# Machine Translation of English Dialogue into Korean on the Basis of Contextual Information*

## Dong-Young Lee

The purpose of this research lies in revealing that contextual information such as information about dialogue participants and social status information must be used in a machine translation of English dialogue into Korean. Unlike a single sentence, a dialogue is the conversation between dialogue participants. Although there is no indication of honorification in English dialogue, such indication always appears in Korean dialogue. This means that depending on the relative order of social status among the people involved in English dialogue, translated Korean dialogue must vary.

## 1. Introduction

There are a few studies (Choi et al. 1994, Weinstein et al. 1996) on how to translate English text into Korean using the computer. None of them, however, dealt with a machine translation of written English dialogue into Korean.

This paper shows that an appropriate translation of English dialogue into Korean must make use of the situation in which English dialogue is held. For example, the result of translating an English dialogue that is held between a father and his son is different from that of translating an English dialogue that is held between a father and his friend. The reason is that although English dialogue has the same form regardless of who are the speaker and the addressee of utterances occurring in dialogue, the form of Korean dialogue varies greatly depending on the relative order of social status for the people such as the speaker, the addressee, and the individuals mentioned in dialogue. The system for translating English dialogue into

Korean is developed on a Sun SPARCstation using the Prolog programming language.

The layout of this paper is as follows: In Section 2 we present the difference in linguistic phenomenon between English dialogue and Korean dialogue. Section 3 illustrates the reason contextual information is needed in the translation of English dialogue into Korean. The architecture of our dialogue translation system and a method to implement the system are discussed in detail in Section 4. In the final section we discuss the appropriateness and usefulness of our translation system and the application of the system to other areas.

## 2. Differences between English Dialogue and Korean Dialogue

Since we are concerned with how to translate English dialogue into Korean dialogue using a computer, let us consider differences between dialogues of these two languages in the aspect of linguistic phenomena.

### 2.1. English Dialogue

In English utterance the speaker may use the word *sir* to show 'honor' to the addressee when the social status of the latter is higher than that of the former. Let us consider the utterance appearing in (1).

(1) I did it, sir.
    (Speaker: a military officer, Addressee: another military officer)

Since the utterance in (1) is used in a military community, it is possible to infer that the social status of the addressee is higher than that of the speaker.

In everyday life, however, the word *sir* is frequently used to show 'politeness' to the addressee. For example, let us look at the utterance shown in (2), which is used in restaurants.

(2) Are you ready to order, sir?
    (Speaker: a waiter, Addressee: a customer)

From the utterance in (2), it is not possible to infer that the social status

of the addressee is higher than that of the speaker. Thus from the use of the word *sir* in an utterance we cannot always infer the relative order of social status between the speaker and the addressee.

In general, however, from an English utterance we cannot get information about the relative order of social status among the people involved in the utterance. Let us look at the utterance shown in (3).

(3) Did you meet him yesterday?
   (Speaker: John, Addressee: Paul)

From the utterance in (3) we cannot obtain information about the order of social status among the three people: John, Paul, and the referent of the pronoun *him*.

Likewise, English dialogue, which consists of utterances, does not provide any information about the order of social status among the people involved. Let us consider the dialogue illustrated in (4).

(4) a. I gave Mr. Smith a brochure last Tuesday.
      (Speaker: Harold, Addressee: Susan)
    b. Is he interested in multimedia?
      (Speaker: Susan, Addressee: Harold)
    c. Yes, he is. He has learned it from Mr. Hazlitt.
      (Speaker: Harold, Addressee: Susan)

In dialogue (4) conversation is held between Harold and Susan. In their conversation two persons (that is, Mr. Smith and Mr. Hazlitt) are men-tioned. Thus four persons are involved in dialogue (4). From the dialogue, however, we cannot get any information about the relative order of social status among those four persons.

Although the use of the word *sir* may give a clue to the order of social status between the speaker and the addressee of an utterance occurring in English dialogue, the clue is not always reliable. No other words or morphemes used in English dialogue convey information about social status. Therefore, from English dialogue we cannot infer any definite order of social status among the people involved in dialogue.

## 2.2. Korean Dialogue

In contrast to English dialogue, honorification phenomenon is related to

Korean dialogue. The phenomenon of honorification occurs when the speaker of an utterance shows honor to a person (this person may be the addressee or the one who is mentioned in the utterance).[1] Honorification is linguistically realized in Korean dialogue through the use of some specific morphemes such as honorific case markers, an honorific suffix, an honorific verbal ending, and humble verb forms.

### 2.2.1. Morphemes Related to Honorification

As discussed in Lee (1996a), there are several morphemes that are related to honorification in Korean.

First, the morpheme *nim* is attached to an NP in order to show that the speaker honors the referent of that NP.

Second, honorific case markers as well as nonhonorific case markers exist. Unlike English, a grammatical function of an NP in Korean is determined by the case marker that follows the NP. The system of case markers is as illustrated in (5).

(5) Case Markers:

|             | nominative | genitive | dative | accusative |
|-------------|------------|----------|--------|------------|
| nonhonorific | ka, i     | uy       | eykey  | ul, lul    |
| honorific    | kkeyse    | –        | kkey   | –          |

As illustrated in (5) honorific case markers exist only for nominative case and dative case. An expression of honor to the referent of genitive NP or accusative NP is indicated by the honorific morpheme *nim* which is inserted before the case marker.[2]

Third, the honorific morpheme *si* is used after the root of a verb when the speaker shows honor to the person who is referred by the subject NP.

Fourth, if the speaker honors the addressee, honorific verbal endings are used in an utterance. The system of honorific verbal endings and nonhonorific verbal endings is as illustrated in (6).

---

[1] For honorification phenomenon in Korean, refer to the following researches: Han 1991, Kim 1988, Kuno and Kim 1985, Park 1991, and Suh 1978.

[2] In the case of a nominative NP or a dative NP the honorific morpheme *nim* can be followed by an honorific case marker.

(6) a. Declarative Verbal Ending:

|  | formal | informal |
|---|---|---|
| nonhonorific | ta | e, a |
| honorific | (su)pnita | (e,a,ye)yo |

b. Interrogative Verbal Ending:

|  | formal | informal |
|---|---|---|
| nonhonorific | (nu)nka | ni, e, a |
| honorific | (su)pnikka | (e,a,ye)yo |

If the situation in which conversation occurs is formal (for example, the situation of business negotiation), formal verbal endings are used in dialogue. On the other hand, if a situation is informal (for instance, the situation of conversation between family members), informal verbal endings are used. For example, the use of the verbal ending *supnikka* in an utterance means that the speaker of the utterance honors the addressee and conversation is held in a formal situation.

Finally, when the speaker shows honor to the referent of an object NP and a separate humble form of a verb is available, that form of the verb is used in an utterance (for example, the verb *tuli* 'give' is the humble form of the verb *cwu*).

### 2.2.2. Occurrence of Honorification

According to the kinds of honorific morphemes that are used in an utterance, the person who is honored by the speaker can be recognized.

First, let us look at an utterance in which a subject referent is honored.

(7) Robinson-ssi-kkeyse      kicha-lul      nohchi-si-ess-e.
         Mr.-nom (hon)      train-acc      miss-hon-past-dec
     'Mr. Robinson missed a train.'
     (Speaker: Ralph, Addressee: Peter)

In utterance (7) the honorific nominative case marker *kkeyse* and the honorific morpheme *si* are used. The use of these two morphemes indicates that the person Mr. Robinson is honored by the speaker Ralph in the utterance.

Second, by using the honorific morpheme *nim*, an honorific case marker, or a humble form of a verb, the speaker can show honor to an object referent.

(8) David-ka      Malone-pwusacang-nim-ul      poy-ess-e.
         nom          vice-president-hon-acc    meet-past-dec
    'David met vice president Malone.'
    (Speaker: Albert, Addressee: Larry)

Since honorific accusative case marker does not exist in Korean, the honorific morpheme *nim* is inserted between an object NP and a case marker, as shown in utterance (8). In addition, the verb *poy* is the humble form of a verb *manna* 'meet' and thus that form is used in (8).

Third, the speaker can show honor to the addressee by using an honorific verbal ending.

(9) Michael-i    George-eykey    phyenci-lul    pwuchi-ess-eyo. [3]
         nom            dat       letter-acc    mail-past-dec (hon)
    'Michael mailed a letter to George.'
    (Speaker: Tim, Addressee: Mr. Wilson)

The use of the honorific informal verbal ending *eyo* in utterance (9) indicates that the speaker honors the addressee and the situation of conversation is informal.

Finally, more than one individual may be honored by the speaker.

(10) Ewing    pwucang-nim-i           Garfield    cenmwu-kkey
         department director-hon-nom       executive director-dat (hon)
     selyu-lul         ponay-si-ess-supnikka?
     document-acc    send-hon-past-int (hon)
     'Did department director Ewing send a document to executive director Garfield?'
     (Speaker: Herbert, Addressee: Andrew)

In utterance (10) the use of the honorific morpheme *nim* before a nomi-

---

[3] The verbal ending *eyo* contains an inserted vowel *e*. The vowel *e* is inserted before the ending *yo* when the last syllable of a preceding morpheme has a dark vowel (for example, *e*, *wu*, *i*, and *u*) as its nucleus or when a preceding morpheme is the morpheme indicating past tense.

native case marker and that of the honorific morpheme *si* in a verb indicate that the speaker honors the subject referent. In addition, by the use of the honorific dative case marker *kkey*, the speaker honors the indirect object referent. Moreover, the speaker honors the addressee by using the honorific formal verbal ending *supnikka*. Thus the speaker of utterance (10) honors three individuals: the subject referent, the object referent, and the addressee.

### 2.2.3. Constraints on Occurrence of Honorification [4]

An occurrence of honorification in Korean is constrained by the relative order of social status among the people involved in an utterance. [5]

First, in order for a subject referent to be honored by the speaker, the constraints shown in (11) must be satisfied. [6]

(11) a. $Ref_{sbj} > Ref_{spk}$
    b. $Ref_{sbj} \geq Ref_{adr}$

The constraint in (11a) means that the social status of the subject referent must be higher than that of the speaker. The constraint in (11b) means that the social status of the subject referent must be equal to or higher than that of the addressee (in other words, the social status of the subject referent must *not* be lower than that of the addressee). The reason the constraint in (11b) must be satisfied is that when the social status of the subject referent is lower than that of the addressee, the speaker cannot show honor to the subject referent even though the social status of the subject referent is higher than that of the speaker.

Second, in order to honor an object referent, the constraints shown in

---

[4] The constraints discussed here are a revised version of the templates for a relation of social status that are explained in Lee (1996b).

[5] In the concept of social status the concepts such as social rank, seniority, and kinship are included. Although a relative order of social status is a predominant factor in the occurrence of honorification in Korean, it is not the sole and absolute factor. The following factors also play a role in honorification: the degree of closeness between the speaker and the addressee, the characteristics of the group to which people who are involved in dialogue belong, and the degree of proximity in social rank. These delicate factors are not considered in this paper.

[6] Here we use the symbols '>' and '≥' to stand for the relations 'higher than in social status' and 'equal to or higher than in social status', respectively. For example, the notation '$Ref_{adr} > Ref_{spk}$' indicates that the social status of the addressee is higher than that of the speaker.

(12) must be met.

    (12) a. $Ref_{obj} > Ref_{spk}$
       b. $Ref_{obj} \geq Ref_{adr}$

The constraints in (12) mean that the object referent must be higher than the speaker and must be equal to or higher than the addressee in social status.

    Third, in order to use a humble form of a verb, the constraints shown in (13) must be satisfied.

    (13) a. $Ref_{obj} > Ref_{spk}$
       b. $Ref_{obj} \geq Ref_{adr}$
       c. $Ref_{obj} > Ref_{sbj}$

The first requirement for the use of a humble form of a verb is that an object referent must be honored by the speaker (accordingly, the constraints in (13a) and (13b), which are the same as those in (12a) and (12b), are needed). The second requirement is that the social status of the object referent must be higher than that of the subject referent (accordingly, the constraint in (13c) is needed). This means that if the social status of the object referent is *not* higher than that of the subject referent, a humble form of a verb cannot be used even when the social status of the object referent is higher than that of the speaker.

    Finally, in order for the speaker to honor the addressee, the constraint shown in (14) must be satisfied.

    (14) $Ref_{adr} > Ref_{spk}$

The constraint in (14) means that the social status of the addressee must be higher than that of the speaker. Thus in the case of honoring the addressee, only the constraint on the relative order of social status between the speaker and the addressee is needed.

### 2.2.4. Honorification in Dialogue

Depending on the relative order of social status among the people involved in dialogue, the form of dialogue varies. Let us think about a dialogue in which four people are involved.

(15) a. Brown-ssi-kkeyse       K-nim-ul          poy-ess-e.
           Mr.-nom (hon)        hon-acc    meet (hum)-past-dec
      'Mr. Brown met K.'
      (Speaker: Sarah, Addressee: Molly)
    b. K-nim-i        choan-ul       kemthoha-si-ess-eyo?
          hon-nom     draft-acc    examine-hon-past-int (hon)
      'Did K examine a draft?'
      (Speaker: Molly, Addressee: Sarah)

In dialogue (15) conversation is held between Sarah and Molly. Two individuals (namely, Mr. Brown and K) are mentioned in their conversation. In utterance (15a) Mr. Brown and K are honored by the speaker (that is, Sarah), whereas the addressee (i.e., Molly) is not honored. In utterance (15b) K and the addressee (in this case, Sarah) are honored by the speaker (that is, Molly). It follows from this that the form of dialogue (15) is appropriate only to the situation in which the social status information shown in (16) is valid.

(16) K > Mr. Brown, Mr. Brown > Sarah, Sarah > Molly

The social status information contained in (16) is manifested through the honorification occurring in dialogue (15). Thus we can infer that information from the dialogue by using the constraints explained in Section 2.2.3.

(17) a. · Mr. Brown > Sarah, Mr. Brown ≥ Molly
          (from the honorification of the subject referent in (15a))
        · K > Sarah, K ≥ Molly
          (from the honorification of the object referent in (15a))
        · K > Mr. Brown
          (from the use of a humble verb in (15a))
        · Sarah ≥ Molly
          (from the fact that the addressee is not honored in (15a))
    b. · K > Molly, K ≥ Sarah
          (from the honorification of the subject referent in (15b))
        · Sarah > Molly
          (from the honorification of the addressee in (15b))

The social status information shown in (17a) is extracted from utterance (15a), whereas the information shown in (17b) is extracted from utterance

(15b). The information illustrated in (17a) and (17b) can be collapsed into
the information illustrated in (18a) and (18b), respectively. [7]

(18) a. K > Mr. Brown, Mr. Brown > Sarah, Sarah ≥ Molly
    b. K ≥ Sarah, Sarah > Molly

When we further collapse the social status information illustrated in (18a)
and (18b), we obtain the information shown in (16). Consequently, the
dialogue (15), which consists of utterance (15a) and utterance (15b),
correctly reflects the social status information shown in (16).

Let us now look at another dialogue, which is illustrated in (19).

(19) a. Minho-ka        K-lul        manna-ss-eyo.
                 nom         acc        meet-past-dec (hon)
      'Minho met K.'
      (Speaker: Sarah, Addressee: Molly)

    b. K-ka          choan-ul        kemthoha-yess-ni?
             nom         draft-acc        examine-past-int
      'Did K examine a draft?'
      (Speaker: Molly, Addressee: Sarah)

A close look at both the dialogue in (19) and the dialogue in (15) reveals
that although the contents of these dialogues are similar, their forms are
different. In the first utterance of dialogue (19), only the addressee is
honored by the speaker (namely, Sarah). In the second utterance (that is,
utterance (19b)) no one is honored by the speaker (i.e., Molly). Thus
honorification occurring in dialogue (19) is different from honorification
occurring in dialogue (15). This difference results from the difference in the

---

[7] In the process of collapsing information, the rules shown in (a) and (b) are used.

(a) Less informative information is discarded.
(b) Redundant information is discarded.

For example, since the information Mr. Brown > Sarah and Sarah ≥ Molly in (17a)
implies the information Mr. Brown > Molly, the less informative information Mr.
Brown ≥ Molly, which appears in (17a), is discarded during merging process in
accordance with the rule stated in (a). Similarly, the information K ≥ Sarah and
Sarah > Molly in (17b) implies the information K > Molly and thus the latter informa-
tion, which appears in (17b), is discarded during the process of merging in
accordance with the rule stated in (b).

relative order of social status among the people involved in each dialogue. The form of dialogue in (19) is appropriate only under the situation in which the social status information shown in (20) holds true.

(20) Molly > Sarah, Sarah ≥ Minho, Sarah ≥ K

Accordingly, social status information, which is provided through morphemes used in dialogue, plays a crucial role in determining the form of dialogue.


## 3. Context-Based Translation

In English dialogue the phenomenon of honorification does not occur. On the other hand, in Korean dialogue every utterance indicates whether or not honorification occurs in it. Thus in translating English dialogue into Korean dialogue on a computer, this difference must be taken into account.[8] Let us consider the utterance shown in (21).

(21) Section chief Johnson attended the meeting.
(Speaker: Carol, Addressee: Nick)

In utterance (21) three individuals are involved (namely, the speaker Carol, the addressee Nick, and Johnson, who is mentioned in the utterance). Since honorification does not occur in English, the form of utterance (21) does not change according to the relative order of social status among those three individuals involved in that utterance.

On the other hand, the phenomenon of honorification is always related to utterances in Korean. This means that when utterance (21) is translated into Korean, the result of translation must vary depending on the relative order of social status among the people involved in the utterance.

For example, when the order of social status shown in (22a) holds true with respect to the people involved in utterance (21), the result of translating that utterance into Korean must be the utterance appearing in (22b).

---

[8] The effects of honorification on the resolution of pronouns, the recovery of missing constituents, and coherence of dialogue in the parsing of Korean dialogue are discussed in Lee (1998). But the research deals with neither the generation of Korean dialogue, nor machine translation.

(22) a. Johnson > Nick, Nick > Carol

　　 b. Johnson　　kwacang-kkeyse　　　　　ku　　moim-ey
　　　　　　　　　　section chief-nom (hon)　　the　meeting-postp
　　　　　chamsekha-si-ess-eyo.
　　　　　attend-hon-past-dec (hon)

Since the social status of the speaker of utterance (21) is lower than that of the addressee and the subject referent as shown in (22a), honorific morphemes such as *kkeyse*, *si*, and *eyo* are used in (22b).

In the case where the order of social status shown in (23a) holds true, the translated Korean utterance corresponding to utterance (21) must be (23b), not (22b).

(23) a. Carol > Nick, Nick > Johnson

　　 b. Johnson　　kwacang-i　　　ku　　moim-ey
　　　　　　　　　　section chief-nom　the　meeting-postp
　　　　　chamsekha-yess-e.
　　　　　attend-past-dec

In the social status information specified in (23a), the social status of the speaker (that is, Carol) is higher than any other people. Thus honorific morphemes are not used at all as shown in (23b).

Moreover, when the relative order of social status shown in (24a) is valid with respect to the people involved in utterance (21), the appropriate translation of the utterance into Korean must be the utterance illustrated in (24b).

(24) a. Johnson > Carol, Carol > Nick

　　 b. Johnson　　kwacang-kkeyse　　　　　ku　　moim-ey
　　　　　　　　　　section chief-nom (hon)　　the　meeting-postp
　　　　　chamsekha-si-ess-e.
　　　　　attend-hon-past-dec

As shown in (24a), the social status of Johnson is higher than that of Carol and thus the subject referent (that is, Johnson) is honored by the speaker (namely, Carol) in utterance (24b). The addressee, however, is not honored by the speaker since the social status of the former is not higher than that of the latter as illustrated in (24a).

Likewise, in translating English dialogue, which consists of utterances,

into Korean, the use of social status information is needed as in the case of translating English utterance. As an example, let us assume that the dialogue illustrated in (25) is held under the situation in which the order of social status shown in (22a) holds true.

(25) a. Section chief Johnson attended the meeting.
  (Speaker: Carol, Addressee: Nick)
 b. Did he suggest a new plan?
  (Speaker: Nick, Addressee: Carol)
 c. Yes, he did.
  (Speaker: Carol, Addressee: Nick)

In this case the result of translating dialogue (25) into Korean is the dialogue illustrated in (26).

(26) a. Johnson  kwacang-kkeyse  ku  moim-ey
     section chief-nom (hon)  the  meeting-postp
   chamsekha-si-ess-eyo.
   attend-hon-past-dec (hon)  (same as utterance (22b))

 b. kupwun [9]-kkeyse  saylowun  kyeyhoyk-ul
   he (hon)-nom (hon)  new  plan-acc
   ceyanha-si-ess-ni?
   suggest-hon-past-int
 c. yey[10],  ha-si-ess-eyo.[11]
   yes (hon),  do-hon-past-dec (hon)

---

[9] In Korean there are two pronouns that correspond to English pronoun *he*. They are shown in the table illustrated below.

| | singular third-person pronoun for a male person |
|---|---|
| honorific | kupwun |
| nonhonorific | ku |

When the referent of a third-person pronoun is higher than the speaker in social status, the pronoun *kupwun* is used. Otherwise, the pronoun *ku* is used.

[10] There are two words that correspond to English word *yes*, as illustrated in the following table.

| | word for an affirmative reply 'yes' |
|---|---|
| honorific | yey |
| nonhonorific | ung |

Since social status information related to dialogue (26) is the same as that related to utterance (22b), the form of utterance (26a), which occurs in the dialogue, is also the same as that of utterance (22b). In utterance (26b) the subject referent is section chief Johnson, whose social status is higher than that of the speaker (i.e., Nick), and thus the honorific pronoun *kupwun* is used. In addition, since the social status of the speaker is higher than that of the addressee in utterance (26b), the former does not honor the latter, as indicated by the use of a nonhonorific verbal ending. In the final utterance of dialogue (26) the social status of the addressee is higher than that of the speaker. Accordingly, the honorific form *yey* which corresponds to English word *yes* and an honorific verbal ending are used.

Let us now consider the case in which the social status information stated in (23a) holds true with respect to the people involved in the dialogue which is illustrated in (25). In this case the appropriate translation of dialogue (25) into Korean is the dialogue shown in (27).

(27) a. Johnson      kwacang-i          ku      moim-ey
                  section chief-nom      the   meeting-postp
            chamsekha-yess-e.
            attend-past-dec                          (same as utterance (23b))

      b. ku-ka          saylowun    kyeyhoyk-ul
            he-nom          new          plan-acc
            ceyanha-yess-eyo?
            suggest-past-int (hon)

      c. ung,   ha-yess-e.
            yes,   do-past-dec

As shown above, the form of dialogue (27) is different from that of dialogue (26) since the social status information related to dialogue (27) is not the same as that related to dialogue (26). In both utterance (27a) and utterance (27c) the social status of the speaker (namely, Carol) is higher than that of any other people involved, and thus no honorification occurs in

---

If the social status of the addressee is higher than that of the speaker, the word *yey* is used. Otherwise, the word *ung* is used.

[11] When translating an English pronoun whose referent is the same as that of the pronoun occurring in the preceding utterance, it is more natural not to translate the pronoun at all. Thus, as shown in utterance (26c), the pronoun *he* occurring in English utterance (25c) is skipped in translation.

these two utterances. In utterance (27b) only the addressee is higher than the speaker (that is, Nick) in social status, and thus only the addressee is honored. Since the referent of the pronoun used in utterance (27b) is not higher than the speaker in social status, a nonhonorific pronoun *ku* is used.

Finally, let us look at the case in which the social status information stated in (24a) is valid with respect to the people involved in dialogue (25). In this case the appropriately translated dialogue is the dialogue illustrated in (28) (thus, neither the dialogue in (26), nor the dialogue in (27) is an appropriate translation in this case).

(28) a. Johnson      kwacang-kkeyse          ku      moim-ey
             section chief-nom (hon)       the   meeting-postp
        chamsekha-si-ess-e.
        attend-hon-past-dec            (same as utterance (24b))
    b. kupwun-kkeyse          saylowun    kyeyhoyk-ul
        he (hon)-nom (hon)        new        plan-acc
        ceyanha-si-ess-eyo?
        suggest-hon-past-int (hon)
    c. ung,     ha-si-ess-e.
        yes,   do-hon-past-dec

The subjects of utterance (28a) and utterance (28b) refer to the same person (that is, section chief Johnson). Since the social status of the subject referent is higher than that of the speaker in both utterances, the former is honored, as indicated by the use of an honorific morpheme and an honorific pronoun. In utterance (28b) the addressee is also honored by the speaker since the social status of the former is higher than that of the latter. In utterance (28a) and utterance (28c), however, the addressee is not honored. This is indicated by the use of a nonhonorific verbal ending or a nonhonorific responding word.

Therefore, in order to translate English dialogue into Korean dialogue appropriately, we have to make good use of social status information. If such information is not considered, a translated Korean dialogue would be very awkward since it cannot capture the context in which an English dialogue takes place.

# 4. Our Translation System and Implementation

Let us first take a brief look at the architecture of our translation system, which uses contextual information as well as linguistic information, and then consider how to implement that system on the computer.

## 4.1. Overview of the System

In our machine translation system we take the interlingua approach (Goodman and Nirenburg 1991; Hutchins and Somers 1992; Nirenburg et al. 1992), which is neutral between the source language and the target language. The architecture of the system is as illustrated in (29).

(29) Architecture of Our Machine Translation System:
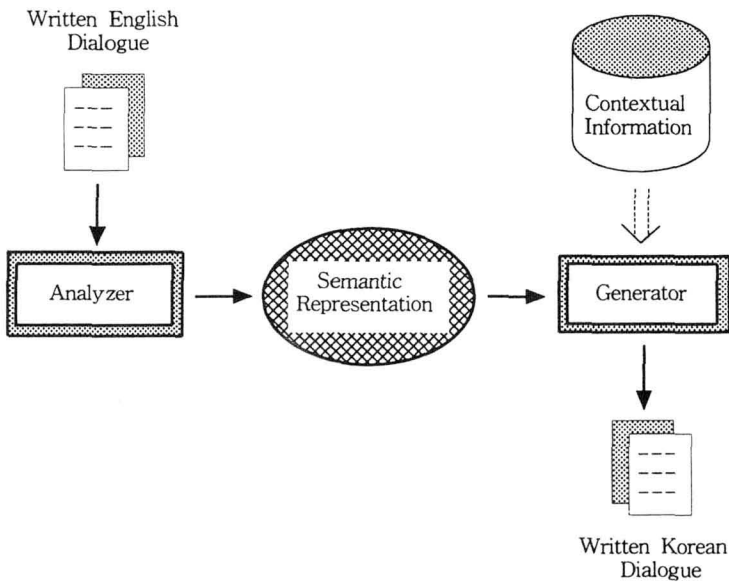


In our system the analyzer processes the utterance of the source dialogue and produces their semantic representation. This representation acts as the interlingua, whose form is independent of any natural language. From this interlingua the generator produces the appropriate target dialogue using contextual information.

## 4.2. Implementation of the System

We assign a speaker and an addressee to each utterance of an English dialogue and designate the relative order of social status for the individuals involved in the dialogue. We encode all relevant pieces of information needed in translating English dialogue into Korean dialogue utilizing the Prolog programming language.[12] Based on those pieces of information we produce appropriate Korean dialogue using a Sun SPARCstation which is a UNIX computer. Thus we implement machine translation of English dialogue in the way that when an English dialogue is given as an input along with contextual information, its corresponding and appropriate Korean dialogue is produced as an output.

The main predicates for a translation of English dialogue into Korean are as shown in (30).

(30)

```
% mtdiag_e2k/5
mtdiag_e2k(ContxtInfo,[UttPtc|RestUttPtcs],
        [Rel|RestRels],[EngUtt|RestEngUtts],KorDiag) :-
        engkor_diag_transl(ContxtInfo,[UttPtc|RestUttPtcs],[Rel|RestRels],
                        [EngUtt|RestEngUtts],KorDiag),nl,
        pretty_print(KorDiag),
        nl,user_response.

% engkor_diag_transl/5
engkor_diag_transl(ContxtInfo,[UttPtc|RestUttPtcs],[Rel|RestRels],
                [EngUtt|RestEngUtts],[KorUtt|RestKorUtts]) :-
        engkor_utt_transl(ContxtInfo,UttPtc,Rel,EngUtt,KorUtt),
        engkor_diag_transl(ContxtInfo,RestUttPtcs,RestRels,
                        RestEngUtts,RestKorUtts).
engkor_diag_transl(_,[],[],[],[]) :- !.
```

---

[12] For the Prolog programming language itself, refer to Bratko 1990, Clocksin and Mellish 1987, O'Keefe 1990, and Sterling and Shapiro 1994. In addition, for an application of Prolog to natural language processing, refer to Carpenter and Penn 1995, Covington 1994, Gal et al. 1991, Gazdar and Mellish 1989, Lee 1998, and Pereira and Shieber 1987.

```
% engkor_utt_transl/5
engkor_utt_transl(ContxtInfo,UttPtc,Rel,EngUtt,KorUtt) :-
    s(Sem,EngUtt,[]),                              % s/3
    ks(Sem,ContxtInfo,UttPtc,Rel,KorUtt,[]).       % ks/6
```

In the Prolog code shown in (30), the predicate *s/3* constitutes the analyzer of English utterances. The predicate processes the input English utterance and then gives us its semantic representation as illustrated in (31).

(31)
    | ?- s(SemIntpt,[section_chief_johnson,attended,the,meeting],[]).

    SemIntpt = specific(_A,meeting(_A),attended(section_chief_johnson,_A)) ? ;

    no
    | ?-

In (31) the value of the variable *SemIntpt* is the semantic representation of the input utterance.

The predicate *ks/6* shown in (30) constitutes the generator of Korean utterances. Based on the semantic interpretation such as the one illustrated in (31) and contextual information related to the input utterance (that is, the relative order of social status for the people involved in the utterance, the speaker and the addressee of the utterance, and the relationship between them), the predicate gives us the result of generating the output utterance. An example run of the predicate appears in (32).

(32)
    | ?- ks(specific(_A,meeting(_A),attended(section_chief_johnson,_A)),
            [hg(sc_johnson,nick),hg(nick,carol)],[carol,nick],informal,KorUtt,[]).

    KorUtt = [johnson_kwacang_kkeyse,ku,moim_ey,chamsekha_si_ess_eyo] ? ;

    no
    | ?-

The predicate *engkor_utt_transl/5*, which consists of the predicate *s/3* and *ks/6*, translates English utterance into Korean using contextual information as illustrated in (33).

(33)

  | ?- engkor_utt_transl([hg(carol,nick),hg(nick,sc_johnson)],[carol,nick],informal,
                            [section_chief_johnson,attended,the,meeting],KorUtt).

  KorUtt = [johnson_kwacang_i,ku,moim_ey,chamsekha_yess_e] ? ;

  no
  | ?-

On the other hand, when the information about social status or the relationship between the speaker and the addressee changes, the result of translation also changes even if the input utterance remains the same. This is illustrated in (34).

(34)

  | ?- engkor_utt_transl([hg(sc_johnson,carol),hg(carol,nick)],[carol,nick],informal,
                            [section_chief_johnson,attended,the,meeting],KorUtt).

  KorUtt = [johnson_kwacang_kkeyse,ku,moim_ey,chamsekha_si_ess_e] ? ;

  no
  | ?-

Thus, our translation system produces appropriate output utterance reflecting the situation under which the input utterance is spoken.

    The predicate *engkor_diag_transl/5* shown in (30) translates English dialogue, which consists of utterances, into Korean incrementally using contextual information. The top-level predicate *mtdiag_e2k/5* arranges the result of dialogue translation in an easily readable format and prompts user's response. The run of the predicate that translates the dialogue shown in (25) is illustrated in (35).

(35)

  | ?- mtdiag_e2k([hg(sc_johnson,nick),hg(nick,carol)],
               [[carol,nick],[nick,carol],[carol,nick]],
               [informal,informal,informal],
               [[section_chief_johnson,attended,the,meeting],
                [did,he,suggest,a,new,plan],
                [yes,he,did]],KorDiag).

<<< Translated Korean Dialogue >>>

[

[johnson_kwacang_kkeyse,ku,moim_ey,chamsekha_si_ess_eyo],

[kupwun_kkeyse,saylowun,kyeyhoyk_ul,ceyanha_si_ess_ni],

[yey,ha_si_ess_eyo]

]

Further Translation? y.

no

| ?-

As in the case of the translation of utterances, when the relative order of the social status for the people involved in the input dialogue or the relationship between dialogue participants varies, the result of translated dialogue must also vary. This is illustrated in the run appearing in (36).

(36)

| ?- mtdiag_e2k([hg(sc_johnson,carol),hg(carol,nick)],

[[carol,nick],[nick,carol],[carol,nick]],

[formal,formal,formal],

[[section_chief_johnson,attended,the,meeting],

[did,he,suggest,a,new,plan],

[yes,he,did]],KorDiag).

<<< Translated Korean Dialogue >>>

[

[johnson_kwacang_kkeyse,ku,moim_ey,chamsekha_si_ess_ta],

[kupwun_kkeyse,saylowun,kyeyhoyk_ul,ceyanha_si_ess_supnikka],

[ung,ha_si_ess_ta]

]

Further Translation? y.

no

| ?-

Therefore, our machine translation system produces different result of

translation according to the context in which the input dialogue is held, even if the contents of the dialogue is the same.[13]

## 5. Concluding Remarks

Up to now there have been a few attempts at machine translation of English text into Korean text. Those attempts, however, have never dealt with English dialogue. So this research is the first attempt to translate English dialogue into Korean dialogue. Since dialogue participants are involved in dialogue and linguistic phenomena occurring in English dialogue are different from those occurring in Korean dialogue, we have to make systematic use of pragmatic (or contextual) information as well as syntactic, morphological, and lexical information. By implementing machine translation of English dialogue into Korean dialogue in Prolog on a UNIX computer, this research showed that without using contextual information such as information about dialogue participants and social status information, we cannot get a corresponding Korean dialogue which is appropriate to the situation in which an English dialogue is held.

Our dialogue translation system, which is sensitive to the context related to dialogue, may be used in the translation of written dialogue appearing in the web sites on the internet. Furthermore, the system may also be applied to the speech-to-speech machine translation, which is very useful in the remote video conference. Thus, our system has various practical value and can be applied in many ways to the machine translation of written or spoken dialogue.

## References

Bratko, Ivan (1990) *Prolog Programming for Artificial Intelligence*, Second Edition, Addison-Wesley Publishing Company, Reading, MA.

Carpenter, Bob and Gerald Penn (1995) *ALE* (Attribute Logic Engine), Version 2.0.1 (available at the web site, http://macduff.andrew.cmu.edu/

---

[13] Although the input dialogue discussed here is rather short, the idea and method used in the process of translation can be equally applied to the translation of long English dialogue into Korean.

ale/ or by anonymous ftp from ftp.cs.cmu.edu:/user/ai/lang/prolog/
code/parsing/ale/), Choi, Key-Sun, Seungmi Lee, Hiongun Kim,
Deok-Bong Kim, Cheoljung Kweon, and Gilchang Kim (1994), 'An
English-to-Korean Machine Translator: MATES/EK,' in *Proceedings
of the Fifteenth International Conference on Computational Linguistics*
(COLING-94), 129-133, Kyoto.

Clocksin, William F., and Christopher S. Mellish (1987) *Programming in
Prolog*, Third, Revised and Extended Edition, Springer-Verlag, New
York, NY.

Covington, Michael A. (1994) *Natural Language Processing for Prolog
Programmers*, Prentice Hall, Englewood Cliffs, NJ.

Gal, Annie, Guy Lapalme, Patrick Saint-Dizier, and Harold Somers (1991)
*Prolog for Natural Language Processing*, John Wiley & Sons, New
York, NY.

Gazdar, Gerald and Chris Mellish (1989) *Natural Language Processing in
Prolog: An Introduction to Computational Linguistics*, Addison-
Wesley Publishing Company, Reading, MA.

Goodman, Kenneth and Sergei Nirenburg, editors (1991) *The KBMT
Project: A Case Study in Knowledge-Based Machine Translation*,
Morgan Kaufmann Publishers, San Mateo, CA.

Han, Eunjoo (1991) *Honorification in Korean*, Unpublished Manuscript,
Stanford University.

Hutchins, W. John and Harold L. Somers (1992) *An Introduction to
Machine Translation*, Academic Press, San Diego, CA.

Kim, Yong-Bum (1988) *A Fragment of Korean Phrase Structure Grammar*,
Doctoral Dissertation, Department of Linguistics, University of
Victoria, Available from Hanshin Publishing Company, Seoul.

Kuno, Susumu and Young-Joo Kim (1985) 'The Honorific Forms of
Compound Verbals in Korean,' in Susumu Kuno, John Whitman,
Ik-Hwan Lee, and Young-Se Kang, eds., *Harvard Studies in Korean
Linguistics*, 178-189, Hanshin Publishing Company, Seoul.

Lee, Dong-Young (1996a) 'An HPSG Account of the Korean Honorification
System,' in Claire Grover and Enric Vallduví editors, *Edinburgh
Working Papers in Cognitive Science*, Volume 12, Studies in HPSG,
169-190, Centre for Cognitive Science, University of Edinburgh.

_____ (1996b) 'Computation of Relative Social Status on the
Basis of Honorification in Korean,' in *Proceedings of the Sixteenth
International Conference on Computational Linguistics* (COLING-96),

693-698, Copenhagen.

_____ (1998) *Information-Based Processing of Korean Dialogue with Reference to English*, Thaehak Publishing Company, Seoul.

Nirenburg, Sergei, Jaime Carbonell, Masaru Tomita, and Kenneth Goodman (1992) *Machine Translation: A Knowledge-Based Approach*, Morgan Kaufmann Publishers, San Mateo, CA.

O'Keefe, Richard A. (1990) *The Craft of Prolog*, The MIT Press, Cambridge, MA.

Park, Byung-Soo (1991) *The Semantic and Pragmatic Nature of Honorific Agreement in Korean: An Information-Based Approach*, Unpublished Manuscript, Kyung Hee University.

Pereira, Fernando C. N. and Stuart M. Shieber (1987) *Prolog and Natural-Language Analysis*, CSLI Lecture Notes Number 10, Center for the Study of Language and Information (CSLI), Stanford, CA.

Sterling, Leon and Ehud Shapiro (1994) *The Art of Prolog*, Second Edition, The MIT Press, Cambridge, MA.

Suh, Chung-Soo (1978) 'Remarks on Subject Honorification,' in Chin-Wu Kim, editor, *Papers in Korean Linguistics*, 297-304, Hornbeam Press, Columbia, SC.

Weinstein, Clifford, Dinesh Tummala, Young-Suk Lee, and Stephanie Seneff (1996) 'Automatic English-to-Korean Text Translation of Telegraphic Messages in a Limited Domain,' in *Proceedings of the Sixteenth International Conference on Computational Linguistics* (COLING-96), 705-710, Copenhagen.

Department of English Language and Literature
Sejong University
Kunja-dong 98, Kwangjin-gu
Seoul 143-747, Korea
E-mail : leedy@kunja.sejong.ac.kr