TR-I-0269

English-Japanese Transfer by ASURA Framework (interim report)

Mutsuko Tomokiyo Masami Suzuki Noriyoshi Uratani July 1992

ATR Interpreting Telephony Research Laboratories © ATR Interpreting Telephony Research Laboratories

Abstract

English-Japanese transfer is done through Illocutionary force types which are intermediary concept between two languages: input data is feature structures from analysis results which are produced by the HPSG and unification grammars. The top level in semantic representation of a sentence is IFT. Input feature structures are changed into feature structures for target language feature stuctures by rewritting rules on transfer phase.

Almost of rewritting rules are defined with in and out. in represents logico-semantic pattern with which should be made pattern match operations of actually incoming logico-semantic pattern of source text. out represents logico-semantic pattern to be output of target language.

Transfer task bases on formalizing differences of semantic representation between source and target languages, by checking whole a sentence.

Acknowledgements

We would like to express our gratitude to Dr.Akira Kurematu for his comprehension about our research. Further thanks are due to him as well as Tsuyoshi Morimoto and Hitoshi Iida for helping us to write this report. Heartily thanks should go to Dr.Yves Lepage for seting up experimental environment and to all colleagues for interesting discussions.

Keywords and abreviations

feature structure, machine translation(MT), rewriting system(RWS), IFT, speech act, spoken Japanese, syntax, semantic, aspect, ellipsis, tense, topicalization

Research period of time

Transfer system was developed by T. Hasegawa as Japanese-English transfer system in November of 1990 in ATR. Our researh about English-Japanese transfer started in April of 1992 and transfer experiments are actually ongoing on SUN04 machine.

Contents

1	Inti	roduction	4	
2	Rev	vriting system	6	
3	Rewriting rules			
	3.1	Main rule and others	8	
	3.2	Main rule	8	
	3.3	Aspect and tense calculation rules	0.	
	3.4	Labeling rules	0	
	3.5	Production rules of the illocutionary force type 1	.1	
	3.6	Lexical rules	.3	
	3.7	Rule application order and rule workings	4	
		3.7.1 Order	4	
		3.7.2 Rule workings	.5	
4	Linguistic issues			
	4.1	Honorifics	.6	
	4.2	Topicalization	21	
	4.3	Pronoun ellipsis	23	
	4.4	Illocutionary force types	26	
		4.4.1 Constraints for IFTs 2		
		4.4.2 IFT variety	28	
		4.4.3 IFT and person		
		4.4.4 IFT and polarity		
		4.4.5 IFT and generality/specificity 3		
	4.5	Denial expression		
	4.6	Aspect and tense	15	

1 Introduction

This report aims to describe the overview of transfer phase activities in English-Japanese MT on ASURA system, as well as previous studies concerning the corpus grammatical and linguistic issues of taken corpus.

Transfer processes base on firstly extracting necessary information from analysis results¹, secondarily applying rewriting rules, finally outputting transfer results to generation phase. Transfer from English to Japanese consists largely in semantic transfer and pragmatic transfer. The former is done by intermediate languages which are invented as illocutionary force type². This comes at the top level of the semantic representation and its object is verbs of main sentence or such grammatical issues as the interrogative, negative, etc. Pragmatic transfer consists in changing communicative conditions between English and Japanese languages.

Transfer by IFT depends on which kind of IFTs are extracted from semantic feature structures. The good point of this method is to be able to neglect syntactic differences existing between source and target languages in transfer process. The problem is that there exists difficulty in determining unique IFT from extracted information. For example, grammatical negation is not necessarily negation from the viewpoint of IFT and also interrogative utterance is not necessarily questions from speaker. Each of IFT labels should be conditioned with pertinent de-

```
[[PARM ?X]
[RESTR [[RELN CONFERENCE_OFFICE-N]]
[ENTITY ?X]]
```

However, this formulae serves nothing for whole process of MT. And so in the experiments, the semantic representation of nouns is simplified as follows:

[[RELN CONFERENCE_OFFICE-N]]

¹English-Japanese analysis is not yet done. Analysis results which are used in the experiments are one that was previously presumed, by looking at Japanese-English transfer results of sentences in the same corpus. E-J analysis results have minimal changes from J-E analysis results: for example, in J-E semantic representation formulae, a noun is represented, based on situation semantics:

²Illocutionary force type (IFT) will be mentioned later.

tails, when a IFT is taken. This is ongoing research and so difficulty stays in determining which kind of contexual situation a IFT could cover.

Properly speaking, transfer rules should be rules to be applied beyond one sentence: sentences as an utterance should be taken as transfer unit, at least a unit of 2 sentences, or caller-and- receptionist utterance unit in our task domain. We will have two steps as transfer phase activities: firstly, application of sentence-by-sentence transfer rules, secondarily application of rules concerning the sentence units. This report will describe only the sentence-by-sentence transfer and so the transfer including discourse issues is omited, because it's not yet implemented in our system.

Our task domain is telephone conversation of which topic is limited to inquiries about international conference. Speakers on telephone are a caller and receptionist of the conference office.

2 Rewriting system

out.

Rewriting system is a system in rewriting feature structures by rules defined by unique formulae: rewriting rule is constitued with rule index, rule definition and rule body.

```
e.g. EXPECT-VT-1
(rws:defrwschema2 def528 V EX
                                      rule index
"on <OBJE RELN> EXPECT-VT-1 in :PHASE :E-J :Type :Default
                                      Rule definition
 in= [[RELN UNKNOWN-IFT]
                                      rule body
      [AGEN ?AGEN]
      [RECP ?RECP]
      [OBJE [[reln EXPECT-VT-1]
              [SUBJ ?SUBJ]
              [OBJE ?OBJE]
              [COMPLEMENT ?COMPLEMENT]
              ?rest]]]
                      set parameter :IFT :INFORM
 out= [[RELN INFORM]
        [OBJE [[RELN 予定する-1]
               [SUBJ ?SUBJ]
               [OBJE ?OBJE]
               [COMPLEMENT ?COMPLEMENT]
             ?rest]]]
end")
   Rule definition is constituted with
     feature path name: OBJE RELN
     feature name: EXPECT-VT-1 and
     rule application condition: in :PHASE :E-J :TYPE :default.
Rule body is constituted with
     in
            and
```

in, the right side of of an equation is semantic feature structure as input and out is also semantic feature structure as output. Analysis results of source language are checked by pattern matching operations with feature structures of in. When pattern matching is successfully done, input feature structures are rewritten into such feature structures as shown by out feature structures. out feature structures are target language structure corresponding to source text.

3 Rewriting rules

This section describes rewriting rule varieties and aims of using these rules.

3.1 Main rule and others

Rules are built to rewrite purposively analysis results into generation input forms. They are rules which bridge between source and target languages by rewriting feature structures, since semantic representation is described with logico-semantics by feature and its value. Input feature structures represent semantic structure of source language and output is one of target language.

There exists firstly main rule to set up the environment where others rules function. This rule contains rule application order and rule application way: recursively, loop or once.

Other rules are as follows according to their aims:

IFT rules
Labeling rules
Tense and Aspect rules
Prag-feature rules
Idiom rules
Lexical rules

3.2 Main rule

We assume that input feature structure which is object of the transfer is as follows:

e.g. This is the conference office.3

³This example will be used in whole this section.

```
[IDEN [[RELN CONFERENCE_OFFICE-1]

[INDEX [[DETEM SPECIFIED]

[GENDER NEUT]

[PERSONA 3RD]

[NUMBER SING]]]]]]]]]

[PRAG [[HEARER !X3]

[SPEAKER !X2]]]]
```

Main rule sets up rule application environment for IFT and transcribes SEM of these structures into following structures:

The value UNKNOWN-IFT is at the top of the semantic representation and will be filled by IFT label in good time. Speaker and hearer labels are also set up for agen and recp. This means that agen and recp are filled up with speaker and hearer and that deixis I and you are speaker and hearer of the conversation, respectively. IFT is speaker or hearer's speech intention and doesn't necessarily belong to clear-cut linguistic system.

3.3 Aspect and tense calculation rules

Aspect calculation is in general done in analysis phase. However, English aspect is independent on tense, contrastively with Japanese aspect which includes tense concept. Furthermore verb index which is necessary for aspect calculation are used only for it. This isn't information which would be supplied with lexicon. This is the reason why the aspect calculation is done as transfer component.

Aspect rules are constituted with 4 different types of rules.

1. initialization rules which serves:

to combine tense and grammatical aspect: perfect, nonperfect, state, progressive

2. verb index labeling rules which serve:

to add verb index at the point of view of time concept for each verb

3. English aspect rules which serve:

to determine English aspect

4. Japanese aspect rules which serve:

to produce Japanese aspect

Rules 1,2,3 are applied at the first stage of the rewriting proces and Rule 4, after general or default rule application. Aspect and tense are mentioned below.

3.4 Labeling rules

Labeling rules are needed for adding linguistic information to analysis results according to different linguistic phenonema between two languages: for example, this of deixis as subject of a sentence is translated into Japanese with topic indication marker wa in certain context. For indication of particle usage of wa or ga in Japanese, such a concept as topicalization is introduced as common paradigm between source and target languages. Be-verb expresses a current state. When the pronoun

this as subject is combined with current state verbs, topicalization occurs in a sentence. Labeling rules serve to indicate the paradigm by labeling to concerning words and to use it for the sake of having more japanese-like translation result.

Almost of Labeling rules are as follows:

Labeling rule takes back following rewriting result, because of rules for the pronoun this and the verb BE-VI-5 in this context. TOPIC indication features are added to the pragmatic feature structures and it has OBJECT, TOPIC_MODE and SCOPE as values.

```
[SPEAKER !X1]
```

[PRAG [[HEARER !X2]

[TOPIC [[OBJECT !X1] [TOPIC-MODE WA] [SCOPE !X3]]]]]

3.5 Production rules of the illocutionary force type

IFT depends on what illocutionary act has a token from the view point of situation semantic. IFT is concretely established by setting up the parameter to input feature structures. IFT labels are actually as follows:

Request Inform Question

Promise

Phatic

Response

Expressive

The same IFTs as mentioned above are taken in J-E transfer on ASURA system. However, it remains to conduct further examination of illocutionary forces.

Transfer result by using IFT rule is as follows:

[[RELN INFORM]] which is located at the top level of semantic representation, means simple statement from *SPEAKER* to *HEARER*. inform label is default value and other values are determined by context-dependent cases.

```
[[SEM [[RELN INFORM]
      [AGEN !X1[[LABEL *SPEAKER*]]]
      [RECP !X2[[LABEL *HEARER*]]]
       [OBJE !X3[[RELN CURRENT_STATE]
                 [OBJE [[RELN だ-IDENTICAL]
                        [TENSE PRESENT]
                        [OBJE [[RELN OUTWARD_POINTING]
                               [OBJE [[RELN とちら-PRON-1]]]]]
                        [IDEN [[RELN NAMED]
                               [IDEN [[RELN 会議事務局-1]
                                      [INDEX [[DETEM SPECIFIED]
                                              [GENDER NEUT]
                                              [NUMBER SING]
                                              [PERSONA
                                                3RD]]]]]]]
                        [SEM-ASPE JYOUTAI]]]]]
```

3.6 Lexical rules

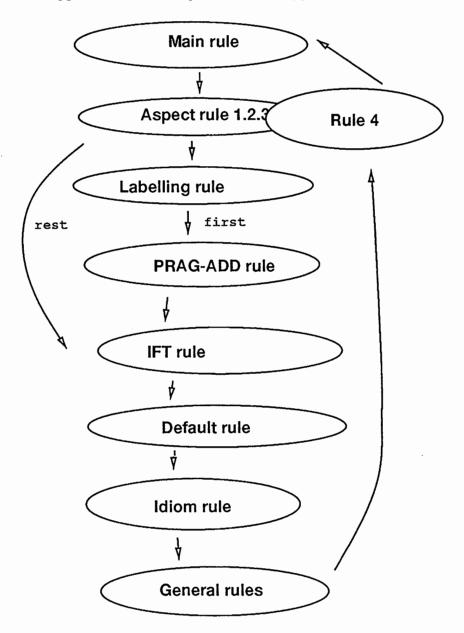
Lexical rules switch lexies from English to Japanese. For example, rule of conference_office is as follows:

```
(rws:defrwschema2 def303 N CON
"on <IDEN RELN> Conference_Office-1 in :PHASE :E-J :TYPE :default
 in= [[RELN NAMED]
      [IDEN [[RELN Conference_Office-1]]
             [INDEX [[DETEM SPECIFIED]
                     [GENDER NEUT]
                     [PERSONA 3RD]
                     [NUMBER SING]]]
                     ?REST]]]
out= [[RELN NAMED]
      [IDEN [[RELN 会議事務局 -1]
             [INDEX [[DETEM SPECIFIED]
                     [GENDER NEUT]
                     [PERSONA 3RD]
                     [NUMBER SING]]]
                            ?REST]]]
end")
```

3.7 Rule application order and rule workings

3.7.1 Order

Rule application order depends on rule type or transfer convenience:



3.7.2 Rule workings

Main rule works at the beginning stage and the last stage of transfer process. Other rules work recursively and/or in loop. And some kind of rules such as labeling rule works only recursively. One of pragmatic rules works only once.

4 Linguistic issues

This section describes what kind of linguistic phenomena are currently handled on transfer phase of English-Japanese MT on ASURA.

4.1 Honorifics

Honorifics are constituted with politeness and formality. English politeness is expressed by:

- a. Intonation
- b. Tense
 - e.g. When will you be visiting us again?
- c. Question form
 - e.g. Would you please shut the door?
- d. Permission form
 - e.g. May I suggest that we postpone the meeting until tomorrow?
- e. Possibility form
 - e.g. Could I suggest that we postpone the meeting until tomorrow?
- f. Sentence style
 - e.g. I wonder if you would mind shutting the door?
- g. Adding words or a chain of words
 - e.g. please
- h. Negative and interrogative form
- i. Full-sentence form

English conversation tends to be more polite when someone is talking to a person he does not know well, or a person senior to himself in terms of age or social position. Japanese conversation is the same. However, Japanese is more delicate in a way to make a polite attitude. For example in

Japanese, there is polite expression towards 3rd person who is not on the actual spot and is concerning person to hearer. The other hand, English polite is more tactful: tactful imperative form is used, for example between a boss and his secretary in order to avoid causing offense or distress to the later in business. In Japanese, there doesn't exist such kind of polite usage.

English formality is expressed according to cases by:

a. Formal words

quantifier: many, much

pronoun: we

time and place: within

adverb: subsequently, ultimately

frequency: per

manner, means: with what...

cause: on account of purpose: in order condition: in case of contrast: whereas ,etc.

b. Verbs

e.g. wish

- c. Permission form
 - e.g. May I...
- d. Sentence style

e.g. listing - firstlysecondarily,

Japanese politeness is expressed by:

- a. Polite auxiliaries:
 - e.g. teitadakitainodesuga
- b. Negative and interrogative form
 - e.g. teitadakemasendeshouka
- c. Prefix

e.g. o-name

d. Full-sentence form

Japanese formality in conversation is expressed by:

- a. Formal auxiliaries
 - e.g. desu, masu
- b. Kanji-word usages

English formality is used mainly in writen language, contrastively with Japanese formality which occurs in spoken language. And spoken formality is used as a kind of politeness. Formality is represented by pragmatic features, while being included in politeness representation. Politeness and formality by lexies are analyzed in analysis phase and the other formal and polite expressions are handled at the transfer phase, by rewriting semantic features and then moving them to pragmatic features. Actual workings are shown by a lexie PLEASE-INTERJ-1.

e.g. Please tell me about the topic of the Conference.

```
[[SEM [[RELN IMPERATIVE]
       [AGEN !X1[[LABEL *HEARER*]]]
       [RECP !X2[[LABEL *SPEAKER*]]]
       [OBJE [[RELN TELL-VT-2]
              [ASPT NON_PERFECT]
              [TENSE PRESENT]
              [OBJE !X2]
              [TOPIC [[reln OF-MODIFICATION]
                       [ARG-1 [[RELN NAMED]
                               [IDEN [[RELN CONFERENCE-N-1]
                                      [INDEX [[DETERM SPECIFIED]
                                               [GENDER NEUT]
                                               [NUMBER SING]
                                               [PERSONA 3RD]]]]]]
                      [AGR-2 [[RELN TOPIC-N-6]
                              [INDEX [[DETERM SPECIFIED]
                                      [NUMBER SING]
```

```
[GENDER NEUT]
[PERSONA NULL]]]]]]
[ATTITUDINAL [[RELN PLEASE-INTERJ-1]]]]]]
R !X1]
```

[PRAG [[HEARER !X1] [SPEAKER !X2]]]]

Sentence, Please tell me.. is polite imperative. Following rule for please is applied and attitudinal feature is removed from input features. Honorific features are added to pragmatic features.

Rule:

[PRAG [[HEARER ?X1] [SPEAKER ?X2]]]]

delete ATTITUDINAL from input.sem

add [HONORIFIC [[RELN POLITE]

[RESPECTANT ?X1]

[RESPECTEE ?X2]

[STRATEGY [[OBJECT ?OBJE]]]]] to

input.prag

 [RESPECTANT ?X1]
[RESPECTEE ?X2]
[STRATEGY [[OBJECT ?OBJE]]]]]]]]]

Honorifics have three values: respectee (receiver of politeness), respectant (addresser of politeness) and strategy(representation way). In Japanese, it might be possible to make a degree of politeness such as polite, more polite and the most polite. However, common point to be found between two languages, is which kind of strategies are taken for expression of politeness. So [STRATEGY [[OBJECT ?OBJE]]] is taken as values of the structure.

Transfer result:

```
[[SEM [[RELN REQUEST]
       [AGEN !X2[[LABEL *SPEAKER*]]]
       [RECP !X1[[LABEL *HEARER*]]]
       [OBJE [[RELN 教える-1]
              [TENSE PRESENT]
      [SEM-ASPE MUSOU]
              [OBJ !X2]
              [TOPIC [[RELN OF-MODIFICATION]
                      [ARG-1 [[RELN 内容-1]
              [INDEX [[DETERM SPECIFIED]
              [GENDER NEUT]
      [NUMBER SING]
      [PERSONA 3RD]]]]
      [ARG-2 [[reln 会議-1]
      [INDEX [[DETERM SPECIFIED]
              [GENDER NEUT]
              [NUMBER SING]
      [PERSONA 3RD]]]]
                                                  1111111
 [PRAG [[HONORIFIC [[RELN POLITE]
                [RESPECTANT !X2]
                [RESPECTEE !X1]
                [STRATEGY [[OBJECT ?OBJE]]]]
        [HEARER !X1]
        [SPEAKER !X2]]]
```

4.2 Topicalization

In Japanese, there exits wa-type topicalization and ga-type one which are obligatory in some context: particle wa and ga indicate topicalized part of speech.

e.g. This is the conference office.

this and BE-verb are rewritten by using labeling rules as mentioned above, and Topic indication is made in pragmatic features.

Transfer result is as follows:

```
[[SEM [[RELN INFORM]
        [AGEN !X1[[LABEL *SPEAKER*]]]
        [RECP !X2[[LABEL *HEARER*]]]
        [OBJE !X3[[RELN CURRENT_STATE]
                  [OBJE [[RELN だ-IDENTICAL]
                         [TENSE PRESENT]
                         [OBJE [[RELN OUTWARD_POINTING]
                                 [OBJE [[RELN とちら-PRON-1]]]]]
                         [IDEN [[RELN NAMED]
                                 [IDEN [[RELN 会議事務局
-1]
                                        [INDEX [[DETEM SPECIFIED]
                                                [GENDER NEUT]
                                                [NUMBER SING]
                                                [PERSONA
                                                    3RD]]]]]]]
                         [SEM-ASPE JYOUTAI]]]]
 [PRAG [[HEARER !X2]
        [SPEAKER !X1]
        [TOPIC [[OBJECT !X1]
                            [TOPIC_MODE HA]
                            [SCOPE !X3]]]]
```

e.g. There's a special application form.

In this example, there which is called introductory-there is obligatorily translated into topicalized Japanese with marker ga.

So introductory there is labeled as shown below and finally ga-topicalization is represented in pragmatic feature in the same way as topicalized-wa.

[[RELN INTRODUCTORY] [OBJECT THERE-ADV-1]]

4.3 Pronoun ellipsis

There are two varieties of pronoun ellipsis to be resolved in English-Japanese MT process: ellipsis in source language and ellipsis to be done in target language and two ellipsis aren't same. In our corpus, Conversation-a,B,1-10 there are 246 sentences and sentences with ellipsis in English version is only one sentence, contrastively with Japanese version which has 65 sentences(26%) of subject pronoun omission. Therefore the question is what is obligatory ellipsis in Japanese: following omission are observed in Japanese version of our corpus.

In a affirmative sentences when,

- a. grammatical subject is 1st person singular and modal auxiliary means desire.
- b. grammatical subject is 1st person singular and modal auxiliary means humble polite.
- c. grammatical subject is 1st person and predicate verb is performative one.
- d. grammatical subject is 2nd person in conditional clause.
- e. grammatical subject is 1st person singular and modal auxiliary means intention.
- f. grammatical subject is general person which is not fixed.

In interrogative sentences,

- a. grammatical subject is 1st person and modal auxiliary means humble polite.
- b. grammatical subject is 2nd person and modal auxiliary means respect.

Contrarily, ga or wa case in Japanse is regarded as being obligatory in following contexts:

a. grammatical subject as topic of old information in a discourse...wa

- b. grammatical subject as contrast of new information in a discourse...wa
- c. grammatical subject as exhaustive listing...ga
- d. grammatical subject of copula sentence....ga, wa
- e. emphatic subject...deha, de
- e.g. I think I'd like to present a paper at the conference.

This is translated into following Japanese:

e.g. Kaigini ronbunwo happyousitaito omotte irunodesuga.

The subject of omou (think) and happyousitai (would like to present) are omited in Japanese version. The subject of think is 1st person singular and think is a kind of commissive verb⁴. The subject of would like to (...tai) also is 1st person singular and means personal emotion. When commissive verbs occur on surface form with 1st person singular, the sentence subject by pronouns is omoitted in Japanese.

The rule concerning Japanese ellipsis productions is as follows:

[PRAG [[HEARER ?RECP] [SPEAKER ?AGEN] ?rest]]]

⁴Taxonomy of verbs will be mentioned below.

```
add [ELLIPSIS [[RELN ELLIPSIS] [OBJECT ?SUBJ]]] to input.prag
```

Initial pragmatic features are changed into following feature structures by this rule:

```
[PRAG [[HEARER !X2]
[SPEAKER !X1]
[ELLIPSIS [[RELN ELLIPSIS]
[OBJECT !X1]]]]]
```

The other case of ellipsis is genitive of pronouns in Japanese.

```
e.g. May I have your address, please?
```

Pronoun your is omitted in Japanese version and your address is translated into gojyuusho. Prefix go is marker of politeness expression and so it's clear who's address is asked. So rule for your has index of 2nd person and ellipsis in Japanese is indicated by following rule.

⁵We can write only complex type for add function at the current version. However it will be revised. In fact it isn't needed to make complex type description in this example.

4.4 Illocutionary force types

IFT is defined in this report as logico-semantic patterning labels of token, while standing on the side of speaker who utters a sentence. The logico-semantic means we deal with trichotomy of syntax, semantics, and pragmatics for semantic representation through whole a framework of ASURA. IFT caracterizes the semantic representation. IFT also contains communicative competence that doesn't belong to linguistic system proper. Strictly speaking, one which belongs to the linguistic system and one which belongs to the use of linguage should be separated. IFT has no the boundary between both, because it's convenient to limit communicative environment and to use it in order to describe meanings for the sake of transfering source into target linguage in MT.

The concept of illocutionary act is borrowed from Searle's theory. It is formalized as F(P), where F is illocutional force and where P is propositional content. This formulae should be controlled by following four conditions: propositional content condition, preparatory condition, sincerity condition, essential condition⁶. For example, promise is restricted with future tense as propositional content condition.

Futhermore, IFT depends largely on stress of token. However, we don't currently deal with accents, intonations and pitches features for logico-semantic structures. So this direction is one of important researches in future.

4.4.1 Constraints for IFTs

Each IFT has necessary conditions in order to be established. IFT is valid with following conditions:

Preparatory condition Propositional content condition

Preparatory conditions restrict the scene of a conversation:

⁶Sincerity condition and essential condition by Searle are presumed to be conditions satisfied in our domain.

AGEN is *speaker* and RECP is *hearer*. Relationship of *speaker* and *hearer*

Propositional content conditions restrict poropositional content:

tense
aspect
subject type etc.
noun index
verb index
adverb type etc.

Propositional content conditions depend on each IFT, contrastively with preparatory conditions which can be determined previously. Restrictions are represented in feature structures of rule body.

Request: e.g. [[verb [current state -]], [tense future][voice active][subject *HEARER*], [prag HONORIFIC]
Question: e.g. [semantic label [question type]]
Promise: e.g. [verb [verdictives +]], [tense future or infinitive] [subject *SPEAKER*]
Inform: e.g. [grammatical label [negate or not]]
Question: e.g. [semantic label [question type]]
Phatic: e.g. [pos interjection]
Response: e.g. [grammatical label [affirmative +]]⁷
Expressive: e.g. [subject *SPEAKER*][tense past]⁸

IFT should be necessarily detailed according to these restriction conditions.

e.g. order, request, demand, Plea

⁷Interrogative response is generally possible. However our corpus has no one. ⁸In case of apology

4.4.2 IFT variety

IFTs are determined with lexies, grammatical forms, semantic labels, pragmatic labels and transfer conventions.

Lexies mean, for example which category a verb belongs to in case of verbs. Following taxonomy for verbs⁹ of our corpus is actually taken¹⁰.

- e.g. state, describe, claim, explain, assert, remark, boast, understand, tellpresent, specialize, etc. (representatives)
- e.g. request, command, ask, order, pray, plead, beg, entreat, permit, advice, direct, prohibit, forbid, need, etc. (requestives)
- e.g. advice, recommand, suggest, counsel, warn, caution, indicate, etc. (advisives)
- e.g. thank you, congratulate, condole, apologize, deplore, welcome, etc, (expressives)
- e.g. permit, authorize, allow, grant, etc. (permissives)
- e.g. declare, marry, resign, christen, appoint, nominate, etc (declaratives)
- e.g. promise, vow, pledge, convenant, contract, swear, gurantee, embrace, shall, favor, bet, consent, oppose, intend, mean to, threaten, resove, expect, etc. (comissives)
- e.g. bet, offer, wager, etc. (propositive)
- e.g. send, apply, have, attend, do, use, pay, follow, contain, deposit, include, make, take, check, stay, etc. (verdictives)

Grammatical forms mean whether a sentence is:

- affirmative or negative
- interrogative or declarative

⁹This list contains some verb other than our corpus.

¹⁰This taxonomy is done by refering to classification of illocutionary act by Austin, Katz and Searle.

- indicative, imperative or subjunctive
- coordinate or subordinate clause ,etc.

Semantic labels mean what kind of interrogative sentence, what kind of verbs, what kind of nouns, or what kind of adjunct clauses, etc.

Almost of information about grammatical forms and some kind of semantic labels come from analysis phase. However, in general IFT is related with whole a sentence feature structures. Therefore some kind of labels or index are added as local information on transfer phase and IFT is determined while checking whole a sentence.

e.g. a. Please tell me about the topic of the Conference.

(Kaigino naiyounituite osietekudasai)

e.g. b. I would like to like to apply for the Conference.

(Kaigini mousikomi tainodesuga)

e.g. c. May I have your name and your address?

(Onamaeto gojyushowo onegaisimasu)

Example a.b, c have imperative, declarative, interrogative sentence types, respectively as grammatical form, but IFT of three sentences is request. Example a. is rewritten by imperative rule and b. is done by semantic label [[RELN 1st_PERSON_WISH]] with OBJE which has the same subject as [[RELN 1st_PERSON_WISH]] as shown below. Underlined value, [AGEN !X07] in semantic features is tagged with AGEN of IFT. c. is done by interrogative rule and permit rule.

Input of a.:see above

Input of b.:

```
[OBJE [[RELN APPLY_FOR-VT-1]

[SUBJ !X07]

[OBJE [[RELN NAMED]

[IDEN [[RELN CONFERENCE-N-1]

[INDEX [[DETERM SPECIFIED]

[NUMBER SING]

[PERSON NULL]

[GENDER 3RD]]]]]]]
```

[ASPT NON_PERFECT]]]]]]

[PRAG [[HEARER !X13] [SPEAKER !X07]]]]

Rule to be applied for a. is as follows:

SET PARAMETER : IFT : REQUEST

Rule to be applied for b. is as follows:

in= [[RELN UNKNOWN-IFT]
[AGEN ?AGEN]

SET PARAMETER : IFT : REQUEST

These two requests are not the same. First request is direct_request and the second is hypothetical_request. The difference of them is shown by honorific in pragmatic feaures.

Finally transfer conventions mean such a case as there is conversational implication. A IFT is classified into some details on purpose to be translated into more real Japanese. English version a, b, c are corresponding to different surface expressions in Japanese. The difference among three English expressions lies in whether hypothetical conditional clause is hided or not, contrastively with the difference of Japanese versions which lies in direct or indirect request manners. Therefore request is indirectly classified into hypothetical_request and direct-request.

- a. would like to ... tainodesuga
- b. Please... tekudasi
- c. May I....onegaishimasu

4.4.3 IFT and person

IFT varies according to the person and the polarity when it occurs in a interrogative or declarative sentence. Following examples show IFT of would like to depends on the person.

- a. I would like to apply for the conference. (request)
- b. You (He, She, They) would like to apply for the conference. (inform)

In example a. speaker is asking to apply for the conference. In example b. speaker is speaking for 2nd or 3rd persons. IFT of two sentences is completely different because of grammatical persons. And so would like to firstly is rewritten to 1st-person-wish and finally 1st-person-wish is rewritten into request.

Rule: would like to

4.4.4 IFT and polarity

IFT varies according to the polarity of a sentence.

e.g. a. I would like to apply for the conference. (request)

e.g. b. I wouldn't like to apply for the conference. (inform)

Sentence a. is asking to apply for the conference, comparing with sentence b. which is simple statement.

Affirmative form isn't explicitly conditioned in rules and denial porality has Negate label.

4.4.5 IFT and generality/specificity

IFT varies according to whether marked or unmarked sentence style is, especially in the interrogative sentence or comparative form.

e.g. a. Why don't you send me a registration form?

(request or complain)

e.g. b. Why do you send me a registration form? (question)

In general affirmative form is an unmarked sentence in declarative style. However, for example in comparative sentence,

e.g. a. Paul is taller than Michel.

e.g. b. Paul is shorter than Michel.

As for sentence a. it's possible that Paul and Michel, both are shorter than the other. Sentence b. Paul and Michel are comparatively shorter than the other. tall is an unmarked adjective and short is a marked adjective in comparative sentence. This means there are adjectives which indicate generality/specificity (marked/unmarked), when it's the predicate of comparative form. IFT of example a. b. is simple statement, but there remains the difference of implicastion between two. Specifity by negation is transferred, but one by adjectives particularity is not yet tried.

4.5 Denial expression

Adverbs of degree involves with explicit denial expression in the point of view of lexical choice of the target language.

```
e.g. a. I don't understand Japanese at all. e.g. b. It's interesting that he came at all.
```

Adverb clause at all in negative sentence will be translated into Japanese as zenzen and in affirmative sentence it will be translated as somosomo. Therefore lexical rules of at all are separately described in 2 ways, as shown by cases of not ...at_all and at_all

```
(rws:defrwschema2 idiom5037 PP AT
"on <OBJE MANN RELN> AT_ALL-1 in :PHASE :E-J :TYPE :IDIOM
    in= [[RELN NEGATE]
         [OBJE [[RELN ?RELN]
                [MANN [[RELN AT_ALL-1]]]
                        ?rest]]]
   out= [[RELN NEGATE]
         [OBJE [[RELN ?RELN]
                [MANN [[RELN 全然 -1]]]
                  ?rest]]]
END")
(rws:defrwschema2 idiom5100 V MO
"on <OBJE MANN RELN> AT_ALL-1 in :PHASE :E-J :TYPE :IDIOM
    in= [[RELN ?RELN]
         [MANN [[RELN AT_ALL-1]]]
                        ?rest]
   out= [[RELN ?RELN]
         [MANN [[RELN そもそも -1]]]
                         ?rest]
END")
```

4.6 Aspect and tense

Meanings of English aspect and time are determined by whether predicate of a sentence is perfective form or not and verb meanings. Following list shows relationship between aspectual meanings and surface forms.

A.present

- 1.State (present tense)11
- 2. Single event (present tense)
- 3. Habitual (present tense)
- 4. Temporary (present progressive)
- 5. Temporary habit (present progressive)

B.past

- 1.State up to present time (present perfect)
- 2.Indefinite event (present perfect)
- 3. Habit up to present (present perfect)
- 4. With present result (present perfect)
- 5. Temporary state up to present time (present perfect progressive)
- 6. Temporary habit up to present time (present perfect progressive)
- 7. Temporary, with present result (present perfect progressive)
- 8.Definite state (past tense)
- 9.Definite event (past tense)
- 10.Definite habit (past tense)
- 11.Definite temporary (past progressive)
- 12.Past before past time (past past perfect)
- 13.State up to past time (past past perfect)
- 14. Temporary state up to past time (past perfect progressive)

C.future

¹¹comments with brackets express English surface form

- 1. Future time (neutral-will, shall)
- 2. Future time (arising from present time-be going to)
- 3. Future time (plan or arrangement-present progressive)
- 4. Future time (as fact simple present)
- 5. Future time (as matter of course-will/shall+progressive)
- 6. Future time (temporary-will/shall+progressive)
- 7. Past in Future time (will/shall+perfect)

Japanese aspect/tense is as follows:

A.present

- 1.State(RU-form)
- 2. Single event (RU-form)
- 3. Habitual (RU-form)
- 4. Temporary (TEIRU)
- 5. Temporary habit (TEIRU)

B.past

- 1.State up to present (TEIRU)
- 2. Indefinite events (TAKOTOGAARU)
- 3. Habit up to present time (TEIRU)
- 4. With present result (TESIMATTA)
- 5. Temporary state up to present time (TEIRU)
- 6. Temporary habit up to present time (TEIRU, HAJIMETA)
- 7. Temporary, with present result (TEIRU, TEKITA)
- 8. Definite state (TEITA)
- 9. Definite event (TA)
- 10.Definite habit (TEITA)

- 11.Definite temporary (TEITA)
- 12.Past before past time (TEITA)
- 13. State up to past time (TEITA)
- 14. Temporary state up to past time (TEITA)

C.future

- 1. Future time as neutral (RU-form, DAROU)
- 2. Future time as arising from present time (UTOSITEIRU, SOUDA)
- 3. Future time as plan or arrangement (RU-form, KOTONINATTEIRU, TUMORIDA)
- 4. Future time as fact (RU-form)
- 5. Future time as matter of course (RU-form)
- 6. Future time as temporary (RU-form)
- 7. Past in Future time (TEIRUDAROU)

Aspectual meanings vary according to surface form (perfect/non-perfect, progressive, state), tense and verb meanings. Therefore verbs has classified with following index: +STATE, +/-CONT(tenious), +/-PERFECT(ive), +/- PERFORM(ative), +ACT(ive).

These index are put at the first stage of aspect calculation by labeling rules.

e.g. This is the conference office.

Firstly, Rule.1 and rule.2 is recursively applied to verb BE-VI-5 and internal structures change into following state.

Be-verb has an index [[STATE +]] and interim feature structures are as follows:

```
[GENDER NEUT]
[NUMBER SING]
[PERSONA 3RD]]]]]]]]
```

Secondarily, Rule.3 is applied and then following calculated result is produced.

Finally, rule 4 which is Japanese aspect is applied to Japanese verb and [SEM-ASPE [[STATE +]]] changes into [SEM-ASPE [[JOUTAI +]].

```
[SEM [[RELN INFORM]

[AGEN !X1[[LABEL *SPEAKER*]]]

[RECP !X2[[LABEL *HEARER*]]]

[OBJE [[RELN だーIDENTICAL]

[TENSE PRESENT]

[OBJE [[RELN OUTWARD_POINTING]

[OBJE [[RELN とちらーPRON-1]]]]]

[IDEN [[RELN NAMED]

[IDEN [[RELN NAMED]

[IDEN [[RELN 会議事務局-1]

[INDEX [[DETEM SPECIFIED]

[GENDER NEUT]

[NUMBER SING]

[PERSONA 3RD]]]]]]]
```

[SEM-ASPE JYOUTAI] is Japanese aspect name and [TENSE PRESENT] is correspond to Japanese tense RU-form.

ANNEX 1

The list describes grammatical compendium to which our courpus belongs, for the sake of future researches to be done:

Y means grammatical items are included within our corpus Conversation, A, B, 1-10. N means grammatical items are not included.

- 1. Adjective patterns
- 1.1 prepositional phrase Y
- e.g. be close to NP
- 1.2 that-clause N
- e.g. I'm sure that he'll be late.
- 1.3 to-infinitive or -ing N
- e.g. He was furious to hear about it.
- 2. Adjectives
- 2.1 Attributive adjectives Y
- e.g. extensive sphere
- 2.2 Predicative adjectives Y
- e.g. be close to
- 2.3 Postmodifiying adjectives Y
- e.g. research related to Interpreting Telephony
- 2.4 Adverbial adjectives N
- e.g. He was dead drunk.

It all happened so quick.

- 3. Adjectives as head N
- e.g. the English, the rich

The things went from bad to worse.

- 4. Adverbials
- 4.1 Adverbs Y
- e.g. very much

- 4.2 Prepositional adverbs Y
- e.g. I'll attend with my wife.
- 4.3 Finite clauses Y
- e.g. If I apply for the conference now, how much...
- 4.4 Non-finite clauses Y
- e.g. We expect those people, too, to attend.
- 4.5 Verbeless clauses N
- e.g. Peter was playing, unaware of the danger.
- 4.6 Noun phrases Y
- e.g. this time
- 4.7 Noun phrase followed by ago, long etc N
- e.g. Three years ago
- 4.8 Adverbial positions
- 5. Adverbial denoting manner, means, and instrument
- 5.1 Place Y
- e.g. at the Kyoto conference center
- 5.2 Time adverbials
- e.g. When a substistute is selected, I'll..
- 5.3 Time duration adverbials N
- e.g. for the summer
- 5.4 Time frequency adverbials N
- e.g. every day
- 5.5 Degree adverbials Y
- 5.6 Two or more adverbials N
- e.g.monthly this year
- 5.7 Sentence adverbials Y
- e.g. I don't understand Japanese at all.

 Could you possibly refund the registration fee?
- 6. Adverbs
- 6.1 Adverbs used as a modifier Y

- e.g. very much, about how much..., all right
- 6.2 Adverbs is a complement of a preposition N
- e.g. I don't know anybody around here.
- 7. Aposition N
- 7.1 Apposition N
- e.g. Fred Lang, a neighbour of you
- 7.2 Explicit apposition N
- e.g. For example, especially, in particular
- 8. Articles Y
- 9. Auxiliary verbs
- 9.1 Primary auxiliary verbs Y
- e.g. do, does, did, be, have, has (had)
- 9.2 Modal auxiliaries Y
- e.g. can, could, would, may, might, shall, should, will, must, ought to, need
- 10. Case
- 10.1 Noun and pronouns cases Y
- e.g. the conference office (common case) your name (genitive case)
- 10.2 Verb cases Y
- e.g. subject, object etc.
- 11. Basic verb patterns
- 11.1 SVC (be) Y
- 11.2 SVO (cancel)Y
- 11.3 SVOV (told) N
- 11.4 SVOO (teach) N
- 11.5 SVOC (consider) N
- 11.6 SV (Stay) Y
- 12. Voice
- 12.1 Active voice Y

- 12.2 Passive voice Y
- 12.2.1 Be-passives Y
- e.g. The proceedings and the reception fee are included in the attendance fee.
- 12.2.2 Get-passives N
- e.g. The boy got hurt on his way home from work.
- 13. Copula
- 13.1 be Y
- 13.2 others N
- 14. Clauses
- 14.1 Finit clauses Y
- e.g. This is the conference office.
- 14.2 Non-finit clauses
- 14.2.1 -ed participle clauses N
- e.g. The jobs finished, we went...
- 14.2.2 -ing participle clauses N
- e.g. drink up, go on, take off
- 41.2 Prepositional verbs Y
- e.g. apply for
- 42. Plurals Y
- 43. Postmodifiers
- 43.1 Relative clauses Y
- e.g. something which you don't know
- 43.2 Prepositional phrases Y
- e.g. thirty thousand yen per person
- 43.3 Non-finit clauses Y
- e.g. research related to Interpreting telephony
- 43.4 Appositive clauses N
- e.g. the fact she's good-looking is not the only reson why..
- 43.5 Clauses of time, place, manner, and reason Y

- e.g. In case a presentation is made in Japanese..
- 43.6 Adverbs N
- e.g. The way out is over there.
- 43.7 Adjectives N
- e.g. There is something odd about her.
- 44. Premodifiers
- 44.1 Adjectives Y
- e.g. a simultaneous translation
- 44.2 Participles Y
- e.g. a meeting place
- 44.3 Nouns Y
- e.q. the conference office
- 44.4 Compound premodifiers Y
- e.g. Osaka-shi
- 45. Prepositional phrases Y
- 45.1 Noun phrases
- e.g.Dinner is included in the attendance fee.
- 45.2 Wh-clauses Y
- e.g. From when..
- 45.3 -ING cluases N
- 45.4 Functions Of prepositional phrases
- 45.4.1 Adverbials Y
- e.g. Dinner is included in the attendance fee.
- 45.4.2 Postmodifiers in a noun phrase Y
- e.g. 10,000 yen per person
- 45.4.3 Verb complements Y
- e.g. The meeting place is in front of the conference center.
- 45.4.4 Complements of adjectives Y
- e.g. is closer to the...
- 45.4.5 Subjects, complements N

- e.g. Before breakfast is when I do my best work.
- 45.5 Prepositional adverbs N
- e.g. according to, due to, by means of, along with, except for, in comparison with, as for

ã

- 46. Pronouns Y
- 47. Proper nouns Y
- 47.1 Proper nouns with definit particle N
- e.g. The Netherlands
- 48. Quantifiers
- 48.1 Determiners Y
- e.g. some questions
- 48.2 Of-construction Y
- e.g. some of the speakers ..
- 48.3 Others N
- e.g. everybody, everyone
- 49. Questions
- 49.1 Yes-no questions Y
- e.g. Would you please take a look at it?
- 49.2 Wh-questions Y
- e.g. What sort of topic is this?
- 49.3 Tag questions N
- e.g. The boat has left, hasn't it?
- 50. Reciprocal pronouns N
- e.g. John and Mary like each other.
- 51. Relative clauses Y
- 51.1 Subjects Y
- e.g. The title of the papers which will be presented at the Conference
- 51.2 Object Y
- e.g. something which you don't understand..

- 51.3 Restrictives Y
- 51.4 Non-restrictives N
- e.g. Then he met Mary, who invited him to a party.
- 51.5 Senetence relatives N
- e.g. He admires Mrs Brown, which surprised me.
- 52. Sentences
- 52.1 Simple sentences Y
- e.g. That's right.
- 52.2 Complex sentences Y
- 52.2.1 Coordination Y
- 52.2.2 Subordination Y
- 53. Subjects Y
- 53.1 Subjects as Agent or experiencer Y
- e.g. I have heard that..
- 53.1 Subject as Object Y
- e.g. This is the conference office.
- 54. Subjunctives N
- e.g. It is necessary that every member should inform himself of these rules.

If she were to do something like that..

- 55. Signals of surbordination Y
- 55.1 Signals Y
- e.g. that, if, where
- 55.2 Subordinating conjunctions
- 55.2.1 Simple subordinating conjunctions Y
- e.g. if..
- 55.2.2 Compound subordinating conjunctions N
- e.g. so that, in order that, such that
- 55.2.3 Ending with as Y
- e.g. as far as that goes

- 55.2.4 Ending with than N
- e.g. rather than sooner than
- 55.3 Corelative surbordinative conjunctions N
- e.g. if...then, more.....than, such...as,
- 56. Subclause function
- 56.1 Subject N
- e.g. That we badly need new equipment is obvious.
- 56.2 Direct object Y
- e.g. Tell me how I should go..
- 56.3 Indirect object N
- e.g. I gave whoever it was a drink.
- 56.4 Subject complement N
- e.g. the point is that they're leaving right away.
- 56.5 Object complement N
- e.g. I can't imagine John overcome with grief.
- 56.6 Adverbial Y
- e.g. As far as that goes,...
- 57. Verb patterns
- 57.1 linking verbs
- 57.1.1 Copula Y
- e.g. be
- 57.1.2 Resulting linking verbs Y
- e.g. go
- 57.2 Verbs with object Y
- 57.2.1 Object is a noun Y
- e.g. Could you suggest some place?
- 57.2.2 Object is infinitive N
- e.g. Can I help clean the window?
- 57.2.3 Object is to-infinitive Y
- e.g. I would like to inf

- 57.2.4 Object is -ing form N
- e.g. We all enjoyed seeing them again.
- 57.2.5 Object is that-clause Y
- e.g. I've heard that there is a city tour ..
- 57.2.6 Object is 'so' or not N
- e.g. I think so.
- 57.2.7 Object is that-clause with putative should N
- e.g. I request that she should go alone.
- 57.2.8 Complement is Wh-clauses Y
- e.g. I'd like you to tell me how I should go to..
- 57.2.9 Object is non-finit verbs N
- e.g. I don't know how to address this letter.
- 57.2.10 Complement is infinitive Y
- e.g. Let me know...
- 57.2.11 Complement is to-infinitive N
- e.g. He allowed the neighbours to use his car.
- 57.2.12 Complement is -ing N
- e.g. I don't like him being rude to you.
- 57.2.13 2 objects are nouns N
- e.g. He gave George the money.
- 57.2.14 Object + that clause Y
- e.g. tell me that...
- 57.2.15 Object is so or not N
- e.g. I told you so.
- 57.2.16 Object + finit wh-clause Y
- e.g. tell me how I should go..
- 57.2.17 object + Complement N
- e.g. The parents named the baby Susan.
- 57.3 Verbs without object and complement Y
- e.g. From when you will stay?

- 58. Verb combinations
- 58.1 with modal Y
- e.g. you can't make refunds.
- 58.2 with perfectives Y
- e.g. I've heard..
- 58.3 with progressive N
- e.g. we'll be expecting.
- 58.4 with passive N
- e.g. Several letters were being typed by him.
- 59. Tense and aspect
- 59.1 Tense (Present, past, future) Y
- e.g. we also expect people who specialize ...
- ·59.2 Aspect(perfect, Non-perfect, progressive) Y
- e.g. You have already deposited 85,000 yen..

ANNEX 2

Meanings in situation (at random)

Symbol ⊲ means meanings lie in following grammatical items.

- 2.1 Definite and indefinite meanings ⊲ articles
- 2.2 Restractive and non-destractive meanings ⊲ modifiers
- 2.3 Time ⊲ prepositions, nouns
- 2.4 Tense and Aspect ⊲ auxiliaries
- 2.6 Frequency ⊲ noun phrases, prepositions
- 2.7 Place, direction and distance ⊲ adverbs, prepositions, nouns
- 2.8 Manner, Means and instrument ⊲ prepositional phrases, adverbs
- 2.10 Condition and contrast ⊲ conjunctions, adverbs
- 2.11 Degree ⊲prepositional phrases, adverbs
- 2.12 Role, at and and viewpoint ⊲ adverbs, prepositions, nouns
- 2.14 Addition, exception and restriction ⊲ adverbs, conjunctions
- 2.15 Denial and affirmation ⊲ adverbs
- 2.16 Agreeement and disagreements adverbs
- 2.17 Fact, hypothesis and neutrality ⊲verbs
- 2.19 Permission and obligation ⊲ modal auxiliaries
- 2.21 Honorifics ⊲ modal auxiliaries

- 2.22 Spontaneity ⊲ verbs
- 2.23 Ability ⊲ modal auxiliaries
- 2.24 Hearsay ⊲ modal auxiliaries
- 2.25 Volition ⊲ modal auxiliaries
- 2.26 Performative ⊲ verbs
- 2.27 Invitation ⊲ modal auxiliaries
- 2.28 Reason ⊲ conjunctions, prepositional phrases
- 2.29 Recommandation ⊲ modal auxiliaries
- 2.30 Explanatory ⊲ verbs
- 2.31 Experience and habit ⊲ verbs
- 2.32 Understanding ⊲ phrases
- 2.33 Perlocutionary ⊲ verbs
- 2.34 Topics and scope ⊲ phrases
- 2.35 Polarity ⊲ adverbs, antonym word

References

- [1] Toshiro Hasegawa Details of the transfer process in the AŞURA translation system 1990 TR-I-0188 ATR
- [2] Toshiro Hasegawa Feature Structure Rewriting Manual 1990 TR-I-0187 ATR
- [3] Geoffrey Leech and Jan Svartvik A Communicative Grammar of English 1975 Longman
- [4] Mutsuko Tomokiyo and Tsuyoshi Morimoto Communicative Functions of Spoken Japanese and its Meaning Interpretation on MT System 1992 TR-I-0260 ATR
- [5] Petr Sigall Eva Hajičová and Jarmila Panevova The meaning of the Sentence in its Semantic and Pragmatic Aspects 1926 D.Reidel Publishing Company
- [6] Ikuo Kudo and Mutsuko Tomokiyo Ellipsis resolution Mechanism by using Japanese Predicate Particularity 1992 paper accepted by Densijyouhoutuusingakkai
- [7] Searle.J.R and M.Bierwisch Speech act theory and pragmatics 1980 Dordrecht
- [8] Laurel Fais and Gen-ichiro Kikui Determining Surface Form for Indirect Speech Acts in English 1991 TR-I-0235 ATR
- [9] Minoru Yasui and al. Imiron Taisyuukansyoten 1983