# SENTENCE-STRUCTURE PLANNING FOR NATURAL EXPRESSIONS OF KOREAN

## 2B-7

Jong Hyeok LEE* Dong Un AN* Gil Chang KIM* Key-Sun CHOI** Kazunori MURAKI**

(*Korea Advanced Institute of Science and Technology,  **NEC Corp.)

## 1. INTRODUCTION

The conceptual representation, produced from Japanese or English texts by PIVOT system [1], is a meaning-oriented dependency structure with modality and pragmatic information such as TOPIC, GIVEN, NEW, and FOCUS.[2] A major difficulty in generating good styled expressions from conceptual representations is an absence of surface-level information on the target sentence-structure. This paper describes an attempt to plan the target sentence-structure with capability of generating more natural Korean expressions, which includes pragmatic and stylistic decision on target voice, stylistic changes of sentence-structures, and replacement of function word with content word having a supplementary meaning.

## 2. PRAGMATIC AND STYLISTIC DECISION ON TARGET VOICE

The voice generation is important to synthesis of a natural target sentence retaining information structure of a source sentence. But there are several difficulties in generating Korean voice, which are caused by the following characteristics of Korean passives:

- Complex passive verb system with low productivity:

  The passive-verb system of Korean is not so uniform as that of English or Japanese. There are three conventional passive forms and the use of each passive forms is restricted to a specific type of verbs.

- Heavy passive constraints:

  Being compared with English or even Japanese, Korean has many passive constraints such as verb constraints, object constraints, and meaning constraints. Especially Korean does not allow an indirect passive.

To cope with the difficulties, scope of passives is extended to notional passives like activo-passives which are morpho-syntactically active but semantically passive, and information on exceptions to passivization is encoded into each lexical item. Autonomous passives, by which spontaneous or accidental processes are expressed, are generated directly from the dictionary, not through passivization.

The decision on the target voice depends upon both pragmatic and stylistic factors:

- Pragmatic factor (topicality):

  The passive is preferable to the active when a low-activity deep case is more topical than high-activity ones [3]. The most topical one among several deep cases(NPs) is identified using the topicality hierarchy:

  1. TOPIC
  2. GIVEN > NEW
  3. Human(1st > 2nd > 3rd) > Animate > Inanimate
  4. Agent > Reason > Instrument > Recipient > Object > Material

- Stylistic factor:

  In Korean, even though the high-activity deep case is more topical, the target voice becomes passive when it is a non-agent.

Then the sentence-structure planning phase checks the passive constraints and, if necessary, performs

an activo-passivization, by which conceptual symbol for predicate is replaced with a new one having passive meaning, and the deep cases are also changed. Figure 1 shows the stages of deciding the target voice [4].

## 3. STYLISTIC CHANGE OF MEANING STRUCTURE

Korean has several stylistic differences from English and even Japanese, so a same event in the world has to be often expressed differently in Korean from in other languages. Because the conceptual representation is designed to be language-independent as an interlingua of MT, the sentence-structure planning for good styled expressions is needed, which results in structural/lexical changes of sentence-structures:

- for activo-passivization,
- for predicative use of quantitative adjectives,
- for denominal verbs, etc.

## 4. REPLACEMENT OF FUNCTION WORD WITH CONTENT WORD

Surface realization of deep cases using only function words may often cause meaning ambiguities. In such a case, the planning procedure makes a structural change of sentence-structure. For example, a genitive case-marking particle in Korean, whose typical deep case is Possessor, may be supplemented with content words with meanings such as Agent, Location and so on. [2]

## 5. CONCLUSION

Some techniques of sentence-structure planning for natural expressions are introduced. But for the full stylistic planning, the use of real-world pragmatic information must be studied in a long run.

REFERENCE

[1] Muraki, K.(1986) "VENUS: Two-phase Machine Translation System", In Future Generation Computer Systems, 5.
[2] 村木 (1984.12.17)「知識ベースと言語に独立の中間表現とを用いた日英機械翻訳システム」、日経エレクトロニクス、pp.195-220.
[3] Givon, T.(1976) On Understanding Grammar, New York: Academic Press.
[4] Lee, J. and Kim, G.(1988) "Generation of Voice from Conceptual Representation: syntactic, semantic, and pragmatic aspects", to be appeared in Literary and Linguistic Computing, Vol.3, No.4.
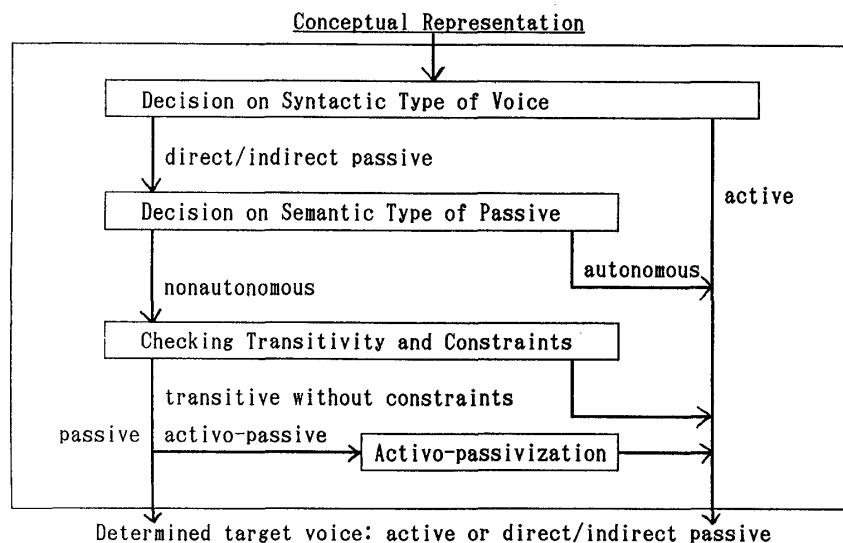
Conceptual Representation



Figure 1. Decision on Target Voice