

Illocutionary Act Based Translation of Dialogues

Kiyoshi KOGURE*

NTT Basic Research Laboratories
9-11, Midori-cho 3-chome, Musashino-shi, Tokyo, 180 Japan

Masako KUME and Hitoshi IIDA

ATR Interpreting Telephony Research Laboratories
Sanpeidani, Inuidani, Seika-cho, Soraku-gun, Kyoto, 619-02 Japan

Abstract

A new dialogue translation method called the illocutionary act translation method is designed for translating speakers' intentions conveyed by rich expressions in dialogues between humans. One of the most significant characteristics of this method is that it is pragmatics-based. That is, translation of utterances representing illocutionary acts, which consist of illocutionary forces and propositional contents, is possible while most conventional MT methods are based on syntactic or semantic information and concerned with propositional contents only. The new method extracts illocutionary act relationships from input utterances by using semantic-pragmatic representations called illocutionary act types. These types consist of language-independent illocutionary act concepts and language-dependent propositional content concepts. The method then transfers these into target language illocutionary act types, and finally generates target language utterances. This method combines the features of the transfer approach and the interlingual approach.

1 Introduction

In a dialogue, a speaker expresses his/her communicative intentions (or performs communicative acts) and a hearer tries to understand them by exploiting (1) shared beliefs regarding communicative act achievement modes held in the language community to which the dialogue participants belong, and (2) shared beliefs regarding the dialogue situation including social relationships between participants and discourse relationships such as topics and foci. Based on these beliefs, a wide variety of communicative acts such as stating, questioning, requesting and promising are performed. Moreover, for a single communicative act, various linguistic devices can be used. For example, in requesting, the linguistic devices used depend on, for example, the speaker's attitude towards the social relationship with the hearer.

To translate dialogues, an MT system must at least translate communicative acts correctly. However, this is difficult for conventional MT approaches. Conventional approaches are mostly designed for treating written texts such as technical documents, which mainly contain assertive sentences. Almost all of these approaches treat syntactic information or semantic information related only to the propositional contents of sentences. However, dialogue utterances involve various kinds of communicative acts. Thus, the following properties of dialogue utterances cause problems. First, very different kinds of linguistic devices are used to represent communicative intentions in different languages. Second, linguistic devices with separate locations are jointly used to represent a communicative intention. These make it difficult to make correspondences between syntactic objects. As a result, conventional approaches are inadequate. An alternative approach is required which is

* Most of the research reported in this paper was done while the author was at ATR Interpreting Telephony Research Laboratories.

designed for treating communicative intentions directly, i.e., an approach that extracts communicative intentions from the source language and generates expressions for these intentions in the target language.

This paper proposes a new dialogue translation method, called the illocutionary act translation method (IATrans method for short), which was developed in line with the intention translation approach[5][7][2]. One of the most significant characteristics of the IATrans method is that it translates the illocutionary acts in the speaker's utterances. In general, an illocutionary act consists of an illocutionary force and a propositional content[9]. The IATrans method extracts illocutionary acts consisting of language-dependent illocutionary force concepts and source language propositional content concepts. The method then transfers them to illocutionary acts including target language propositional concepts, and finally generates the target language utterances.

Section 2 first discusses the kinds of semantic or pragmatic objects treated in dialogue translation. The section then describes how to represent these objects. The framework of the illocutionary act translation method is also presented. Sections 3 and 4 explain the analysis, transfer and generation processes of the method.

2 Overview of Illocutionary Act Based Translation

2.1 Translation of Illocutionary Acts

The minimal units of human communication can be taken as speech acts of a type called illocutionary acts such as stating, questioning, requesting, promising, apologizing, and so on[9]. Dialogue translation at least requires treatment of illocutionary acts. For example, the English sentence

(1) *Can you pass the salt?*

is literally translated into the Japanese sentence¹

(2) Shio o tewatase masu ka?
Salt ACC pass-can POL QUEST
Can you pass the salt?

but this translation is often inadequate, because uttering the English sentence can perform two different illocutionary acts, that is, the speaker's question of the hearer's ability to pass the salt, and the speaker's request that the hearer pass the salt. In other words, from the hearer's point of view, a direct interpretation of the uttered English sentence can be the question act and an indirect interpretation can be the request act. However, uttering the Japanese sentence can perform only the question act. When the speaker performs the request in uttering the English sentence (a frequent case), the Japanese translation is inadequate. An adequate Japanese sentence for performing the request is, for example the following expression which includes a benefactive predicate

(3) Shio o tewatashite itadake masu ka?
Salt ACC pass BENE-can POL QUEST
Can I receive your favor of passing the salt ?

For keeping a speaker's intention in dialogue translation, indirect interpretations of illocutionary acts must be treated. This paper, therefore, proposes an illocutionary-act-based translation of dialogue utterances. This paper focuses on conventional illocutionary acts for the reason described below.

The degree of indirectness in illocutionary acts varies with the situation. Some illocutionary acts are performed rather indirectly but others are performed very indirectly. For example, consider the following sentences.

(4) Gakusei waribiki o shite kudasai.
Student discount ACC do REQUEST
Please give me a student discount.

(5) Gakusei waribiki o shite itadake masu ka.
Student discount ACC do BENE-can POL QUEST
Could you give me a student discount.

¹ ACCusative case marker, BENEFicative subsidiary verb, POLite auxiliary verb, QUESTion particle, COMPLEMENTizer, MODeration particle, COPULA verb

| | |
|------------|--|
| relation | <i>(illocutionary force relationship name)</i> |
| agent | <i>(speaker)</i> |
| recipient | <i>(hearer)</i> |
| object | <i>(propositional content type)</i> |
| manner | <i>(achievement mode type)</i> |
| politeness | <i>(set of honorific types)</i> |

Figure 1: Definition of the illocutionary act type

- (6) Gakusei na no desu ga.
 Student COPULA COMP COPULA MOD
I'm a student, you know.

In these examples, each sentence can, in certain adequate situations, have the interpretation that the speaker requests the hearer to give him/her a student discount. The differences are in degree of indirectness. Utterance of sentence (4) has the direct interpretation of a request, while (5) and (6) have indirect interpretation of the request. Sentence (5) is interpreted as the request less indirectly than (6). This difference in indirectness is related to convention. Sentence (5) is interpreted conventionally as a request while sentence (6) is not. A new question now arises. Is it sufficient to treat only conventional indirect interpretations?

Unconventional interpretations are obtained from conventional interpretations and plan-based inferences referring to the utterance situation. In this inference process, linguistic conventions do not play any significant role. This process proposes additional interpretations which can not be derived from linguistic conventions. Thus, this process is language-independent. An unconventional interpretation can be obtained from a conventional interpretation by the hearer even when he does not know the linguistic convention in the source language. Therefore, treating conventional direct and indirect interpretations is sufficient². Moreover, if unconventional interpretations are assumed and generation from unconventional interpretations is permitted, too wide a range of utterances would be generated, most of which would be inappropriate for the utterance situation.

2.2 Illocutionary Act Types

The IATrans method extracts from an input utterance the type of illocutionary act the speaker performed, as a semantic-pragmatic representation called the illocutionary act type. An illocutionary act type represents an illocutionary act relationship. The type consists of an illocutionary force relationship name, and the agent (i.e., the speaker), recipient (i.e., the hearer), object, manner, and politeness roles. In this paper, such types are represented by feature structures (as shown in Figure 1).

The illocutionary force relationship name is a subtype of ASSERTIVE, COMMISSIVE, DIRECTIVE, ROGATIVE-IF, ROGATIVE-REF, DECLARATIVE, or EXPRESSIVE.

The object role value is a propositional content type representing a propositional content relationship. It is represented by source or target language relationship names. The the manner role value is an achievement mode type for the illocutionary act relationship. It consists of an achievement mode relationship name, and optional roles to represent subtle information. The achievement mode relationship name is a subtype of DIRECTLY or INDIRECTLY, e.g., INDIRECTLY-BY-ASKING-HR-ABILITY, INDIRECTLY-BY-REPORTING-SP-DESIRE³. The politeness role value is a set of honorific relationship types between the speaker, the hearer, and the objects referred to in the utterance; each relationship is named RESPECT, CONDESCEND, or POLITE.

An illocutionary act type, excluding its object role, is called an illocutionary force type. The illocutionary force type is expressed language-independently; only the propositional content type is expressed language-dependently.

2.3 The Framework of the Illocutionary Act Translation Method

² This does not mean unconventional interpretations are not useful. Analyzing unconventional interpretations are useful in examining dialogue consistency and in resolving ambiguities of conventional interpretations.

³ SPeaker, HeaRer

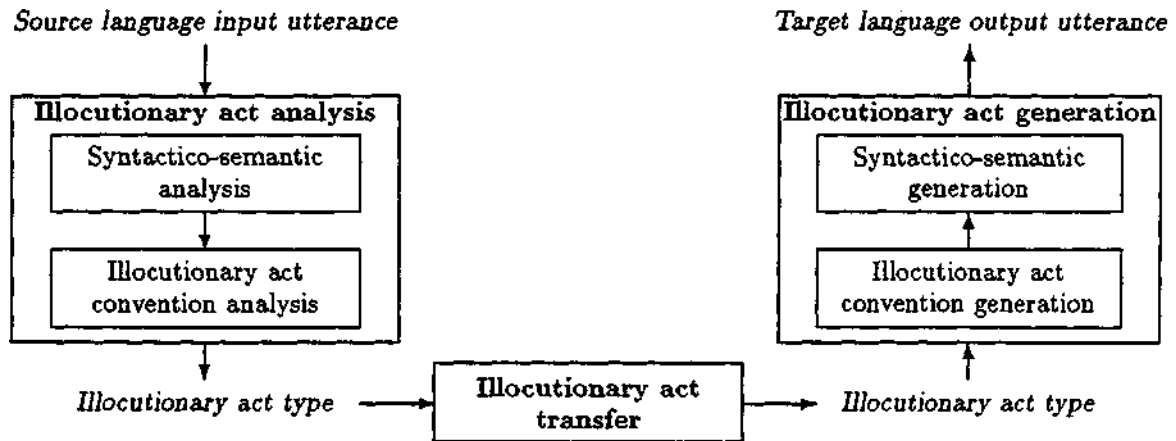


Figure 2: An overview of the illocutionary act translation method

The IATrans method is essentially similar to methods taking the semantic transfer approach, and consists of the analysis, transfer, and generation processes. The analysis process extracts from an input utterance an illocutionary act type whose propositional content type is represented in terms of source language relationship names. The transfer process then converts the illocutionary act type into an illocutionary act type with the same illocutionary force type and corresponding propositional content type in terms of the target language relationship names. Finally, the generation process generates a target language expression. The IATrans method is neither syntax-based nor semantics-based, but pragmatics-based.

In a sense, this method can be considered to be intermediate between the transfer method, the most popular method, and the interlingual method, which has difficulty in defining language-independent interlingual relationship names, but is efficient at developing translation systems among more than two languages. This transfer process is simpler than conventional transfer process because this method's transfer process converts only propositional contents. It also reduces the difficulty of defining interlingual relationship names because only relationship names for illocutionary force types need to be defined language-independently, and because there are much fewer of them than interlingual language-independent relationship names.

The following section will illustrate the translation of Japanese dialogue utterances into English utterances based on the IATrans method.

3 Illocutionary Act Analysis

The IATrans method treats conventional interpretations of illocutionary acts including indirect interpretations. This method requires analysis of direct interpretations and of indirect interpretations obtained from direct ones by using linguistic conventions.

A two-stage analysis approach is, therefore, adopted. It consists of a syntactico-semantic analysis process and an illocutionary act convention interpretation process. The first process takes a spoken-style Japanese sentence as its input and extracts a direct or literal illocutionary act type which represents the illocutionary act performed in the utterance of the input sentence. The second process maps a direct illocutionary act type into an illocutionary act type which represents the primary illocutionary act by using illocutionary act convention rules. The output of the first process, i.e., a direct illocutionary act type, consists of relationships in terms of language-dependent relationship names and thus permits the second process to access language-dependent information.

3.1 Syntactico-semantic Analysis of Input Utterances

In the syntactico-semantic analysis process, a unification-based syntactico-semantic approach is adopted so as to compositionally construct direct illocutionary act types from input utterances.

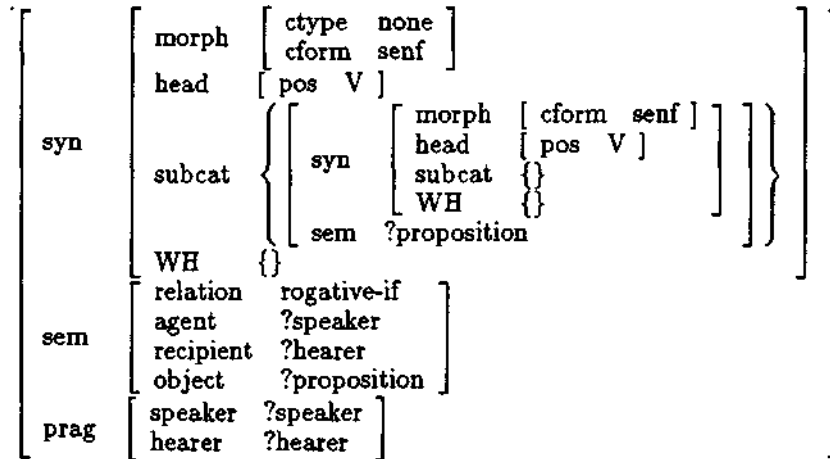


Figure 3: Lexical description of the sentence final particle "ka"

In the above matrix notation of feature structures, "?" is the prefix of the tag and structures with the same tag are token identical. A feature value with a tag but no value specification can be unified with any feature structure. "{" and "}" enclose feature structure lists; "{}" denotes an empty list.

The process takes a spoken-style Japanese sentence as its input and constructs phrase structures, including syntactic, semantic and pragmatic information, represented by feature structures.

The syntactico-semantic analysis process uses a unification-based lexico-syntactic grammar for treating spoken-style Japanese sentences. The grammar is essentially based on HPSG[8] and uses features from JPSG[1]. The reason for this is listed below.

1. A (feature-structure-) unification-based approach permits integrated descriptions of information from various sources such as syntax, semantics, and pragmatics. That is, constraints among them can be described in terms of typed feature structures. Thus, a simultaneous analysis of syntax, semantics, and pragmatics is allowed.
2. The unification-based approach permits partial descriptions.
3. A lexico-syntactic approach is modular in the sense that most of the grammatical information is specified in descriptions of lexical items. It is therefore easy to extend a grammar by adding new lexical items to the lexicon and by adding new information to the lexical items.
4. The HPSG and JPSG frameworks can capture constraints between complex predicate constituents and their complements. These constituents are very important in expressing illocutionary forces in spoken-style Japanese.

Spoken-style Japanese sentences often have complex sentence-final predicate phrases consisting of main predicates (verbs, adjectives, and nominal adjectives), combinations of auxiliary predicates, and sentence-final particles. Auxiliary predicates such as the causative, "seru", benefactive, "morau", "itadaku" etc., and sentence-final particles such as question particle, "ka", and moderation particle, "ga" are very important in expressing illocutionary forces.

In such a predicate phrase, the head constituent stipulates the attributes such as part of speech, conjugational type and form, and unfilled complements of its complement occurring immediately to its left. Such stipulations are easily described in the SUBCAT feature value in the head. A SUBCAT feature value is a list of complement constituent specifications.

For example, Figure 3 shows the lexical description of the sentence-final particle "ka" which indicates a question. The SUBCAT feature value specifies that the particle takes as its complement a verb phrase with CFORM (conjugational form) feature value SENF (sentence final form). The complement's WE feature value and SUB CAT feature value are empty lists. The SEM feature value expresses that a sentence with this particle as its sentence-final constituent can be used to perform an ROGATIVE-IF illocutionary act when its complement satisfies the above specifications. The

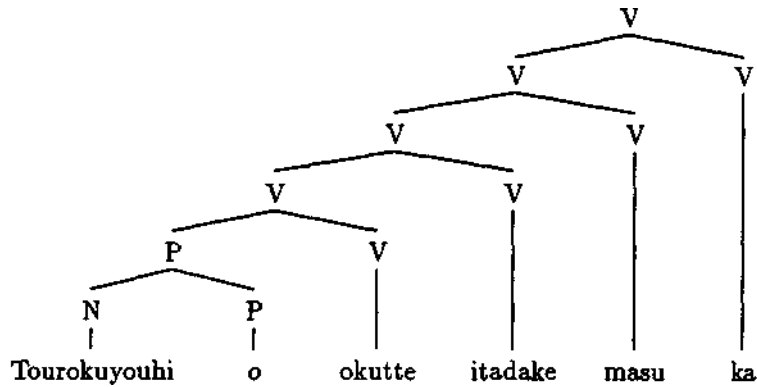


Figure 4: Derivation tree of sentence (7)

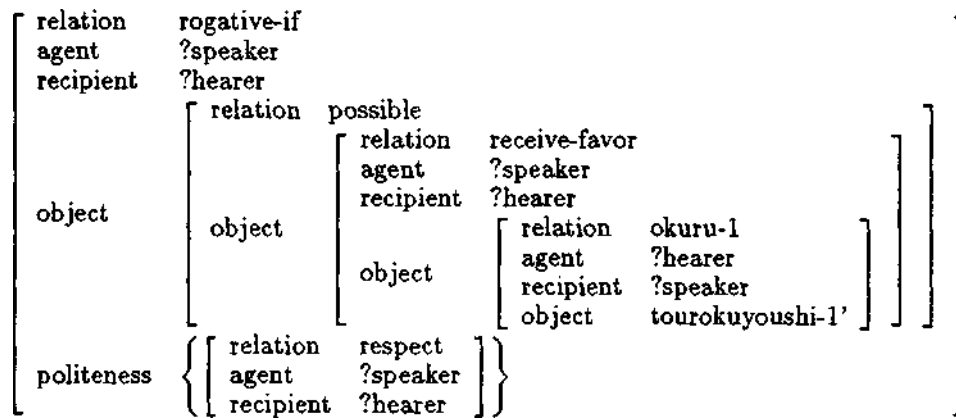


Figure 5: The direct illocutionary act of sentence (7)

agent of the act is the speaker and the recipient is the hearer. The propositional content (the object role value) is specified by the complement's *SEM* feature value.

This analysis method outputs syntactic, semantic and pragmatic information in terms of feature structures. An utterance's direct illocutionary act type is represented by the *SEM* feature value. For example, the sentence

- (7) Tourokuyoushi o okutte itadake masu ka.
 registration form ACC send BENE-can POLITE QUEST
Could you send me a registration form?

is analyzed in Figure 4. From this analysis, the direct illocutionary act type is shown in Figure 5. The type represents an illocutionary act by which the speaker asks the hearer whether the speaker will be able to receive favor of hearer's sending a registration form or not. This process uses a unification-based parser (e.g., [4]).

3.2 Illocutionary Act Convention Analysis

The illocutionary act convention analysis process takes as its input a direct illocutionary act type and gives as its output a set of primary illocutionary act types. This process uses illocutionary act convention rules to describe relationships between direct and primary illocutionary act types. The process attempts to unify an input type with direct illocutionary act types from rules. If a unification succeeds, the primary illocutionary act type determined by the rule, with additional

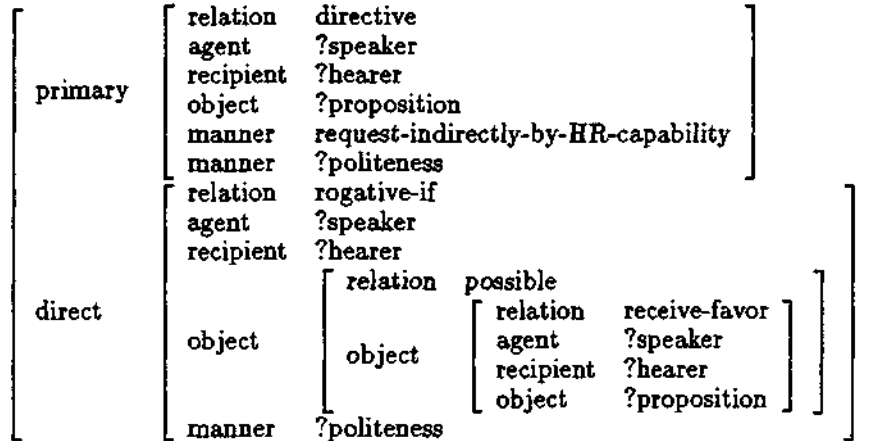


Figure 6: An example of illocutionary act convention rule

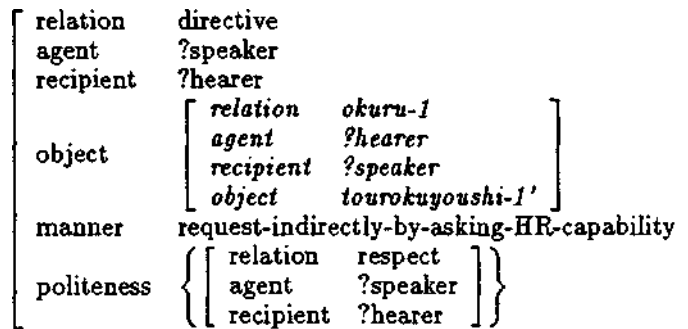


Figure 7: A primary illocutionary act of sentence (7)

information from the unification, is obtained as the primary illocutionary act type corresponding to the input type.

For example, the type in Figure 7 is obtained from the input type (shown in Figure 5) by using the rule shown in Figure 6. The rule represents the linguistic convention that asking the possibility of speaker's receiving favor of a hearer's act means requesting the hearer to perform the act.

This process analyzes linguistic convention for performing illocutionary acts as described above. The process is achieved by using a typed feature structure rewriting system (e.g., [11][7][6]). It often results in several primary act type candidates. These candidates can be filtered by a dialogue plan recognition system (e.g., [3]).

4 Illocutionary Act Transfer and Generation

The illocutionary act transfer process of the IATrans method converts propositional content types in terms of source language relationship names into corresponding propositional content types in terms of target language relationship names. The illocutionary act type in Figure 7 is converted into the type shown in Figure 8. Only the propositional content type (i.e., the OBJECT feature value of the type in *italic*) is modified.

The generation process generates target language expressions. The generation process consists of two stages similar to the analysis process. The first stage applies illocutionary act inference rules in reverse direction to that of the analysis process. Attempts are made to unify the input type with primary illocutionary act types from the rules; if a unification succeeds, then the direct illocutionary act type determined by the rule, with additional information from the unification, is obtained as the

| | | | | | | | | | |
|---|--|-----------------|----------------|--------------|-----------------|------------------|-----------------|---------------|-----------------------------|
| relation agent recipient | directive ?speaker ?hearer | | | | | | | | |
| object | <table style="border-collapse: collapse;"> <tr> <td style="border-right: 1px solid black; padding: 2px 5px;">relation</td> <td style="padding: 2px 5px;"><i>send-1</i></td> </tr> <tr> <td style="border-right: 1px solid black; padding: 2px 5px;">agent</td> <td style="padding: 2px 5px;"><i>?hearer</i></td> </tr> <tr> <td style="border-right: 1px solid black; padding: 2px 5px;">recipient</td> <td style="padding: 2px 5px;"><i>?speaker</i></td> </tr> <tr> <td style="border-right: 1px solid black; padding: 2px 5px;">object</td> <td style="padding: 2px 5px;"><i>registration-form-1'</i></td> </tr> </table> | relation | <i>send-1</i> | agent | <i>?hearer</i> | recipient | <i>?speaker</i> | object | <i>registration-form-1'</i> |
| relation | <i>send-1</i> | | | | | | | | |
| agent | <i>?hearer</i> | | | | | | | | |
| recipient | <i>?speaker</i> | | | | | | | | |
| object | <i>registration-form-1'</i> | | | | | | | | |
| manner | request-indirectly-by-asking-HR-capability | | | | | | | | |
| politeness | <table style="border-collapse: collapse;"> <tr> <td style="border-right: 1px solid black; padding: 2px 5px;">relation</td> <td style="padding: 2px 5px;">respect</td> </tr> <tr> <td style="border-right: 1px solid black; padding: 2px 5px;">agent</td> <td style="padding: 2px 5px;">?speaker</td> </tr> <tr> <td style="border-right: 1px solid black; padding: 2px 5px;">recipient</td> <td style="padding: 2px 5px;">?hearer</td> </tr> </table> | relation | respect | agent | ?speaker | recipient | ?hearer | | |
| relation | respect | | | | | | | | |
| agent | ?speaker | | | | | | | | |
| recipient | ?hearer | | | | | | | | |

Figure 8: The transferred illocutionary act of sentence (7)

direct illocutionary act type. The output type is applied to the lexico-syntactic generation process (e.g., [10]) and a target language sentence is obtained. For example, from the previous illocutionary act type, the following sentence is obtained:

(8) *Could you please send me a registration form?*

This sentence can be used to perform in the target language the illocutionary act that the speaker performs in the source language. As illustrated in the above example, the utterance translated by the IATrans method can be used to perform the speaker's illocutionary act.

5 Conclusion

In this paper, a new dialogue translation method called the illocutionary act translation method was proposed. This method is designed for translating speakers' communicative intentions. It is thus neither syntax-based nor semantic-based, but pragmatics-based. The method extracts source language illocutionary act types which represent what is intended to be achieved in input source language utterances, transfers them into target language illocutionary act types, and finally generates target language utterances. This method consists of three processes, like the transfer approach, and uses language-independent illocutionary force relationship names, like the interlingual approach. Therefore, this method can be thought of as an intermediary among these approaches, and combines their advantages.

Acknowledgement

The authors would like to thank Akira Kurematsu and other colleagues at ATR Interpreting Telephony Research Laboratories for their encouragement and thought-provoking discussions. The authors would like to thank Akira Shimazu for his comments.

References

- [1] T. Gunji. *Japanese Phrase Structure Grammar*. D. Reidel, 1987.
- [2] H. Iida, K. Kogure, K. Yoshimoto, and T. Aizawa. An experimental spoken natural dialogue translation system using a lexicon-driven grammar. In *The Proceedings of the European Conference on Speech Communication and Technology*, 1989.
- [3] H. Iida, T. Yamaoka, and H. Arita. Three typed pragmatics for dialogue structure analysis. In *The Proceedings of COLING-90*, Helsinki, Finland, 1990, *to be published*.
- [4] K. Kogure. Parsing Japanese spoken sentences based on HPSG. In *The Proceedings of the International Workshop on Parsing Technologies*, pages 132-141, Carnegie Mellon University, Pittsburgh, PA, 1989.
- [5] K. Kogure, H. Iida, K. Yoshimoto, H. Maeda, M. Kume, and S. Kato. A method of analyzing Japanese speech act types. In *The Proceedings of the 2nd International Conference on Theoretical and Methodological Issues in Machine Translation of Natural Languages*, Carnegie Mellon University, Pittsburgh, PA, 1988.

- [6] M. Kume and K. Kogure. Inference mechanism of illocutionary force types for MT systems (in Japanese). In *The Proceedings of the IPSJ Fall Meeting*, 1989.
- [7] M. Kume, G. Sato, and K. Yoshimoto. A descriptive framework for translating speaker's meaning. In *The Proceedings of the 4th Conference of the European Chapter of the Association for Computational Linguistics*, 1989.
- [8] C. Pollard. *An Information-Based Syntax and Semantics — Volume 1: Fundamentals. CSLI Lecture Note Number 13*, CSLI, 1987.
- [9] J. R. Searle and D. Vanderveken. *Foundations of Illocutionary Logic*. Cambridge University Press, 1985.
- [10] Y. Ueda and K. Kogure. Generation for dialogue translation using typed feature structure unification. In *The Proceedings of COLING-90*, Helsinki, Finland, 1990, *to be published*.
- [11] R. Zajac. A transfer model using a TFS rewriting system with inheritance. In *The Proceedings of the 27th Annual Meeting of the Association for Computational Linguistics*, The Association for Computational Linguistics, 1989.