4 Topic marking

There is a particular particle, *'wa'*, to mark topics. The first sentence in the below for instance can be topicalised in terms of *'zoo'*(elephant) by marking the word with *'wa'*.

(8) a. zoo no hana ga nagai
elephant (adnominal) nose [NOM] long
The nose of the elephant is long.

b. zoo wa hana ga nagai
elephant [TOPIC] nose [NOM] long
Speaking of the elephant, its nose is long.

Multiple topics are possible. In the below, the two noun phrases, *'Ken'* and *'hon'*(book) can be topicalised as shown in the second sentence. The topicalisation results in an implication such that he may not read other things than books, i.e., he may not read newspapers.

(9) a. Ken ga hon o yomimasu
Ken [NOM] book [ACC] read
Ken reads books.

b. Ken wa hon wa yomimasu
Ken [TOPIC] book [TOPIC] read
Ken reads books.

### 2.2.5 Passives

There are two kinds of passives, one of which is as known in German or English. The other kind of passives is peculiar to Japanese, called *suffering passive*. In this kind of passive, the subject of the sentence suffers because of the action by the agent in the event. The following sentence means the first person suffered because it rained.

(10) a. watashi wa ame ni furare mashita
I [TOPIC] rain (particle) fall
I got wet by rain.

### 2.2.6 Conditional clauses

There are four particles to construct conditional clauses. They are *'ba'*, *'tara'*, *'nara'*, and *'to'*, each of which requires a particular care when constructing conditionals. The particle, *'ba'*, marks the condition and cannot be replaced by *'to'*. The particle, *'to'*, presupposes that the event described in the conditional clause is followed by the event in the main clause as a consequence. The particle, *'tara'*, requires that the event described in the conditional clause has to be finished till the event in the main clause starts. The particle, *'nara'*, indicates that the condition is a presupposition of the event in the main clause.

(11) a. moshi sore de yokere ba kekkoo desu
if with it happy [COND] fine be
If you are happy with it, I am fine.

b. mokuyoobi desu to gogo dake aite imasu
Thursday be [COND] afternoon only be free
If it is on Thursday, I am only free in the afternoon.

c. kochira ni ki tara yotte kudasai
here [DIR] come [COND] drop in please
If you come here, drop in my office, please.

d. roku ji nara korare masu
six o'clock [COND] (I) can come
If it is at six o'clock, I can come.

## 2.3 Pragmatic aspects

### 2.3.1 Zero pronominals

The nominative or accusative argument is often suppressed when it can be determined in the context. Two examples are presented in the below, one for nominative argument dropping and the other for accusative argument dropping, respectively. In the examples, ⟨She⟩ indicates the pronoun can be suppressed. We will see how zero pronominals can interact with other pragmatic factors in §5.1.

(12) a. (Hanako possessed too many stuff to keep them all in her tiny room. Then,)

b. Taro ni hon wo age-ta
[DAT] book [ACC] gave
⟨She⟩ gave Taro a book (books).

(13) a. (There was a book missing in our library. We learnt later that...)

b. Hanako ga Taro ni age-ta
[NOM] [DAT] gave
Hanako gave ⟨it⟩ to Taro.

### 2.3.2 Honorifics

There are certain ways to express speaker's respect to persons who stand socially in higher position. In the below, the first expression is normal form and does not express any special respect to the professor. The second expression in contrast contains the verb, *'irasshaimasu'*, a honorific form of *'kimasu'*(come). The sort of expression is more appropriate when the speaker stands in lower position than the professor, e.g., a student or secretary. The difference is however lost when translated into German or English. We will see how honorifics can interact with other pragmatic factors in §5.1.

(14) a. sensei wa ashita daigaku ni kimasu
The professor [NOM] tomorrow the university [DIR] come
The professor comes to the university tomorrow.

b. sensei wa ashita daigaku ni irasshaimasu
The professor [NOM] tomorrow the university [DIR] come(honorific)
The professor comes to the university tomorrow.

### 2.3.3 Speaker's sex and vocabulary

Different expressions can be employed depending on the speaker's sex. The phenomena can most distinctively be observed in the choice of first person pronouns. Male speakers can call themselves as *'boku'*. (The use of the pronoun is restricted to informal settings.) Sentence final forms can indicate speaker's sex, too. For instance, female speakers tend to end the sentence with *'wa'*, while male speakers tend to use *'zo'* or *'yo'*. The sort of phenomena have however been gradually less observable these days.

# 3 Lexical phenomena

In this section, we study the lexical difference between Japanese and English. The difference arises where items of information are encoded lexically in Japanese while they are non-existent, grammatically constructed, or pragmatically guessed in English. The most visible contrast in lexical can be seen in the varieties of counters in Japanese. To some counters there is no counterpart in English. We investigate these counters in section 3.1. In the next section 3.2, we consider cases of time expressions where a description of time is lexicalised in Japanese while it is grammatically constructed in English. Finally in section 3.3, we study honorific and formal expressions, where the information is lexicalised in Japanese while it can only pragmatically be extracted in English.

## 3.1 Counters

Things are counted in particular ways in languages, e.g., *"a pint of beer"*. There are varieties of counters in Japanese, too. Most of them are more or less compositionally formed, but some expressions undergo morphological changes and are better treated as a single word. More importantly, they involve in quantifier floating. We thus pay particular attention to counters. The relation to quantifier floating will be discussed elsewhere. The examples are taken from [Maynard 1990] and [Makino and Tsutsui 1995].

### 3.1.1 Object classifiers

Numbers are suffixed with particular classifiers, depending on what is counted. 階 *'kai'* is, for example, used for counting stories of building. There exists however further complication due to different ways of counting stories in Japanese and English; In Japanese, flours are counted from the ground flour, that is, 一階 *'ik-kai'* (the first flour) means the ground flour, not the first flour in English sense. Table 1 shows the relation between Japanese and English words for counting stories.

Apart from the difference in counting stories, some numbers undergo phonological changes. For example, 一階 *'ik-kai'* is thought to be composed of two parts, the first part 一 *'ichi'* counting the story as *'1'* and the second 階 *'kai'* classifying the object as flour. The first part is however turned into *'ik'* as the result of phonological change. Table 2 shows how numbers are counted in Japanese. By referring to the table, one can see that 一階 *'ik-kai'*, 六階 *'rok-kai'*, and 十階 *'juk-kai'* undergo phonological changes. The changes can also be observed in the rest of expressions in the table. The morphological changes also occur at

40, 50, 60, 70, 80, and 90.

| in Kanji | in roman | in English |
|---|---|---|
| 一階 | **ik**-kai | the ground flour |
| 二階 | ni-kai | the first flour |
| 三階 | san-kai | the second flour |
| 四階 | yon-kai | the third flour |
| 五階 | go-kai | the fourth flour |
| 六階 | **rok**-kai | the fifth flour |
| 七階 | nana-kai | the sixth flour |
| 八階 | hachi-kai | the seventh flour |
| 九階 | kyuu-kai | the eight flour |
| 十階 | **juk**-kai | the ninth flour |
| 十一階 | juu-**ik**-kai | the tenth flour |
| 十六階 | juu-**rok**-kai | the fifteenth flour |
| 二十階 | ni-**juk**-kai | the nineteenth flour |
| 三十階 | san-**juk**-kai | the nineteenth flour |
| 百階 | **hyak**-kai | the ninty-ninth flour |

Table 1: Counting stories

Other object classifiers that might be relevant to Verbmobil are as follows:

- 個 *'ko'*: to count a broad category of small and compact objects such as fruit, balls, and boxes.
- 冊 *'satsu'*: to count bound objects such as books.
- 部 *'bu'*: to count bound documents such as leaflets or reports.
- 台 *'dai'*: to count vehicles such as cars. Small machines such as television sets are also counted using *'dai'*.
- 機 *'ki'*: to count bigger vehicles or machines such as airplanes.
- 枚 *'mai'*: to count flat, thin objects such as papers and dishes.

Table 3 summarises numbers undergoing morphological changes to each classifier.

### 3.1.2 Frequency classifiers

Frequency is counted with either 回 *'kai'* or 度 *'do'* (Table 4). The same morphological changes occur to 回 *'kai'* as 階 *'kai'* undergoes, but no morphological changes occur to 度 *'do'*.

| in Kanji | in roman | in arabic |
|---|---|---|
| 一 | ichi | 1 |
| 二 | ni | 2 |
| 三 | san | 3 |
| 四 | yon | 4 |
| 五 | go | 5 |
| 六 | roku | 6 |
| 七 | nana | 7 |
| 八 | hachi | 8 |
| 九 | kyuu | 9 |
| 十 | juu | 10 |
| 十一 | juu-ichi | 11 |
| 十六 | juu-roku | 16 |
| 二十 | ni-juu | 20 |
| 三十 | san-juu | 30 |
| 百 | hyaku | 100 |

Table 2: Counting numbers

| in Kanji | in roman | morphological changes |
|---|---|---|
| 階 | kai | ik-kai(1), rok-kai(6), juk-kai(10), hyak-kai(100) |
| 個 | ko | ik-ko(1), rok-ko(6), juk-ko(10), hyak-ko(100) |
| 冊 | satsu | is-satsu(1), has-satsu(8), jus-satsu(10) |
| 部 | bu | (hereinafter no morphological changes) |
| 台 | dai | |
| 枚 | mai | |

Table 3: Morphological changes to numbers

The difference between these two classifiers are rather subtle. The attention is paid to each occurence when 回 *'kai'* is applied while the all occurrences are altogether classified by 度 *'do'*. The distinction is similar to that between *'each'* and *'every'* in a sense. The distinction however seems to be lost when translated into English.

| in Kanji | in roman | in Kanji | in roman | in English |
|---|---|---|---|---|
| 一回 | **ik**-kai | 一度 | ichi-do | once |
| 二回 | ni-kai | 二度 | ni-do | twice |
| 三回 | san-kai | 三度 | san-do | three times |
| 四回 | yon-kai | 四度 | yon-do, yo-do | four times |
| 五回 | go-kai | 五度 | go-do | five times |
| 六回 | **rok**-kai | 六度 | roku-do | six times |
| 七回 | nana-kai | 七度 | nana-do | seven times |
| 八回 | hachi-kai | 八度 | hachi-do | eight times |
| 九回 | kyuu-kai | 九度 | kyuu-do | nine times |
| 十回 | **juk**-kai | 十度 | juu-do | ten times |
| 十一回 | juu-**ik**-kai | | | eleven times |
| 十六回 | juu-**rok**-kai | | | sixteen times |
| 二十回 | ni-**juk**-kai | | | twenty times |
| 三十回 | san-**juk**-kai | | | thirty times |
| 百回 | **hyak**-kai | | | hundread times |

Table 4: Counting frequency

### 3.1.3 Order classifiers

Orders are counted with either 等 *'too'* or 番 *'ban'*. The classifier, 等 *'too'*, is employed if the number specifies an order or class. The other classifier, 番 *'ban'*, is adopted if the number specifies an order or number itself.

While no morphological changes occur to 番 *'ban'*, the other classifier, 等 *'too'*, undergoes changes at 1 and 10 (Table 5). The classifier, 等 *'too'*, is also ambiguous in that it can specify either place or class.

Other order classifiers are as follows:

- 日 *'nichi'* is applied to dates, e.g., 11 日 *'11-nichi'* means the 11th date. (c.f. §3.2.1)
- 月 *'gatsu'* is applied to months, e.g., 一月 *'ichi-gatsu'* means *'January'*.

| in Kanji | in roman | in English |
|---|---|---|
| 一等 | **it**-too | the first place or class |
| 二等 | ni-too | the second place or class |
| 三等 | san-too | the third place or class |
| 四等 | yon-too | the fourth place or class |
| 五等 | go-too | the fifth place or class |
| 六等 | roku-too | the sixth place or class |
| 七等 | nana-too | the seventh place or class |
| 八等 | hachi-too | the eigth place or class |
| 九等 | kyuu-too | the ninth place or class |
| 十等 | **jut**-too | the tenth place or class |

Table 5: Counting orders

- 年 *'nen'* is applied to years, e.g., 1997 年 *'1997-nen'* means the year of 1997.
- 年度 *'nendo'* is applied to fiscal years starting from April and ending in next March, e.g., 8 年度 *'hachi-nendo'* denotes the eighth year.
- 期 *'ki'* classifies terms or phase, e.g., 一期 *'ik-ki'* means the first phase. Be aware of the morphological change. Since the classifier can also be used for duration, the prefix, 第 *'dai'*, is most likely added to indicate it as ordinal.
- 学期 *'gakki'* is employed for counting semesters, e.g., 一学期 *'ichi-gakki'* means the first semester. Similar to 期 *'ki'*, the classifier can also be used for duration and the prefix, 第 *'dai'*, is therefore likely added to indicate it as ordinal.
- 号車 *'gousha'* is used to count the N-th wagon of trains or buses.
- 番線 *'bansen'* is used to count the N-th track in stations.

### 3.1.4 Duration classifiers

Duration classifiers specify second, minute, hour, week, month, and year as follows:

- 秒 *'byoo'* classifies seconds, e.g., 一秒 *'ichi-byoo'* means one second.
- 分 *'hun'* classifies minutes, e.g., 二分 *'ni-hun'* means two minutes. Morphological changes occur to the classifier at 一分 *'ip-pun'*(1), 三分 *'san-pun'*(3), 4, 6, and 10 minutes.

- 時間 *'jikan'* classifies hours, e.g., 一時間 *'ichi-jikan'* means one hour.
- 週間 *'shuukan'* classifies weeks, e.g., 二週間 *'ni-shuukan'* means two weeks. Morphological changes occur at 一週間 *'is-shuukan'*(1), 八週間 *'has-shuukan'*(8), and 十週間 *'jus-shuukan'*(10).
- 月 *'tsuki'*, or *'kagetsu'* classifies months. One month is expressed either as 一月 *'hito-tsuki'* or 一ヶ月 *'ik-kagetsu'*. The latter expression is more frequently used these days. Note that 一ヶ月 *'ik-kagetsu'* undergoes a morphological change. Changes also occur at 六ヶ月 *'rok-kagetsu'*(6) and 十ヶ月 *'juk-kagetsu'* (10).
- 年間 *'nenkan'* classifies years, e.g., 1 年間 *'ichi-nenkan'* means one year.
- 期 *'ki'* can also specify the period of terms or phrase, e.g., 一期 *'ik-ki'* means the period during the first phase. Note the morphological changes at 6, 10, and 100.
- 学期 *'gakki'* can also be used for the period during semesters, e.g., 一学期 *'ichi-gakki'* means the period during the first semester.

### 3.1.5 Counting people

People are counted using 人 *'nin'* (person). Table 6 shows the way counting people from one to ten. Special attention should be paid to the first two entries. 一人 *'hitori'* should not read as *'ichi-nin'* and 二人 *'futari'* should not read as *'ni-nin'*. Note also that 四人 *'yo-nin'* should not read as *'yon-nin'*.

Another classifier, 名 *'mei'*, is used as well as 人 *'nin'*. No morphological change occurs to 名 *'mei'*.

### 3.1.6 Counting staying nights

The number of nights to stay at some place other than home is counted using 泊 *'haku'*. Morphological changes occur at 1, 3, 4, 6, 8, and 10 nights, where *'haku'* is turned into *'paku'*. Numbers are also changed at 1, 6, 8, and 10 nights.

### 3.1.7 Counting station distance

To describe how many stations away a particular station is, 駅 *'eki'* (station) is used. One and two station distances are irregular (Table 8). Most frequently used expressions are 一駅 *'hito-eki'* and 二駅 *'futa-eki'*, and other expressions are rare.

| in Kanji | in roman | in Kanji | in roman | in English |
|---|---|---|---|---|
| 一人 | **hitori** | 一名 | ichi-mei | one person |
| 二人 | **futari** | 二名 | ni-mei | two persons |
| 三人 | san-nin | 三名 | san-mei | three persons |
| 四人 | **yo**-nin | 四名 | yon-mei | four persons |
| 五人 | go-nin | 五名 | go-mei | five persons |
| 六人 | roku-nin | 六名 | roku-mei | six persons |
| 七人 | nana-nin, **shichi**-nin | 七名 | nana-mei | seven persons |
| 八人 | hachi-nin | 八名 | hachi-mei | eight persons |
| 九人 | kyuu-nin | 九名 | kyuu-mei | nine persons |
| 十人 | juu-nin | 十名 | juu-mei | ten persons |

Table 6: Counting people

| in Kanji | in roman | in English |
|---|---|---|
| 一泊 | **ip-paku** | one night stay |
| 二泊 | ni-haku | two night stay |
| 三泊 | **san-paku** | three night stay |
| 四泊 | **yon-paku** | four night stay |
| 五泊 | go-haku | five night stay |
| 六泊 | **rop-paku** | six night stay |
| 七泊 | nana-haku | seven night stay |
| 八泊 | **hap-paku** | eight night stay |
| 九泊 | kyuu-haku | nine night stay |
| 十泊 | **jup-paku** | ten night stay |

Table 7: Counting staying nights

| in Kanji | in roman | in English |
|---|---|---|
| 一駅 | **hito-eki** | one station away |
| 二駅 | **futa-eki** | two stations stay |
| 三駅 | san-eki | three stations away |

Table 8: Counting station distance

### 3.1.8 Classifying subdivisions of a city

The street system is not adopted in Japan, but a city is divided into districts and sub-districts. Although the number is thought to be ordinal, it is dubious if the property applies to actual areas because a city is certainly two dimensional, not linear. (One can however notice that Paris adopts the same system.)

- 丁目 *'choume'*: the N-th district. Morphological changes occur at 一丁目 *'ic-choume'*(1), 八丁目 *'hac-choume'*(8), and 十丁目 *'juc-choume'*(10).
- 番地 *'banchi'*: the N-th sub-district

## 3.2 Time expressions

Time expressions can be categorised into three classes. The first class is proper-name like expressions such as *"Tuesday 10th December, 1996"*. This class of expressions are studied in the first section 3.2.1. The second class is like pronoun in the sense that the date is dependent on some other date introduced previously. These include expressions such as *'the next day'*. We study the class of expressions in section 3.2.2. Finally, the third class is indexical expressions in the sense that the date is determined relative to the situation where the sentence is uttered. We study these expressions in section 3.2.3.

### 3.2.1 Date expressions

While dates are expressed as ordinal number such as *'first'* or *'second'* in English, particular words are employed in Japanese to express a date (Table 9). As observed in the table, it is possible to analyse 日 *'ka'* as corresponding to *'date'* and its preceding part as indicating the ordinal number, except of the case of 一日 *'tsuitachi'* (first). Morphological changes however make it difficult to separate them into two parts and treating them as a single lexical item is easier.

### 3.2.2 Anaphoric time expressions

Some anaphoric expressions concerning time, date, week, month, season, and year, are lexical in Japanese while they are composite expressions in English. Table 10 shows a list of indexical expressions, some of which are extracted from [Kameyama 1996].

| in kanji | in roman | in English |
|---|---|---|
| 一日 | tsuitachi | first |
| 二日 | futsuka | second |
| 三日 | mikka | third |
| 四日 | yokka | fourth |
| 五日 | itsuka | fifth |
| 六日 | muika | sixth |
| 七日 | nanoka | seventh |
| 八日 | youka | eighth |
| 九日 | kokonoka | ninth |
| 十日 | touka | tenth |
| 二十日 | hatsuka | twentyth |

Table 9: Date expressions

| in Kanji | in roman | in English |
|---|---|---|
| 翌朝 | yokuasa | the next morning |
| 翌夕 | yokuseki | the next afternoon |
| 翌晩 | yokuban | the next evening |
| 翌翌晩 | yokuyokuban | the evening two days later |
| 翌夜 | yokuya | the next night |
| 翌日 | yokuzitsu | the next day |
| 翌翌日 | yokuyokuzitsu | two days later |
| 前日 | zenjitsu | the previous day |
| 前前日 | zenzenjitsu | two days before |
| 後日 | gojitsu | later (on) |
| 翌週 | yokushuu | the next week |
| 翌翌週 | yokuyokushuu | two weeks later |
| 前週 | zenshuu | the week before |
| 翌月 | yokugetsu | the next month |
| 翌翌月 | lyokuyokugetsu | two months later |
| 前月 | zengetsu | the month before |
| 翌春 | yokushun | the next spring |
| 翌秋 | yokushuu | the next autumn |
| 翌年 | yokunen, yokutoshi | the next year |
| 前年 | zennen | the previous year |
| 後年 | kounen | later, afterward |

Table 10: Anaphoric time expressions

### 3.2.3 Indexical time expressions

Some indexical expressions concerning time, date, week, month, season, and year, are lexical in Japanese while they are composite expressions in English. Table 11 and Table 12 show a list of indexical expressions, some of which are extracted from [Kameyama 1996]. Except of 今 *'ima'* (now), 今日 *'kyou'* (today), 明日 *'ashita'* (tomorrow), and 昨日 *'kinou'* (yesterday), the expressions are composite in English. As is observed in the table, the expressions concerning week, month, and year are compositional to some extent, e.g., 今週 *'konshuu'* (this week) can be separated into 今 *'kon'* (this) and 週 *'shuu'* (week). But treating them as one word makes the translation easier.

| in Kanji | in roman | in English |
|---|---|---|
| 今 | ima | now |
| さっき | sakki | a little while ago |
| 今暁 | kongyou | early this morning |
| 今朝 | kesa, konchou | this morning |
| 今夕 | konseki, konyuu | this afternoon |
| 今晩 | konban | this evening |
| 今夜 | konya | this night |
| 昨夕 | sakuyuu | last afternoon |
| 昨晩 | sakuban | last evening |
| 明朝 | myouasa, myouchou | tomorrow morning |
| 明晩 | myouban | tomorrow evening |
| 明夜 | myouya | tomorrow night |
| 今日 | kyou | today |
| 明日 | ashita, asu, myounichi | tomorrow |
| 明後日 | asatte, myougonichi | the day after tomorrow |
| 明明後日 | shiasatte | three days later |
| 昨日 | kinou, sakujitsu | yesterday |
| 一昨日 | ototoi, issakujitsu | the day before yesterday |
| 一昨昨日 | sakiototoki | three days ago |
| 近日 | kinjitsu | sometime soon |
| 先日 | senjitsu | the other day |
| 先先日 | sensenjitsu | sometime ago |

Table 11: Indexical time expressions (time and date)

| in Kanji | in roman | in English |
|---|---|---|
| 今週 | konshuu | this week |
| 来週 | raishuu | next week |
| 再来週 | saraishuu | the week after next |
| 先週 | senshuu | last week |
| 先先週 | sensenshuu | two weeks before |
| 今月 | kongetsu | this month |
| 本月 | hongetsu | this month |
| 当月 | tougetsu | this month |
| 来月 | raigetsu | next month |
| 再来月 | saraigetsu | the month after next |
| 先月 | sengetsu | last month |
| 先先月 | sensengetsu | two months before |
| 今春 | konshun | this spring |
| 今夏 | konka | this summer |
| 今秋 | konshuu | this autumn |
| 今冬 | kontou | this winter |
| 明春 | myoushun | next spring |
| 昨春 | sakushun | last spring |
| 昨秋 | sakushuu | last autumn |
| 昨冬 | sakutou | last winter |
| 今年 | kotoshi | this year |
| 来年 | rainen | next year |
| 明年 | myounen | next year |
| 再来年 | sarainen | the year after next |
| 明後年 | myougonen | the year after next |
| 去年 | kyonen | last year |
| 昨年 | sakunen | last year |
| 一昨年 | ototoshi | two years ago |
| 一昨昨年 | sakiototoshi | three years ago |

Table 12: Indexical time expressions (week, month, season, and year)

## 3.3 Verbal honorifics

Honorifics express speaker's respect to the person about whom the sentence describes. The respect can sententially be expressed by using particular verbs or by adding particular prefixes to nouns. In this section, we study particular verbs to express respect. The sources of this section include [Kindaichi 1988: pp. 188–192].

### 3.3.1 Respect to agent

When the agent of a sentence is respected by the speaker, the verb describing the action is turned into honorific form. Table 13 shows some examples. Besides these lexically encoded honorific forms, there exist other grammatically constructed forms such as 動かれる *'ugok-areru'* (the honorific form of *'ugoku'*) in contrast to the normal form, 動く *'ugoku'* (move). This sort of construction will be studied elsewhere because we have to go into details of Japanese grammar.

| in Japanese | in roman | in English | honorific form | in roman |
|---|---|---|---|---|
| する | suru | do | なさる | nasaru |
| | | | あそばす | asobasu |
| 行く | iku | go | いらっしゃる | irassharu |
| 来る | kuru | come | おいでになる | oideninaru |
| 居る | iru | stay | | |
| 言う | iu | say | おっしゃる | ossharu |
| | | | 仰せになる | ooseninaru |
| 見る | miru | look | ご覧になる | goraNninaru |
| 着る | kiru | wear | お召しになる | omeshininaru |
| 食べる | taberu | eat | | |
| 呼ぶ | yobu | call | | |
| 求める | motomeru | ask | | |
| くれる | kureru | give | 下さる | kudasaru |

Table 13: Honorific verbs to agent

### 3.3.2 Respect to recipient

The following expressions (Table 14) are to be employed when the addressee is a recipient of the verb and stands in higher position than the speaker. The agent of sentences is usually the speaker.

| in Japanese | in roman | in English | honorific form | in roman |
|---|---|---|---|---|
| 言う | iu | say | 申し上げる | moushiageru |
| 見る | miru | look | 拝見する | haikensuru |
| 聞く | kiku | hear | 伺う | ukagau |
| やる | yaru | give | 差し - あげる | sashi-ageru |
| もらう | morau | let | いただく | itadaku |
| 見せる | miseru | show | お目にかける | omenikakeru |

Table 14: Honorific verbs to recipient

### 3.3.3 Formal verbs

In formal setting, some particular expressions are preferred to normal forms. We call them formal expressions. Such expressions are, for example, preferred in business meetings. Table 15 shows some examples of formal verbs. These forms used to be seen as a honorific expression, but recent studies separate them from other honorific expressions because the status of agent or recipient is irrelevant in choosing the expression.

| in Japanese | in roman | in English | honorific form | in roman |
|---|---|---|---|---|
| 行く | iku | go | 参る | mairu |
| 来る | kuru | come | | |
| する | suru | do | 致す | itasu |
| わかる | wakaru | understand | 承知する | shouchisuru |

Table 15: Formal verbs

In the table, it looks confusing that 参る *'mairu'* corresponds to both 行く *'iku'* (go) and 来る *'kuru'* (come). This seems to suggest that the verb does not involve the point of view in description unlike its normal forms, 行く *'iku'* (go) and 来る *'kuru'* (come). The verb, 来る *'kuru'*, is used to describe the movement towards the speaker and other movements are described using 行く *'iku'*. Such a distinction appears not to affect the use of 参る *'mairu'* and can be a source of confusion. But problems usually do not arise owing to discourse effects like below.

(15) minasan basu-ga **mair**imasita. soredewa **mair**imashou!
Dear all the bus[NOM] has arrived now let us go
(Dear all, the bus has arrived. Now let us go!
皆さん、バスが参りました。それでは、参りましょう。)

# 4 Grammatic phenomena

In this section, we study the phenomena that are encoded grammatically in Japanese but encoded differently in English. The first class includes cases where the information is encoded lexically in English, e.g., singular/plural forms. We explain the phenomena in section 4.1, whose title reflects the fact that most constructions are related to suffixes to noun phrases in Japanese. The second class concerns different grammatical constructions to encode information (§4.2). These include counting expressions.

## 4.1 Suffixes to noun phrases

The distinction between cardinal and ordinal numbers is expressed in Japanese by adding a particular suffix or prefix to noun phrases (§4.1.1) while the distinction is lexicalised in English, e.g., *'one'* to *'first'*. Plural forms of nouns are also constructed with particular suffixes (§4.1.2). Pseudonouns follow noun phrases, resulting in some pragmatic effects (§4.1.3). It is however hard to translate these pragmatic effects into English and we can only remove them at the moment.

### 4.1.1 Ordinal numbers

The distinction between cardinal and ordinal numbers is not so strict in Japanese as is the case in English. Confusingly, there are two systems to count things in Japanese, one is Chinese origin and the other originates from old Japanese. Table 16 shows these two systems in parallel, where group I shows the counting system of Chinese origin and group II that from old Japanese. In group II only the numbers up to 10 are distinct from those in group I. The numbers above 10 cannot be counted in group II.

The systems above can be thought to be of cardinal numbers because a particular prefix or suffix has to be added to turn these numbers into ordinals. A prefix, 第 *'dai'*, has to be added to group I numbers (Table 17) and particular suffix, 目 *'me'*, has to be added to group II numbers (Table 18).

The suffix, 目 *'me'*, needs to be supplemented with counters to count numbers above ninth. For example, to count cars, one has to insert 台 *'dai'* between the number and 目 *'me'* as shown in Table 19. Accordingly the counting system must be switched to group I because the counter prefers the counting system.

Despite of these affixes, the distinction is often not explicitly made in practice. For days, for example, the distinction is made only to *'one'* day and *'first'* day,

| in Kanji | in roman (I) | in Kanji | in roman (II) | in arabic |
|---|---|---|---|---|
| 一 | ichi | 一つ | hito-tsu | 1 |
| 二 | ni | 二つ | futa-tsu | 2 |
| 三 | san | 三つ | mit-tsu | 3 |
| 四 | yon | 四つ | yot-tsu | 4 |
| 五 | go | 五つ | itsu-tsu | 5 |
| 六 | roku | 六つ | mut-tsu | 6 |
| 七 | nana | 七つ | nana-tsu | 7 |
| 八 | hachi | 八つ | yat-tsu | 8 |
| 九 | kyuu | 九つ | kokono-tsu | 9 |
| 十 | juu | 十 | too | 10 |
| 十一 | juu-ichi | – | – | 11 |

Table 16: Cardinal numbers

| in Kanji | in roman (I) | in English |
|---|---|---|
| 第一 | dai-ichi | first |
| 第二 | dai-ni | second |
| 第三 | dai-san | third |
| 第四 | dai-yon | fourth |
| 第五 | dai-go | fifth |
| 第六 | dai-roku | sixth |
| 第七 | dai-nana | seventh |
| 第八 | dai-hachi | eigth |
| 第九 | dai-kyuu | ninth |
| 第十 | dai-juu | tenth |
| 第十一 | dai-juu-ichi | eleventh |

Table 17: Ordinal numbers for group I

| in Kanji | in roman (II) | in English |
|---|---|---|
| 一つ目 | hito-tsu-me | first |
| 二つ目 | futa-tsu-me | second |
| 三つ目 | mit-tsu-me | third |
| 四つ目 | yot-tsu-me | fourth |
| 五つ目 | itsu-tsu-me | fifth |
| 六つ目 | mut-tsu-me | sixth |
| 七つ目 | nana-tsu-me | seventh |
| 八つ目 | yat-tsu-me | eighth |
| 九つ目 | kokono-tsu-me | ninth |

Table 18: Ordinal numbers for group II

| in Kanji | in roman (I) | in English |
|---|---|---|
| 一台目 | ichi-dai-me | first (car) |
| 二台目 | ni-dai-me | second |
| 三台目 | san-dai-me | third |
| 四台目 | yon-dai-me | fourth |
| 五台目 | go-dai-me | fifth |
| 六台目 | roku-dai-me | sixth |
| 七台目 | nana-dai-me | seventh |
| 八台目 | hachi-dai-me | eighth |
| 九台目 | kyuu-dai-me | ninth |
| 十台目 | juu-dai-me | tenth |
| 十一台目 | juu-ichi-dai-me | eleventh |

Table 19: Ordinal numbers for cars

and *'twenty'* days and *'twentieth'* day, as shown in Table 20 [Kindaichi 1988-II: pp.79–79].[1] Since expressions involving days and dates appear frequently in Verbmobil domain, we need some disambiguation procedure to deal with the phenomena.

| in Kanji | cardinal | in English | ordinal | in English |
|---|---|---|---|---|
| 一日 | ichi-nichi | one day | tsuitachi | first day |
| 二日 | futsuka | two days | futsuka | second day |
| 三日 | mikka | three days | mikka | third day |
| 二十日 | nijuu-nichi | twenty days | hatsuka | twentyth day |

Table 20: The distinction between days and dates

### 4.1.2 Plurals

Plural forms of common nouns are lexicalized in English, e.g., *'feet'* to *'foot'* or *'cars'* to *'car'*. In Japanese, plural forms are, on the other hand, constructed by adding a particular suffix to nouns. There are two suffixes, たち *'tachi'* and ら *'ra'*. While ら *'ra'* applies to most nouns, たち *'tachi'* tends to apply only to animate objects as seen in Table 21. Thus, これたち *'kore-tachi'* or あれたち *'are-tachi'* is ungrammatical.

[1] Kindaichi also points out that the distinction is sometimes sloppy even in English, e.g., *"Chapter 1"* reads usually *"chapter one"*, but should be read *"chapter the first"*. In Japanese first chapter is expressed as 第一章 *'dai-is-shoo'*.

| in Kanji | in roman | in English | in Kanji | in roman | in English |
|---|---|---|---|---|---|
| 子供 | kodomo | child | 子供たち | kodomo-tachi | children |
| 男 | otoko | man | 男たち | otoko-tachi | men |
| 私 | watashi | I | 私たち | watashi-tachi | we |
| 彼 | kare | he | 彼たち | kare-tachi | they |
| 彼女 | kanojo | she | 彼女たち | kanojo-tachi | they |
| 子供 | kodomo | child | 子供ら | kodomo-ra | children |
| 男 | otoko | man | 男ら | otoko-ra | men |
| 私 | watashi | I | 私ら | watashi-ra | we |
| 彼 | kare | he | 彼ら | kare-ra | they |
| 彼女 | kanojo | she | 彼女ら | kanojo-ra | they |
| これ | kore | this | これら | kore-ra | these |
| あれ | are | that | あれら | are-ra | those |

Table 21: Plurals

The suffix ら *'ra'* is not always applicable. Some plural forms are derived by repeating noun stems as below.

- 山山 or 山々 *'yama-yama'* is the plural form of 山 *'yama'* (mountain).
- 何何 or 何々 *'nani-nani'* is the plural form of 何 *'nani'* (what).
- 誰誰 or 誰々 *'dare-dare'* is the plural form of 誰 *'dare'* (who).

This sort of plural form might be better treated as a lexical item because some constructions undergo morphological changes, e.g., a plural form of 人 *'hito'* (man) is either 人人 or 人々 *'hito-bito'* (people), where the first consonant of the repeated part gets voiced. The other reason is that the same construction is also applied to construct adverbials such as 時々 *'toki-doki'* (sometimes). There is the other suffix ども *'domo'* to pluralise a noun, but it will be studied later because of its pragmatic effect (humble).

Serious problems are caused by different uses of plurals in Japanese and English. The distinction between singular and plural forms is quite sloppy in Japanese and the hearer is required to infer if the denoted object is a single or multiple item. In the following sentence (16a), for example, it is unclear if the speaker read only one book or several books. To make it clear that he read several books, the books must be counted using the counter, 冊 *'satsu'*, as shown in (16b). (Alternatively, they can be quantified.)

(16) a. watashi-wa hon-wo yon-da
I book read
(I read a book (or books).
私は本を読んだ。)

b. watashi-wa hon-wo suu-satu yon-da
I book several read
(I read several books.
私は本を数冊読んだ。)

The problem in English is that (bare) plurals can be used either to indicate a kind or to refer to individuals. For example, in the sentence (17a), *'hurricanes'* means hurricanes in general, but not a particular one, while it denotes some particular hurricanes in (17b), but not hurricanes in general.[2] We call the first sort use *kind* reading and the second *instance* reading.

(17) a. Hurricanes are dangerous.

b. Hurricanes arose in the South Pacific last night.

When translating singular form nouns from Japanese to English, thus, it must be first checked out if the noun denotes a particular object or a kind. If the noun is interpreted as indicating a kind in general, the translated noun must be plural. If the noun denotes a particular object, one has to check if the denoted object is single or multiple item. But the problem is that it is often unclear if a singular form noun denotes a single or multiple item. One needs to look into broader context, but it may not give a decisive answer. Translating plural form nouns in Japanese may, on the other hand, not be so involving; they can be translated to plural form nouns because they usually refer to particular individuals in Japanese.

Taking the other way around, when translating plural form nouns from English to Japanese, it must be first determined if the noun is of kind reading or instance reading. The plurals with kind reading are to be translated into those without suffixes in Japanese because the noun does not actually denote multiple entities. If the noun denotes a particular multiple object, they can be translated to plural form nouns. But the problem is that not all nouns can be pluralised in Japanese. If no plural form is available, one is forced to translate it into a singular noun. Alternatively, the exact number of items can be counted. Another problem is to choose the most appropriate suffix to pluralise a noun, ら *'ra'* or たち *'tachi'*.

[2]The examples are taken from [Glasbey 1995].

### 4.1.3 Pseudonouns

Pseudonouns are problematic because even its definition is still in the debate among researchers. In the literature [Nagara et al. 1989], Pseudonouns are defined by syntax and thought to convey some pragmatic information such as speaker's intention, but it is yet to be investigated what sort of pragmatic information can be conveyed by each pseudonoun. It seems that pseudonouns have something to do with focus, that is, a noun phrase is marked as a focus by the construction, NP + pseudonoun. If it is the case, the same information may be encoded prosodically in English, but we will leave the question out because we do not know how to deal with such pragmatic information yet.

The pseudonouns considered here are, あたり *'atari'*, ぐらい *'gurai'*, ばかり *'bakari'*, など *'nado'*, and とか *'toka'*. A number of pseudonouns appear in conversation, but we limit our study to most frequent ones in Verbmobil domain. The examples given below are extracted from Verbmobil corpus unless stated otherwise. Sources in this section include [Nagara et al. 1989].

**あたり atari** The pseudonoun, あたり *'atari'* denotes vaguely a time, location, person, or event. In example (18a), the pseudonoun follows the noun phrase, 金曜日午前中 *'kiNyoubi gozeNchuu'* (Friday morning), and is thought to denote a time span. The part involving the example is given below. The pseudonoun can be removed without changing the meaning, that is, the propositional content of the sentence (18b) is exactly the same as (18a) because the noun phrase itself introduces a time reference, thus the pseudonoun is redundunt.

The subtle difference between (18a) and (18b) is that (18b) sounds more straightforward in proposing a time abruptly. The sentence (18a) may, on the other hand, sound more polite because the speaker leaves some room for considertation or gives a pause to the hearer.

(18) a. kiNyoubi gozeNchuu **atari** ga arigatai N desu ga
Friday morning is appreciated
(Friday morning is good to me.
金曜日午前中あたりがありがたいんですが)

b. kiNyoubi gozeNchuu ga arigatai N desu ga
Friday morning is appreciated
(Friday morning is good to me.
金曜日午前中がありがたいんですが)

[034] ACC08: mooshiwake nai N desu ⟨ga⟩ watakushi • ni juu roku nichi wa ichi nichi juu kaigi de tsumat te ori mashi te • ○ dekire ba

• ni juu nana nichi • mokuyoobi • no gogo ka • ni juu hachi nichi no • kiNyoubi gozeNchuu atari ga arigatai N desu ga • seNsee go tsugoo ikaga desu ka . ○

ACC08: 申し訳ないんです〈が〉わたくし • 二十六日は一日中会議で詰まっておりまして • できれば • 二十七日 • 木曜日 • の午後か • 二十八日の • 金曜日午前中あたりがありがたいんですが • 先生ご都合いかがですか。

I am sorry. For the 26th my schedule has already been filled up all the day with meetings. It would be most convenient to me if it will be scheduled in the afternoon on Thursday the 27th or in the morning on Friday the 28th. How convenient is it to you?

The pseudonoun can also follow locations. In example (19),[3] the pseudonoun follows 神田 *'kanda'*, a place in Tokyo well-known for accommodating many secondhand bookshops who are specialised in selling and buying rare books, and denote vaguely the area. The meaning of the sentence is however not changed even if the pseudonoun is removed because the propername, 神田 *'kanda'*, can introduce a location reference by itself.

(19) zeppan ni natta hon-wo sagasu N dattara
to look for books out of print

| | | |
|---|---|---|
| Kanda-**atari** no | furuhoNya-e | ittemitara doudesu-ka |
| in Kanda | secondhand bookshops | you may want to visit |

(To look for books out of print, you may want to visit secondhand bookshops in Kanda. 絶版になった本を捜すんだったら、神田あたりの古本屋へ行ってみたらどうですか。)

**ぐらい gurai** The pseudonoun, ぐらい *'gurai'*, denotes vaguely a time point, following a time expression, while あたり *'atari'* denotes a time span. In example (20), the speaker proposes to see the hearer after three o'clock. The time expression, 午後三時 *'gogo saN ji'* (three o'clock in the afternoon), is followed by the pseudonoun, ぐらい *'gurai'*. Althouth adding the pseudonoun to the noun phrase has an effect similar to *'about'* in English, i.e., *"about three o'clock"*, or *'at latest'*, i.e., *"three o'clock at latest"*, removing the pusedonoun does not change the meaning.

[3] The example is taken from [Nagara et al. 1989].

(20) muika deshi tara gogo saN ji **gurai**
the 6th as for afternoon three o'clock
kara nara ai te ori masu keredomo
after provided I am free
(As for the 6th, I am free after three o'clock in the afternoon.
六日でしたら、午後三時ぐらいからなら空いておりますけれども。)

[040] NAW10: hai . ○ (ee) nanoka wa ● chotto ichi nichi juu muri na N desu keredomo ○ ● muika deshi tara ● gogo saN ji gurai kara nara ● ai te ori masu keredomo . ○ go tsugoo yoroshii deshoo ka . ○

NAW10: はい。(えー) 七日は ● ちょっと一日中無理なんですけれども ● 六日でしたら ● 午後三時ぐらいからなら ● 空いておりますけれども。 ご都合よろしいでしょうか。

O.K. I cannot find a time for the seventh, but can see you after three o'clock on the sixth. Is it convenient to you?

**ばかり bakari** The pseudonoun, ばかり *'bakari'*, follows a time period expression among many usages. The meaning of the pseudonoun looks similar to *'about'* and the use of the pseudonoun indicates that the period is not so long. But the use of ばかり *'bakari'* does not affect the propositional content. In example (21) the speaker indicates that the time period for the meeting is not longer than two hours by saying 二時間ばかり *'nijikan bakari'* (for two hours).

(21) nijikaN **bakari** oai shitai
two hours see want to
(I want to see you for two hours.
二時間ばかりお会いしたい。)

[087] AAW04: hai ● (e) naiyoowa tokuni kuwashiku kiiteinaiN-desu-keredomo ● oota seNsee-to ● ○ (e) sukoshi nijikaN-bakari oai shitai-to ● (ee) itteru-N-desuga . ○

AAW04: はい ● (えー) 内容は特に詳しく聞いていないんですけれども ● 太田先生と ● ○ (えー) 少し二時間ばかりお会いしたいと ● (ええ) 言ってるんですが。 ○

Yes, I haven't heard the detail particularly, but he tells me that he would like to see you, Oota-sensei, for two hours.

**など nado** The pseudonoun, など *'nado'*, emphasises the noun phrase. In example (22) the date is emphasised by being added the pseudonoun. The

pseudonoun does however not contribute to its propositional content and can be eliminated without changing the meaning.

(22) raishuu no ni juu roku nichi kiNyoobi no
next week the 26th Friday
gozeNchuu **nado** ikaga deshoo ka
in the morning how about
(How about scheduling it in the morning on Friday the 26th next week?
来週の二十六日金曜日の午前中など如何でしょうか。)

[159] ABY05: soo desu ne • (ee) deki mashi tara • (ee) raishuu no • (ee) ni juu roku nichi kiNyoobi no • gozeNchuu nado ikaga deshoo ka . ○

ABY05: そうですね • (えー) できましたら • (えー) 来週の • (えー) 二十六日金曜日の • 午前中など如何でしょうか。

Allright. I would be quite happy if we can schedule it in the morning on Friday the 26th next week.

**とか toka** The pseudonoun, とか *'toka'*, functions similar to など *'nado'*, and emphasises the noun phrase it follows.

(23) sochira no tsugoo no ii hi **toka** wa nai deshoo ka
to you convenient day TOPIC isn't there any
(Isn't there any day convenient to you?
そちらの都合のいい日とかは、ないでしょうか)

[016] NAC04: (anoo) kyoo deNwa • sashi te itadai ta keN na N desu keredomo • ○ (anoo) • ree no keNkyuukai no • teema no uchiawase o • shi tai N desu keredomo . ○ (ano) • sochira no • tsugoo no ii hi toka • wa • nai deshoo ka . ○

NAC04: (あのー) 今日電話 • さしていただいた件なんですけれども • (あのー) • 例の研究会の • テーマの打合せを • したいんですけれども。 (あの) • そちらの • 都合のいい日とか • は • ないでしょうか。

The reason why I am calling you today is that I would like to discuss the theme for the workshop. Isn't there any day convenient to you?

## 4.2 Sentence constructions

We study in this section how differently the same information is encoded gramatically in Japanese and English. The section is therefore organised reflecting

semantic units. In listing up semantic units, we adopt ideas from generalised quantifier theory [Barwise and Cooper 1981] and regard a basic proposition as consisting of a relation and some objects standing in the relation, where the objects may be quantified over. A sentence, *"every woman works"*, is for example analysed as (every, $x$, woman($x$))[work($x$)], where $x$ is a reference to the person introduced by the utterance, who is woman in its property and is working in its relation.

We start by looking at how the quantifier is expressed sententially, then proceed to the investigation into the relation. That is, we first look at the part of grammar concerning '(every, $x$, woman($x$))' (§4.2.1), then the part concerning 'work($x$)' (§4.2.2, §4.2.3).

### 4.2.1 Quantifiers

A determiner followed by a noun is regarded as a quantifier in generalised quantifier theory. A phrase, *'every woman'*, is for example a quantifier such that *'every'* is determiner and *'woman'* noun. Such construction is not possible in Japanese because there is no determiner. To form a generalised quantifier an adnominal particle, の *'no'*, is employed to glue a measure adverb and a noun together. The quantifier corresponding to *'every woman'* is, for example, すべての女性 *'subete no josei'*, where すべて *'subete'* is a measure adverb corresponding to *'every'* and 女性 *'josei'* a noun corresponding to *'woman'*. Table 22 lists up measure adverbs with corresponding determiners in English. (All the measure adverbs are observed in Verbmobil corpus.)

An important difference with English is that no monotone decreasing quantifiers exist such as *'no woman'* or *'few women'*. To express the sort of negative information, the proposition must be negated at the predicate. To express the same information encoded using *'few'*, for example, the construction, ほとんど ...な い *'hotoNdo'*...(neg), is adopted, which roughly corresponds to *"Most"*...(neg).

A more interesting and important difference is in the frequent appearance of quantifier floating.[4] Of 600 conversations recorded in Verbmobil corpus, only a few constructions with the adnominal particle, の *'no'*, are observed. We present some examples in the following, all of which are taken from Verbmobil corpus.

Given these two distinctive differences between Japanese and English, two sorts of problems arise in translation. The lack of monotone decreasing quantifiers leads to a complex operation to remove negation from the verb phrase and select

[4]We do not explain about quantifier floating in Japanese. The reader is referred to [Fukushima 1991] for linguistic data and observation.

| in Kanji | in roman | determiner |
|---|---|---|
| すべて | subete | every |
| それぞれ | sorezore | each |
| いずれ ... も | izure...mo | all |
| いずれ ... か | izure...ka | either |
| ずい分 | zuibuN | many |
| かなり | kanari | many |
| 沢山 (たくさん) | takusaN | many |
| 多く | ooku | many |
| ほとんど | hotoNdo | most |
| いくつか | ikutsuka | some |
| 少し | sukoshi | a few |
| 両方 | ryoohoo | both |
| どちら ... も | dochira...mo | both |
| あまり ... ない | amari...(neg) | not so many |
| ほとんど ... ない | hotoNdo...(neg) | few |
| 全然 ... ない | zeNzeN...(neg) | no |
| 全く (まったく)... ない | mattaku...(neg) | no |

Table 22: Measure adverbs

a proper determiner, e.g., *'no'* or *'few'*, to nouns when translating sentences from Japanese to English. In opposite direction, the verb phrase has to be negated and a proper measure adverb has to be chosen. The frequent use of floating quantifiers causes ambiguities such that it is uncertain if an adverb operates on a verb phrase or an adjective.

In the following, we first look at examples involving measure adverbs, すべて *'subete'*, いくつか *'ikutsuka'*, それぞれ *'zorezore'*, and いずれ ... も *'izure...mo'* to see how generalised quantifiers are constructed with measure adverbs. We then look at a sentence with 全然 ... ない *'zeNzeN'*...(neg), the sentence that should be expressed with monotone decreasing quantifiers in English. Finally, we examine a sentence with ずい分 *'zuibuN'*, whose meaning is ambiguous between verb phrase and adjective operators.

**すべて subete** The measure adverb, すべて *'subete'*, corresponds to *'every'*. In example (24a) the adverb adds the measure information to the verb あい *'ai'* (free). The sentence is thought to be a short form of (24b) or (24c). The former is more likely to be the original sentence because the object quantified over is the days from Monday to Friday introduced by the word 今週 *'konshuu'* (this week). We think therefore the quantifier is floating.

(24) a. koNshuu wa subete ai te ori masu
this week TOPIC every is free
(Every day is free this week.
今週は、すべて空いております。)

b. koNshuu wa getsuyoo kara kiNyoo made subete ai te ori masu
this week TOPIC from Monday to Friday every is free
(Every day is free from Monday to Friday this week.
今週は、月曜から金曜まで、すべて空いております。)

c. koNshuu wa subete no hi ga ai te ori masu
this week TOPIC every day[NOM] is free
(Every day is free this week.
今週は、すべての日が空いております。)

[027] AAA07: (e) koNshuu wa • (ee) subete ai te ori masu ga . ○
AAA07: (え) 今週は • (えー) すべて空いておりますが。

**いくつか ikutsuka** The measure adverb, いくつか *'ikutsuka'*, corresponds to *'some'*. Example (25a) is completed with 会議 *'kaigi'* (meeting), which is missing in the original sentence. In the sentence, the measure adverb occupies the first position. The adverb can be also placed just after the subject, 会議 *'kaigi'* (meeting) as shown in example (25b). The quantifier can be expressed with the adnominal particle, の *'no'*, as shown in example (25c), but such a sentence does not appear in the corpus.

(25) a. ikutsuka kaigi ga haittemasu
some meeting[NOM] are scheduled
(some meetings are scheduled.
いくつか会議が入ってます。

b. kaigi ga ikutsuka haittemasu
meeting[NOM] some are scheduled
(some meetings are scheduled.
会議がいくつか入ってます。

c. ikutsuka no kaigi ga haittemasu
some meeting[NOM] are scheduled
(some meetings are scheduled.
いくつかの会議が入ってます。

[B-312] ACU07: hai ○ • soodesunee ○ • raishuuno getsuyoobi juuhachinichidattara • watashi ichinichi aiteruNdesuga ○ • sore igaiwa • ikutsuka • sukoshizutsu haittemasu . ○ seNsee • gotsugoowa • ikagadesuka . ○

ACU07: はい ○ ● そうですねえ ○ ● 来週の月曜日 17 日だったら ● 私、一日あいてるんですが ○ ● それ以外は ● いくつか ● 少しずつ入ってます。 ○ 先生 ● ご都合は ● いかがでしょうか。 ○

Yes, well, on Monday the 17th next week I am free all the day, but in the rest of the week some (meetings) are scheduled from time to time. How is your schedule?

**それぞれ sorezore** The measure adverb, それぞれ *'sorezore'*, corresponds to *'each'*. In example (26a) the adverb operates on 分野 *'buNya'* (subject). The noun is at the same time added a prefix, 各 *'kaku'*, which also means *'each'*. The prefix behaves more like determiner in English and can form a quantifier by itself as seen in (26b) when the noun is followed by a postpositional, について *'nitsuite'* (for).

The phrase, 各分野それぞれ *'kaku-buNya sorezore'*, is a floating quantifier because it can be placed at the first position as seen in example (26c). The construction with の *'no'* is possible, too, as seen in (26d) provided that the noun, 分野 *'buNya'* (subject), is followed by a suffix, ごと *'goto'*, which also means *'each'*. In example (26d) それぞれの *'sorezore no'* can be removed, but sounds slightly odd.

(26) a. keNkyuu naiyoo wa kaku-buNya sorezore yooi sareru
research topic TOPIC each-subject each be prepared
(A research topic will be presented for each subject.
研究内容は各分野それぞれ用意される。)

b. keNkyuu naiyoo wa kaku-buNya nitsuite yooi sareru
research topic TOPIC each-subject for be prepared
(A research topic will be presented for each subject.
研究内容は各分野について用意される。)

c. kaku-buNya sorezore keNkyuu naiyoo wa yooi sareru
each-subject each research topic TOPIC be prepared
(各分野それぞれ研究内容は用意される)

d. keNkyuu naiyoo wa sorezore no buNya-goto ni yooi sareru
research topic TOPIC each subject-each for be prepared
(研究内容はそれぞれの分野ごとに用意される。)

[104] ABB14: soo desu nee ● (ma) ● sono keNkyuu naiyoo nitsuite wa ● kaku buNya sorezore yooi sareru ka to omoi masu node . ○ shootai suru seNseegata no ● sukejuuru ● moshi ari mashi tara onegai shi tai N desu ga . ○

ABB14: そうですねえ • (ま) • その 研究内容については • 各分野それぞれ用意されるかと思いますので。 招待する先生方の • スケジュール • もしありましたらお願いしたいんですが。

Well, as for research topics, I believe that for each subject a topic will be presented. Could I ask you about the schedules of invited lectures, if any?

**いずれ . . . も izure. . .mo** The adverb corresponds to *'all'*. In example (27a) the adverb operates on the dates. The adverb must be followed by a particle, も *'mo'*.

(27) a. watakushi no hoo wa nanoka juu saN juu yoN izure mo ai te ori masu
For me the 7th, 13th, 14th all are free
(As far as I am concerned, the 7th, 13th, and 14th are all free.
わたくしの方は七日十三十四、いずれもあいております。)

[096]: ABA08: soo desu ka . ○ (e) watakushi no hoo wa nanoka juu saN juu yoN • (ee) izure mo ai te ori masu node ○ maa dekire ba hayai hoo ga yoi to omoi masu node • ○ soredewa nanoka wa ikaga deshoo ka . ○

ABA08: そうですか。 (え) わたくしの方は七日十三十四 • (えー) いずれもあいておりますのでまあできれば早い方がよいと思いますので • それでは七日はいかがでしょうか。

I see. As for the 7th, 13th, and 14th, I am free all the days. I believe, the earlier, the better. So how about the 7th?

The construction with the adnominal particle, の *'no'*, is possible. As is observed in example (28a) いずれ *'izure'* (all) and 日 *'hi'* (day) are combined by the particle. In this case, the particle, も *'mo'*, follows the noun.

(28) a. izure-no hi-mo kochirano hoomo yoteega gozaimasu
all day for us too something is planned.
( all the days we are engaged, too.
いずれの日も、こちらの方も、予定がございます。)

[B-390]: ACN11: mooshiwake arimaseN ○ • (ee) izureno himo • kochirano hoomo • yoteega gozaimasu . ○ koNkaiwa • goeNryo sasete itadakutoyuu kotode • yoroshiideshooka . ○

ACN11: 申し訳ありません ○ • (ええ) いずれの日も • こちらの方も • 予定がございます。○ 今回は • ご遠慮させていただくということで • よろしいでしょうか。 ○

I'm sorry that I have engagements all the days. Could you forgive me to turn down your invitation this time?

**全然...ない zeNzeN ...(neg)** There is no monotone decreasing quantifier in Japanese and quantification expressions such as *'no one'* or *'nothing'* can only be constructed by negating the relation. In example (29) the verb, はいってませ *'haitte mase'* (is scheduled), is negated by ん *'N'*, where 全然 *'zeNzeN'* emphasises the negation, producing the similar effect to *'nothing'*.

(29) juuichinichi wa kochira wa zeNzeN yotee ga haitte mase N
11th[TOPIC] I[TOPIC] entirely appointment[NOM] is scheduled (neg)
(On the 11th I have nothing scheduled.
11 日は、こちらは、全然、予定がはいってません。)

[B-543]: NDU06: soodesu . ○ juuichinichiwa ● (e) kochira wa zeN-zeN ● yoteega haittemaseNnode ○ ● sochirano yoteewa doodeshooka . ○

NDU06: そうです . ○ 11 日は ● (え) こちら は 全然 ● 予定がはいってませんので ○ ● そちらの予定はどうでしょうか。○

Yes. On the 11th I have nothing scheduled. How is your schedule?

**ずい分 zuibuN** The adverb corresponds to *'lot'* or *'many'*. In example (30a) ずい分 *'zuibuN'* operates on 自由な日 *'jiyuuna hi'* (free days) and the sentence means that the speaker has many days free. Although the noun, 日 (day), itself is ambiguous between singular and plural forms, only the reading as plural is possible in the context.

The same meaning can also be expressed with the sentence (30b), where the adverb is placed at the first position. The sentence is however ambiguous in that another reading is possible such that the speaker has a day when he is considerably free. In this reading, the adverb, ずい分 *'zuibun'*, operates on the adjective, 自由な *'jiyuuna'* (free), and expresses the degree of his freeness. Despite of this ambiguity, the sentence (30b) is actually uttered. But the context eliminates the wrong reading, the context where they talk about free days in May.

The sort of ambiguity is due to the fact in Japanese that quantification is achieved with adverbs. In fact, the measure adverbs are more often observed to appear as a normal adverb. It would be therefore more appropriate to regard quantification as a particular use of adverbs. The feature is more important to study because the sort of ambiguity is more frequent than the ambiguities due

to quantifier scoping. In Verbmobil corpus very few sentences are constructed with more than one quantifiers.

(30) a. jiyuuna hi ga **zuibuN** aruNdesu
free day[NOM] lot there is(are)
(I have lots of days free.
自由な日がずい分あるんです。)

b. **zuibuN** jiyuuna hi ga aruNdesu
lot free day[NOM] there is(are)
(I have lots of days free.
I have a day when I am considerably free.
ずい分自由な日があるんです。)

[B-305]: ACT05: soodesunee ○ ● kyoowa juugonichidesukara ○ ● (eeto) gogatsuwa ● koohaN ● zuibuN jiyuuna higa aruNdesuyo . ○ doozo osshattekudasai . ○

ACT05: そうですねえ ○ ● 今日は 15 日ですから ○ ● (ええと) 五月は ● 後半 ● ずい分自由な日があるんですよ。○ どうぞ、おっしゃってください。○

Well, today is the 15th. As for the second half of May, I have lots of dates free. Please give me any day convenient to you.

### 4.2.2 Voices

Voice expresses the relationship between, on the one hand, the participant roles of the NP arguments of a verb and, on the other hand, the grammatical relations attributed to those same NPs.[5] Teramura proposes in [Teramura 91: vol. I] to admit the following five voices: 能動 (active), 受動 (passive), 可能 (able), 自発 (spontaneous or middle), and 使役 (causative). Of these, active and passive voices are observed in English, too. The third voice, able, is coded with modals such as *'can'* in English. As for the fourth voice, spontaneous, there is no way to encode the information gramatically in English and it depends on the verb and its argument whether such a usage is possible, e.g., *"Ice cream sells quickly in the summer."* [Teramura 91: vol. 1, pp. 282]. The last voice, causative, is encoded in English using various verbs such as *'force'* or *'let'*.

Although the spontaneous form is one of the characteristic features in Japanese, no use is observed in Verbmobil corpus. The reason is probably that the conversations concern only scheduling and there is no need of describing an event, where spontaneous forms are most prevalent such as 聞こえる *'kikokeru'* (to

[5] The definition is cited from [Trask 1993].

be heard) or 見える *'mieru'* (to be seen). In the following, therefore, we only examine examples of the able and causative forms.

The problems arise when the correspondance between the participant and thematic roles of NPs changes depending on voices. Choice of particles in Japanese is also complicated due to the interaction between verbs and NPs. In what follows, we examine the usage of the able and causative forms, each related to the former and the latter points, respectively.

**Able form** The able form expresses that an action or operation described by the verb is executable. The verb, 取る *'toru'*, is for example inflected as shown in Table 23 for active, passive, and able voices, respectively. Example (31) shows a use of able voice.

| voice | in Kanji | in roman | in English |
|---|---|---|---|
| active | 取る | tor-u | take |
| passive | 取られる | tor-areru | is taken |
| able | 取れる | tor-eru | can be taken |

Table 23: Voices of 取る toru

(31) yotei ga toreru
a time[NOM] can be taken
(I can find a time. 予定が取れる。)

[T8c] ichi ji sugi kara deshi tara nantoka yotei ga toreru N desu ga.

一時過ぎからでしたら何とか予定が取れるんですが。

If the meeting is held after one o'clock, I manage to find a time.

The crucial point is that the target of the verb, 取る *'toru'* (take), is expressed as nominative with the particle, が *'ga'*. To keep the NP, 予定 *'yotei'*, at the subject position, the voice has to be made passive in English, e.g., *"A time can be taken."*, but such an expression sounds odd and sentences such that the target is accusative are preferred, e.g., *"I can find a time."*

In translation, therefore, one has to convert an able form verb into an appropriate active form verb such that the target is naturally fit in as object and to fill in the subject with an agent, which is often suppressed and thus must be inferred.

A note on the term, *able*. The voice form does not imply possibility while English *'can'* does imply it. The sentence (31), therefore, does not mean that it is possible to find a time, but that the speaker has the ability. Thus, we choose the term, *able*.

**Causative form** The causative form expresses that an action or operation described by the verb is forced or allowed to do. Taking an example, the verb, 電話する *'deNwa-suru'* (call up), is inflected for active, passive, and causative forms as shown in Table 24 and sentence (32) means that the speaker is allowed to call up the addressee by his kind permission.

| voice | in Kanji | in roman | in English |
|---|---|---|---|
| active | 電話する | deNwa-suru | call up |
| passive | 電話される | deNwa-sareru | is called up |
| causative | 電話させる | deNwa-saseru | force someone to call up<br>allow someone to call up |

Table 24: Voices of 電話する deNwa-suru

(32) o deNwa sase te itadai ta
HONOUR phone[CAUSATIVE] (state) HONOUR (mood)
(I am allowed to call you up, Sir.
お電話させていただいた。)

[033]:ACB04: enomoto seNsee[6] • (ano) kyoo wa kooeNkai no uchi-awase nitsuite o deNwa sase te itadai ta N desu keredomo . ∘

ACB04: 榎本先生 • (あの) 今日は講演会の打合せについてお電話させていただいたんですけれども。

Enomoto sensee, I call you up today concerning a meeting for the lecture.

Althouth the meaning of 電話させる *'deNwa-saseru'* can be ambiguous between the forcible and permissive readings, the preference to the latter is obvious owing to the fact that the verb is followed by いただく *'itadaku'*, which expresses the speaker's respect to the receipient. (We disregard factors involved in the honorifics, state, and mood as they are not of our interest in this section.) In

[6]The word, 'seNsee', is usually transcribed as 'sensei', but we write it as 'seNsee' as appearing in Verbmobil corpus. The transcription, 'seNsee', reflects the way the speaker utters the word more precisely.

fact, no usage of forcible causative is found in Verbmobil corpus. We therefore conclude that the causative form may not be problematic, but need to investigate if the usage is restricted in new domains, too.

An explanation of the usage of causative form can be found in [Maynard 1990: pp. 372–376]. In short, the problem is concerned with the choice of particles to NPs, を *'wo'* (accusative) or に *'ni'* (dative). The rules presented there refer to types of verbs and direct objects among other things.

### 4.2.3 Verb constructions

Verbs can be constructed by two means. One way is to employ a light verb, する *'suru'*, which roughly corresponds to *'make'* in English. The other way is to combine two verbs to from a compound verb. We examine these two constructions in order.

**する suru** The light verb, する *'suru'*, follows a NP to form a verb. For example, 出版 *'shuppan'* (publication) is turned into a verb, 出版する *'shuppan suru'* (to publish), when followed by the light verb.

(33) hoN wo shuppaN suru
book[ACC] publication light verb
(to publish a book 本を出版する)

> [003]:NAZ02: kita ni desu kedomo . ○ (anoo) • ree no hoN o shuppaN suru tame ni (ano) • seesakuchuu no (ano) geNkoo o mochiyot te keNtoo shi tai N desu kedomo . ○
>
> NAZ02: 木谷ですけども。 (あのー) • 例の本を出版するために (あの) • 作成中の (あの) 原稿を持ち寄って検討したいんですけども
>
> I am Kitani. To publish the book, I wounder if we can hold a meeting, each of us bringing a manuscript.

Table 25 shows some examples of verb constructions with する *'suru'*. To each verb, an English translation is given in the third column and the translation of NP part in the fourth. Translations are mostly unproblematic; Constructed verbs are treated as a single verb. However, some choices must sometimes be made when a translation with a light verb is also possible in English. For example, 掲載する *'keesai suru'* can be translated to *'make a contribution'*, 電話する *'deNwa suru'* to *'make a telephone call'*, and 出席する *'shusseki'* to *'make one's appearance (at)'*.

| in Kanji | in roman | in English | NP | in corpus |
|---|---|---|---|---|
| 掲載する | keesai suru | to publish | publication | [046]:AAC02 |
| 投稿する | tookoo suru | to contribute | contribution | [049]:AAC02 |
| 調整する | choosei suru | to regulate | regulation | [061]:NAI04 |
| 招待する | shootai suru | to invite | invitation | [104]:ABB14 |
| 電話する | deNwa suru | to telephone | telephone | [105]:NAN13 |
| 出席する | shusseki | to attend | attendance | [115]:NAQ10 |
| 希望する | kiboo suru | to hope | hope | [169]:ABW09 |

Table 25: Compound verb constructions with する *'suru'*

In spoken language translation, the light verb construction is preferred in English for three reasons. One reason is that the sentence sounds more colloquial. Another reason is that an argument can be suppressed, e.g., the sentence, *"I made a telephone call"*, does not mention whom the speaker called up. Since the callee is often suppressed in Japanese, the construction is useful. The other reason is that it is easier to encode adverbial information. To understand the last point, let us compare two sentences:

(34) a. I telephone her twice.

b. I make two telephone calls to her.

Both sentences, (34a) and (34b), encode the same information that the speaker calls a person up two times, but the sentence (34b) sounds better probably because the relation between *'two'* and *'telephone calls'* is clearer.

To express the same proposition as above, the sentence construction is like (34a) in Japanese as shown in (35). The difficult part is, therefore, to generate a sentence like (34b) given a sentence like (35), where the syntactic structure of sentence changes.

(35) watashi ga kanojo ni ni do deNwa suru
I [NOM] her [ACC] two times to telephone
(I telephone her twice.
私が彼女に二度、電話する。)

**Compound Verbs** Some verbs can be combined to form a compound verb. Among various constructions we only consider a construction, $\text{Verb}_1$[infinitive] + $\text{Verb}_2$. The meaning of compound verbs is more or less derived from those of element verbs, but the relation between the two verbs is not trivial.

- 持ち - 寄る *'mochi yoru'* is composed of 持つ *'motsu'* (bring) and 寄る *'yoru'* (gather). The meaning is to gather somewhere each bringing something with him. An expression, 原稿を持ち寄って *'geNkoo o mochiyot te'* (to gather each bringing a manuscript), appears at [010]:AAF04.
- 待ち - 合わせる *'machi awase'* is composed of 待つ *'matsu'* (wait for) and 合わせる *'awaseru'* (join together) and means to wait for someone by appointment.
- 取り - 計らう *'tori hakarau'* is composed of 取る *'toru'* (take) and 計らう *'hakarau'* (arrange) and means to arrange a matter as one thinks fit.

The problem in translation is that in most cases there is no directly corresponding single verb in English. To express the information implied in compound verbs, some phrases have to be added to the sentence for explanation. As the result, translated sentences become explanatory and longer, e.g., 私が彼と待ち合わせる *"watashi-ga kare-to machi-awase ru"* is better translated to *"I wait for him by appointment"*, not simply to *"I wait for him"*.

# 5 Pragmatic phenomena

## 5.1 Viewpoint, honorifics and zero-pronouns

Various pragmatic factors can affect the way a single event is described. We consider here as an example how the translation of sentences containing *'give'* can be translated into Japanese. For the English verb, *'give'*, there are at least four possible translations in Japanese: *'kureru'*, *'kudasaru'*, *'ageru'*, and *'sashiageru'*. The use of these words can be summarised as follows:

| pov | normal | honorific | condition |
|---|---|---|---|
| receiver | *'kureru'* | *'kudasaru'* (giver > receiver) | giver $\neq$ 1st per. |
| giver | *'ageru'* | *'sashiageru'* (giver < receiver) | receiver $\neq$ 1st per. |

*'kureru'* is used when the event is described from receiver's point of view (pov) while *'ageru'* is used when the event is described from giver's point of view. *'kudasaru'* is honorific form of *'kureru'*, where the giver stands in higher position than the receiver. Both *'kureru'* and *'kudasaru'* can only be used when the giver is not the first person, typically the speaker in our domain. *'sashiageru'* is honorific form of *'ageru'*, where the giver stands in lower position than the receiver. Both *'ageru'* and *'sashiageru'* can only be used when the receiver is not the first person, say, the speaker.

The argument of these Japanese verbs can often be omitted. Most frequently, the reference to the speaker himself is suppressed when it corresponds to the

viewpoint, e.g., as the receiver of *'kureru'* and as the giver of *'ageru'*. The reference to the hearer, too, can be suppressed, e.g., as the giver of *'kureru'* and as the receiver of *'ageru'*. The same goes for *'kudasaru'* as *'kureru'*, and for *'sashiageru'* as *'ageru'*. In the worst case of translation from Japanese to English, therefore, both the giver and receiver have to be recovered from the context.

Figure 1 depicts the translation relation between *'give'* and {*'kureru'*, *'kudasaru'*, *'ageru'*, *'sashiageru'*}. The rule at the top node, *'ageru'* ⇔ *'give'*, is the default translation rule translating *'ageru'* to *'give'* and vice versa independent of any background condition. From here towards the bottom, the items of information considered in background increase. Moving down leftwards, if the viewpoint is at the receiver and the giver is not the first person, then the rule, *'kureru'* ⇔ *'give'*, is applicable.

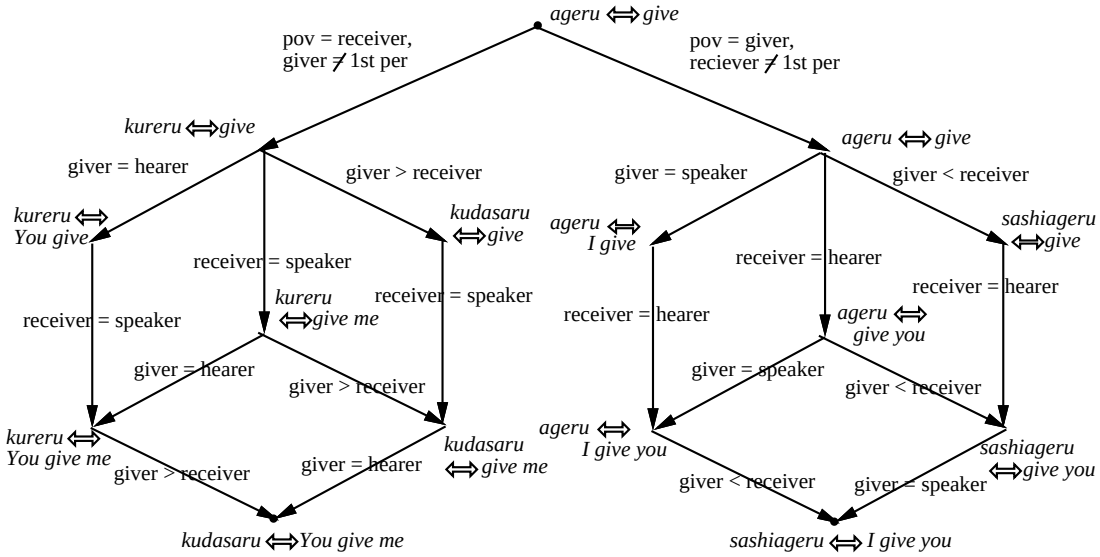


Figure 1: The translation relation concerning *'give'*

Down from here, three more refined rules are found. Moving down rightwards, if the giver stands in higher position than the receiver, the rule, *'kudasaru'* ⇔ *'give'*, is applicable because the honorific form is more appropriate in the context. The other two rules concern zero-anaphora. Backing to the previous point, if the giver is not explicitly realised in the sentence and is identifiable with the hearer, the second person pronoun, *'You'*, is introduced by *'kureru'* ⇔ *'You give'*. Similarly, if the receiver is not expressed and is identifiable with the speaker, the first person pronoun, *'me'*, is introduced by *'kureru'* ⇔ *'give me'*. These three rules are independently applicable as understood from the figure.

The similar relations hold for the cases of *'ageru'* and *'sashiageru'*.

### 5.2 hai/iie answer

The short answers, *'hai'* and *'iie'*, are usually translated to *'yes'* and *'no'*, but they have sometimes to be translated to *'no'* and *'yes'*, respectively. Unlike English *'yes/no'*, the choice between *'hai'* and *'iie'* does not reflect the polarity (positive or negative) of the proposition in question. The choice rather depends on whether or not the speaker agrees to the addressee. That is, *'hai'* means *'agreed'* and *'iie' 'disagreed'*. The difference becomes problematic when the question is of negative form and the speaker would like to agree or disagree with it as shown below.

(36) a. ashita kochira ni korare nai deshoo?
tomorrow here [DIR] cannot come (question)
You cannot come here tomorrow, can you?

b. hai muri desu
no impossible be
No, I can't.

c. iie korare masu
yes can come
Yes, I can.

Another difference is that *'hai'* can also be uttered at the place where *'OK'* is more appropriate in English. In this respect the Japanese *'hai'* is more like the German *'ja'*.

### 5.3 Dialogue acts

It is observed by many researchers that different sentences can mean the same thing or can convey the same intention by the speaker to the hearer. The observation has been developed as Speech Acts Theory and been adopted to analysing the meaning of sentences in dialogue as dialogue acts in Verbmobil [Jekat et al. 1995]. We have applied the ideas to analysing Japanese sentences in dialogue and found that the dialogue acts observed in the Japanese corpus are not dramatically different from those found in German or English. But there are some expressions characteristic to Japanese:

- A word for excuse is often uttered when the speaker turns down the other's suggestion, e.g., すいません *'suimaseN'* (Sorry or Entschuldigung), 申し訳

ございません *'mooshiwake gozai mase N'* (Sorry or Entschuldigung) and *'mooshiwake nai'*.

- A word for cooperation is quite often uttered when the speaker confirms their meeting, e.g., よろしくお願いします *'yoroshiku o negai shi masu'*. The expression is often translated to *'bye'* or *'Auf Wiederhoeren'*, but the literal meaning is to ask for one's good will.

There are also particular ways in negotiation. Namely, a speaker can turn down the other's suggestion by explaining the reason why it does not suit him, without explicitly saying *'no'*. The phenomena is explained in depth in [Alexandersson et al. 1997].

## 6 Conclusion

We have seen what can be problematic when translating sentences between Japanese and German/English. In presenting those problems we categorised them into three classes, i.e., lexical, grammatic, and pragmatic. Of these classes, we have still many problems to consider, especially for pragmatics. We could not discuss the issue of tense and aspect, either. We opt for another report for each subject separately to make the current report accessible.

We could not also discuss how we can approach to the problems presented here since our aim was to explain what the problems are. An approach to pragmatic factors, however, has been discussed by one of the authors in [Fujinami 1997], where he argues that conditional logic helps us to study how most appropriate expressions can be chosen in translation. We know of course that many other problems await us for a solution.

Despite of all these shortcomings, we believe that the report provides the reader with enough material to get an insight into the problems related to Japanese.

## References


Alexandersson, J., Buschbeck-Wolf, B., Fujinami, T., Maier, E., Reithinger, N., Schmitz, B., and Siegel, M.: 1997, *Dialogue Acts in Verbmobil-2*, Verbmobil report, DFKI Saarbrücken, forthcoming

Barwise, J. and Cooper, R.: 1981, Generalized quantifiers and natural language, *Linguistics and Philosophy* 4(2), 159–219

Fujinami, T.: 1997, Combining situated reasoning with semantic transfer minimally, in *Proceedings of Spoken Language Translation Workshop*, pp 92–93, ACL/EACL 97

Fukushima, K.: 1991, Phrase structure grammar, Montague semantics, and floating quantifiers in Japanese, *Linguistics and Philosophy* 14, 581–628

Glasbey, S.: 1995, A situation-theoretic interpretation of bare plurals, in *Proceedings of the Tbilisi Symposium*, Long version

Jekat, S., Klein, A., Maier, E., Maleck, I., Mast, M., and Quantz, J. J.: 1995, *Dialogue Acts in Verbmobil*, Verbmobil Report 65, Universität Hamburg, DFKI Saarbrücken, Universität Erlangen, TU Berlin

Kameyama, M.: 1996, Calendar time inferences and temporal nominal interpretation, in H. Nakagawa (ed.), *Research Report: A Cognitive Study of Situatedness In English and Japanese - Toward a Flexible Natural Language Processing System*, pp 13–28, Faculty of Engineering, Yokohama National University, Yokohama, Japan, An informal collection of papers resulting from a joint research project sponsored by the Japanese Ministry of Education, 1994-96

Kindaichi, H.: 1988, 日本語 *(Japanese) I and II*, 岩波書店 (Iwanami), in Japanese, 岩波新書, 金田一 春彦著

Makino, S. and Tsutsui, M.: 1995, *A Dictionary of Intermediate Japanese Grammar*, The Japan Times

Maynard, S. K.: 1990, *An Introduction to Japanese Grammar and Communication Strategies*, The Japan Times

Nagara, S., Hirota, N., and Nakanishi, Y.: 1989, 形式名詞 *Pseudonouns*, Vol. 2 of *Innovative Workbooks in Japanese*, 荒竹出版 (Aratake), 外国人のための日本語 例文・問題シリーズ

Teramura, H.: 1988–91, 日本語のシンタクスと意味 *(The syntax and semantics of Japanese) I–III*, くろしお出版 (kuroshio), in Japanese, 寺村 秀夫 著

Trask, R. L.: 1993, *A Dictionary of Grammatical Terms in Linguistics*, Routledge