Workshop on Telephony Interpretation
Sep. 1986, at Osaka

Problems of Knowledge Processing in Telephony Interpretation

Hirosato Nomura
NTT Basic Research Laboratories
Musashino-shi, Tokyo, 180, Japan

1. Introduction

Telephony interpretation by computer is an application of multi-lingual natural language processing including machine translation. However, its current technology is definitely insufficient for designing a useful telephony interpretation system. Natural language processing technology has mainly been concerned with processing written text while telephony interpretation technology needs to process spoken text. Spoken texts have sometimes a very different style compared with that of written texts: spoken linguistic expressions are frequently ungrammatical, and the sequence of utterances are occasionally not well organized. Moreover, especially in telephony conversation, there are some difficulties concerning the need for participants to share common knowledge of circumstances to be assumed in carrying out the conversation. Without sharing common knowledge, the two participants might not be able to reach a goal of the conversation. Such kinds of tasks should be accomplished by adopting so-called knowledge processing which can manipulate daily-life knowledge, domain-specific expert knowledge, and linguistic knowledge simultaneously. However, again, current knowledge processing technology is still in very elemental stage for this purpose, and thus we need to develop a new theoretical framework for manipulating knowledge in language understanding and language generation.

This paper discusses fundamental problems of knowledge processing in telephony interpretation and suggests an approach to it. This paper does not concern so-called voice recognition, and also does not concern the details of language processing.

2. Knowledge in language understanding

Knowledge plays important role in language understanding. For example, in the sentence:

She dropped *the plates* on the table and broke *them*.

we can guess that '*them*' is referring to '*the plates*'. This can be resolved by applying a kind of linguistic knowledge, number agreement, since '*them*' is a pronoun whose number is plural and '*plates*' is another noun whose number is also plural and can be referred to by it in the sentence.

However, if the sentence is changed to the following:

She dropped *the plate* on *the table* and broke *it*.

we have no way of knowing whether '*it*' refers to '*the plate*' or '*the table*' without being given, for example, contextual information.

On the contrary, in the following example of a noun phrase involving a conjunction '*and*',

carbon *and* nitrogen tetraoxide,

we need a kind of domain-specific expert knowledge of chemical material to resolve the ambiguity of the conjunctive structure. Linguistically, two structures,

(carbon) *and* (nitrogen tetraoxide) and
(carbon tetraoxide) *and* (nitrogen tetraoxide),

are possible; however, the second structure is not correct since there is no chemical compound carbon tetraoxide. Thus, this problem is solved by domain-specific knowledge but not by linguistic knowledge.

Also we need situational knowledge sometimes to understand the correct meaning of a sentence, for example,

I saw the man *on the hill* *with a telescope*.

It is easy to know that there are at least three possible interpretations, 1) I was holding a telescope and used it to see the man standing on the hill, 2) I saw the man who was standing on the hill which has a telescope on it, and 3) I was standing on the hill, and used a telescope to see the man. However, we cannot identify which is the correct one only by linguistic knowledge, and therefore, we need situational knowledge to resolve the problem.

Also some problems of ellipsis and anaphora need knowledge to identify them in language expressions. For example, in the following sentence:

Taro bought a bicycle. But Hanako *did* not.

we need both linguistic and daily-life knowledge to know that '*did*' represents '*bought a bicycle*'. As for anaphora, in the following sentences:

There is *a* bicycle over there. *The* bicycle is owned by Taro. *A* bicycle is very convenient for students.

we must distinguish the meaning of two '*a*' articles. In the following sentence:

I sold a car, and bought a new bicycle with *the money*.

we need to know the fact that '*the money*' was got by selling the car.

As such, we need much knowledge to understand the correct meaning of sentences. Such kind of knowledge concerns linguistic knowledge, daily-life common sense knowledge, and domain-specific expert knowledge. However, these kinds of knowledge are related each other and should be utilized simultaneously in understanding.

3. Problems in knowledge processing

Research on knowledge processing has been concerned mainly with 1) how to represent knowledge in the computer, 2) how to use stored knowledge to infer a new fact from a given fact, and 3) how to design a so-called expert system by applying these technologies. For the knowledge representation, some representative frameworks such as semantic networks, frame systems, and logical expressions have been proposed. The semantic network has been adopted to represent relationships holding among chunks of knowledge. It has also been adopted to represent analyzed linguistic

structures of a sentence in terms of natural language processing. A frame system is a representative framework for representing typical knowledge about an object, circumstance or situation. Logical expression is also a representative framework for representing facts in the style of logical formula. Some inference mechanisms have been proposed in terms of theorem proving in this framework. However, most attempts have concerned knowledge in a limited domain, treating it as rather static data. As for the strategies for using stored knowledge, they have been manipulating knowledge or data in very straightforward manner. It seems that the inference mechanisms in the human brain are very complicated and their methods and capabilities are very different from the current computer processing methods. As such, current technologies are not sufficient to be applied to designing a really powerful interpreting system and thus we need to come back again to the point from which we can start fundamental study of formalization of knowledge, its representative framework, and inference mechanisms.

Fundamental problems of knowledge processing can be listed as four sub-problems: 1) systematization of knowledge, 2) representation of knowledge, 3) utilization of knowledge, and 4) maintenance of knowledge. Systematization of knowledge relates to formulating each chunk of knowledge and relating them to each other. Representation of knowledge concerns how to represent each chunk of knowledge and organize them in a structured memory. Utilization of knowledge relates to providing inference mechanisms. And, maintenance of knowledge is for modifying the structured memory.

Among these sub-problems, the problem of systematization of knowledge has not been considered seriously. However, we should emphasize it and we should know what knowledge is required for telephony interpretation. We might assume kinds of knowledge as follows:

Linguistic knowledge,	Intention,
Situational knowledge,	Common-sense knowledge,
Expert knowledge,	Meta-knowledge.

Linguistic knowledge consists of knowledge for each language and contrastive knowledge between languages. Intention concerns what is the purpose of utterances or conversation. Situational knowledge specifies the background of the conversation. Common-sense knowledge is daily-life knowledge we can assume that most people know. Expert knowledge is domain-specific knowledge such as that for computers, economics, etc. Meta-knowledge concerns manipulating the stored knowledge during processing.

4. A unified framework for linguistic and non-linguistic knowledge

In language understanding, both linguistic and non-linguistic knowledge are utilized simultaneously. Therefore, both should be organized in memory structure simultaneously. Moreover, if we can formalize them in a unified framework, it is very convenient to apply them in language understanding. For this purpose, this section presents a fundamental consideration for proposing a unified approach, as indicated in figure 1.

We assume a communicative model as shown in figure 2 as a basis of the understanding model. The communicative model consists of participants in communication. The participant might be human, machine or text depending on the situation; however, in any case, each is defined as a transceiver model. The transceiver model process, for example, consists of hearing, interpreting with planning, and speaking as shown in figure 2. These sub-processes relate to cognition, problem solving, and response respectively in terms of the cognitive science view of the human brain, while they can be interpreted as

analysis, transfer and generation respectively in terms of machine translation, as shown in figure 3. In any terms, all these sub-processes are characterized by a knowledge-base and processing mechanism. Therefore, we need to propose a common basis for modeling them. Since the objective of the communication is to transmit the intended meaning by means of language expressions, it is important that a semantic model is proposed as a common basis for modeling the transceiver model, as indicated in figure 4. Such a semantic model can be specified by three sub-models:

- Linguistic model: giving a lexical and grammatical framework,
- Situation model: giving an environment for utterance and interpretation,
- Memory model: giving a representative framework for storing and recalling knowledge.

These sub-problems relate to knowledge as described in the previous section.

By applying the semantic model to the knowledge and the algorithm, the construction of the knowledge-base and the processing-mechanism are obtained. Compared to the human brain, the knowledge-base corresponds to "knowledge", and the processing-mechanism to "intelligence".

The knowledge-base consists of three parts: intra-linguistic knowledge, inter-linguistic knowledge, and extra-linguistic knowledge. Intra-linguistic knowledge concerns the lexicons and grammars of individual languages such as Japanese and English. Inter-linguistic knowledge connects different languages in terms of their lexicons and grammars. Extra-linguistic knowledge consists of common-sense knowledge, expert knowledge, and other types of knowledge listed in the previous section. The processing-mechanism consists of three sub-mechanisms: analysis, transfer, and generation.

By integrating the knowledge-base and the processing-mechanism, we can now design an understanding system, and thus they can be applied to several kinds of application systems such as question answering, text generation, machine translation, and telephony interpretation. Figure 4 shows the model construction of a natural language understanding system and its application to practical systems mentioned above.

From these fundamental assumptions, we can now propose a memory structure which can manipulate both linguistic and non-linguistic knowledge simultaneously in a unified framework. The memory structure consists of three layers: 1) long-term memory, 2) discourse memory, and 3) episodic memory as shown in figure 5. Long-term memory stores knowledge related to linguistic knowledge such as dictionaries and grammar, and non-linguistic knowledge such as experiences, common-sense, and expert and procedural knowledge. Procedural knowledge relates to inferring a fact from a collection of facts. Discourse memory provides situational and contextual information for an utterance environment. Episodic memory stores the meaning of the ongoing segments of utterances. As understanding proceeds, the essence of episodic memory is assimilated into discourse memory and the essence of discourse memory is assimilated into long-term memory.

5. Classification of knowledge

As a basis for organizing linguistic structure, we assume three kinds of semantic primitives: structure, relation, and concept, as shown in figure 6. A variety of structures can be used to represent the hierarchical structure of linguistic knowledge and sentence meaning, as shown in figure 7. There are two types of semantic structures, composite and primitive. A composite structure is made by integrating semantic structures using semantic relations. A primitive structure, by definition, cannot be divided into further sub-structures. In general, a single word corresponds to a primitive structure,

and a phrase corresponds to a composite structure. Since syntactic information can also contribute to define meaning structures, each semantic structure simultaneously incorporates not only meaning information but also syntactic information.

Semantic relations connect semantic structures and then build larger semantic structures, compound structures. For semantic relations in sentence meaning structures, there exist various kinds of semantic relations such as those shown in figure 8.

Concepts are associated with the structures mentioned above. Among them, as shown in figure 9, concepts associated with a word structure represent the word meaning that appears when the word is used in a sentence. A word meaning is represented by a set of principal concepts, supplementary concepts, and their semantic dependencies. Principal and supplementary concepts are defined by using semantic categories, and prepared for nouns, adverbs, verbs, adjective-verbs and modalities. Semantic dependencies are defined by using semantic relation frames and semantic structure frames. Semantic categories, semantic relation frames and semantic structure frames have the following characteristics: 1) They have two types of concepts: prototype and instance. Prototypes partially determine selectional constraint and define semantic dependency structures. Instances show an assimilated structure that satisfies the selectional constraints. 2) They show semantic commonness and analogy between two structures. This allows the system to share information and to provide facilities for paraphrase. 3) They make up a hierarchical structure. This provides the system with inheritance ability and information sharing.

Non-linguistic knowledge is also defined in terms of the three primitives mentioned above. Figure 10 shows three kinds of such knowledge, concept relation, event state relation, and meta-knowledge. As such, non-linguistic knowledge can be defined and manipulated in the same way as linguistic knowledge.

References

[1] Nomura, H., A Model for Knowledge Acquisition Systems with Problem Solver, Proceedings of the Fourth International Congress of Cybernetics and Systems, Aug. 1978.

[2] Nomura, H., Artificial Intelligence Approach to Machine Translation, ACM SIGART Newsletter, Jan. 1982.

[3] Nomura, H., Towards the High Ability Machine Translation, European-Japanese Workshop on Machine Translation, Nov. 1983.

[4] Nomura, H., Experimental Machine Translation Systems: LUTE, Second European-Japanese Workshop on Machine Translation, Dec. 1985.

[5] Shimazu, A., Naito, S., Nomura, H., Japanese Language Semantic Analyzer Based on an Extended Case Frame Model, Proceedings of the International Joint Conference on Artificial Intelligence, Aug. 1983.

[6] Nomura, H., The LUTE Project, in Nirenburg, S. (ed) Special Section on Machine Translation of Natural Languages, ACM SIGART Newsletter, No.92, May 1985.

[7] Nomura, H., Modeling and Representative Framework for Linguistic and Non-linguistic Knowledge in Natural Language Processing, Second Germany-Japan Science Seminar, May, 1986.

[8] Nomura, H., Naito, S., Katagiri, Y., Shimazu, A., Translation by Understanding: A Machine Translation System LUTE, Proceedings of the International Conference on Computational Linguistics, Aug., 1986.

Fig. 1 Unifying Linguistic and Non-linguistic Knowledge

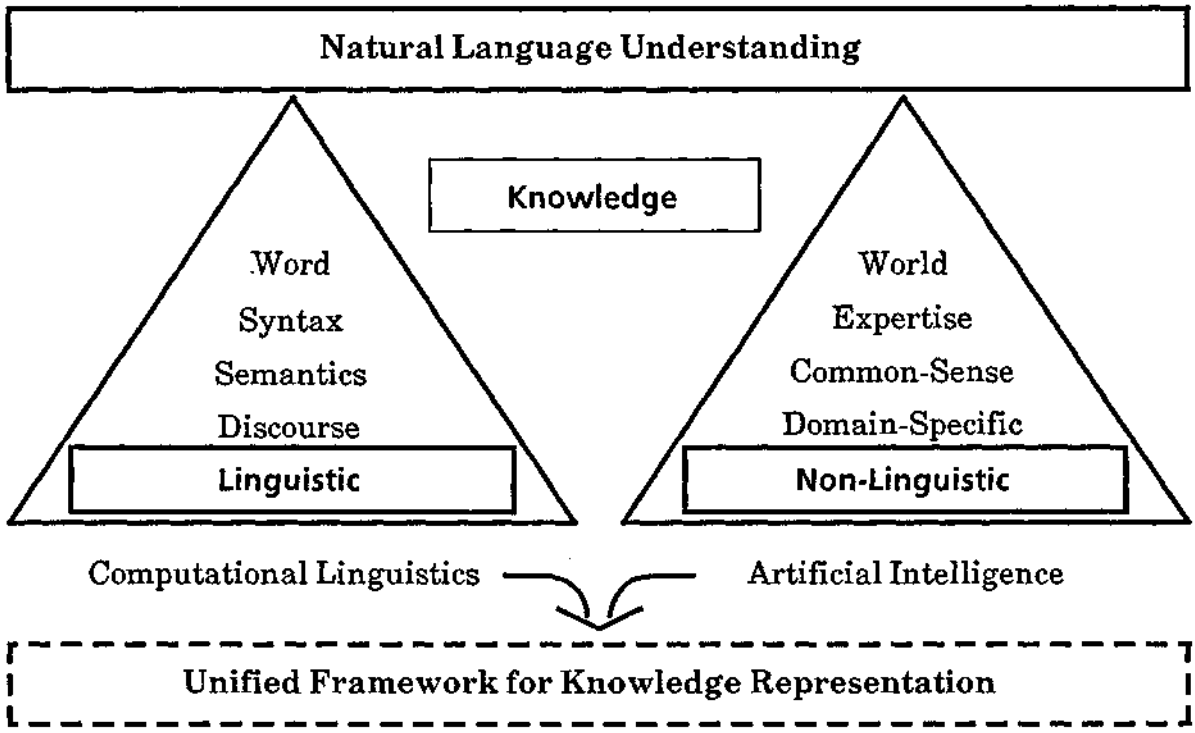

Fig. 2 Communicative model

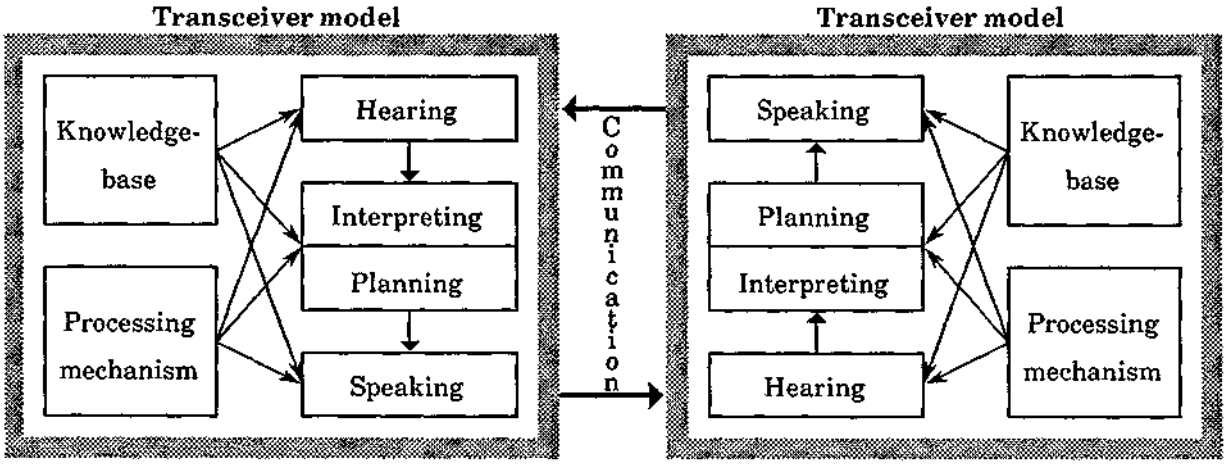

Fig. 3 Meaning Understanding Model

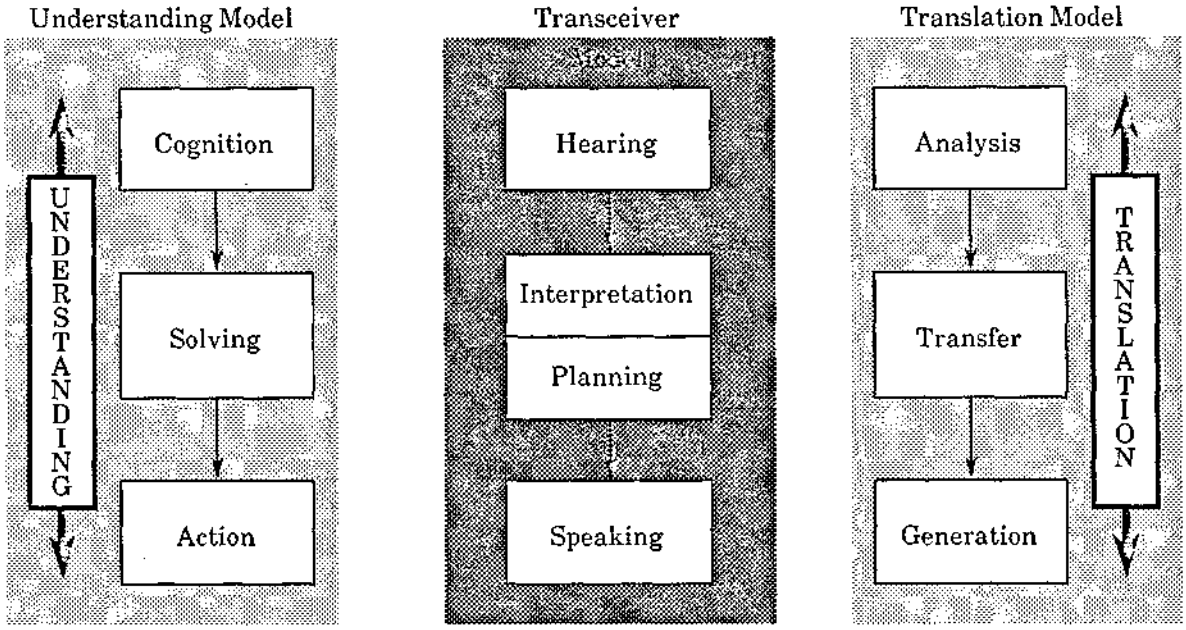

Fig. 4 Model for natural language understanding and its applications

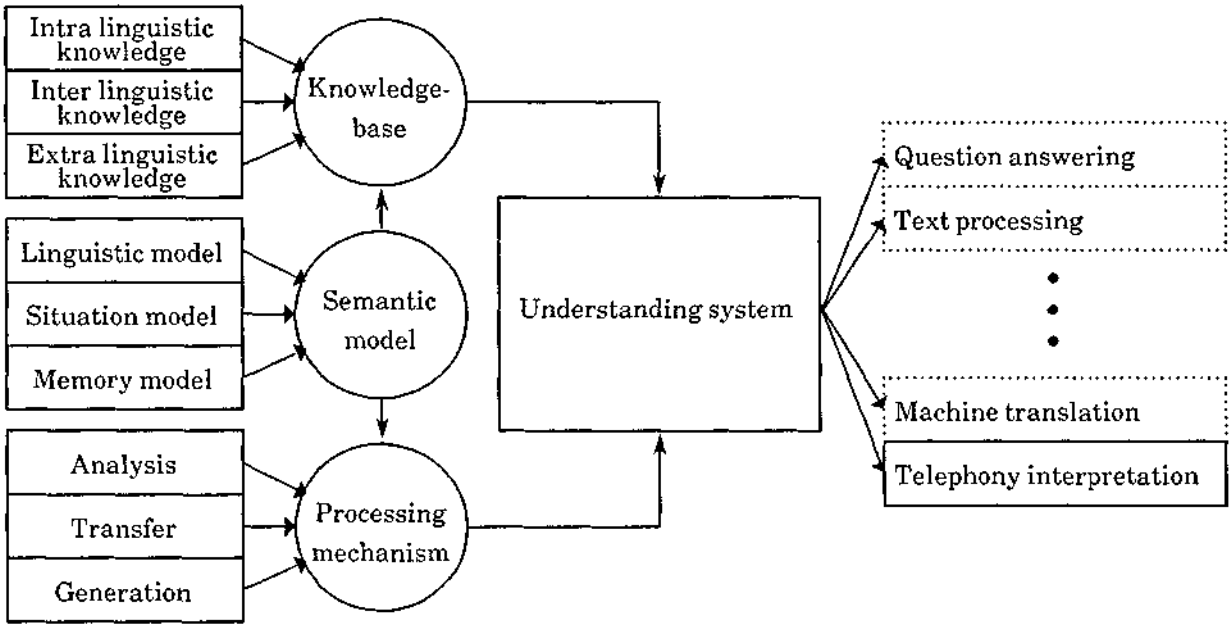

Fig. 5 Memory Structure

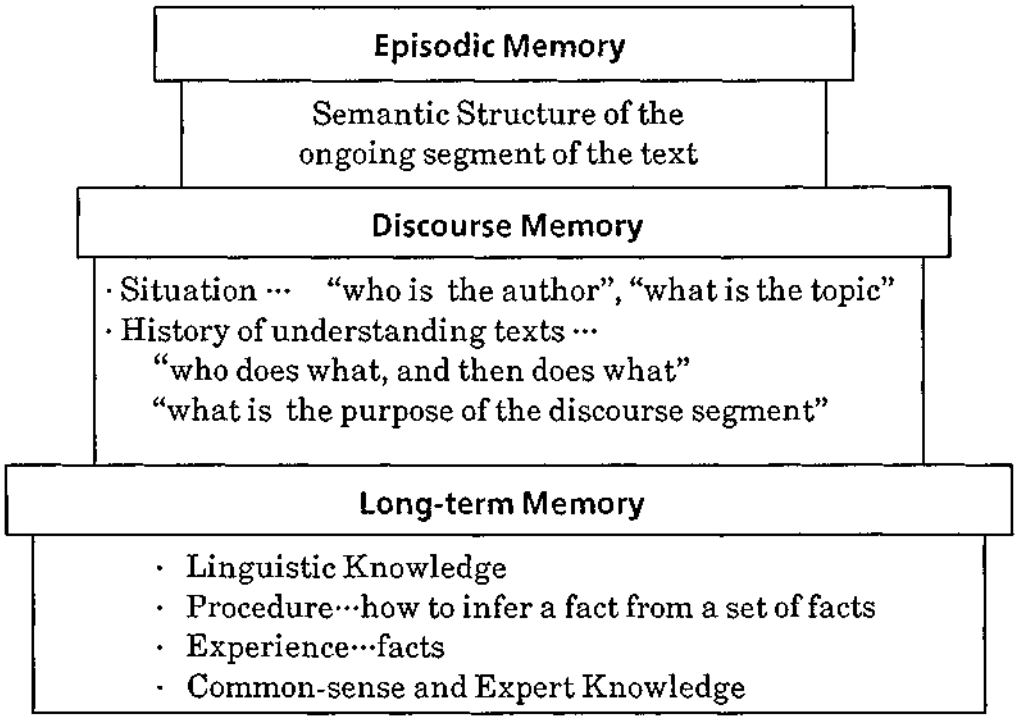

Fig. 6 Primitives for Memory Structure

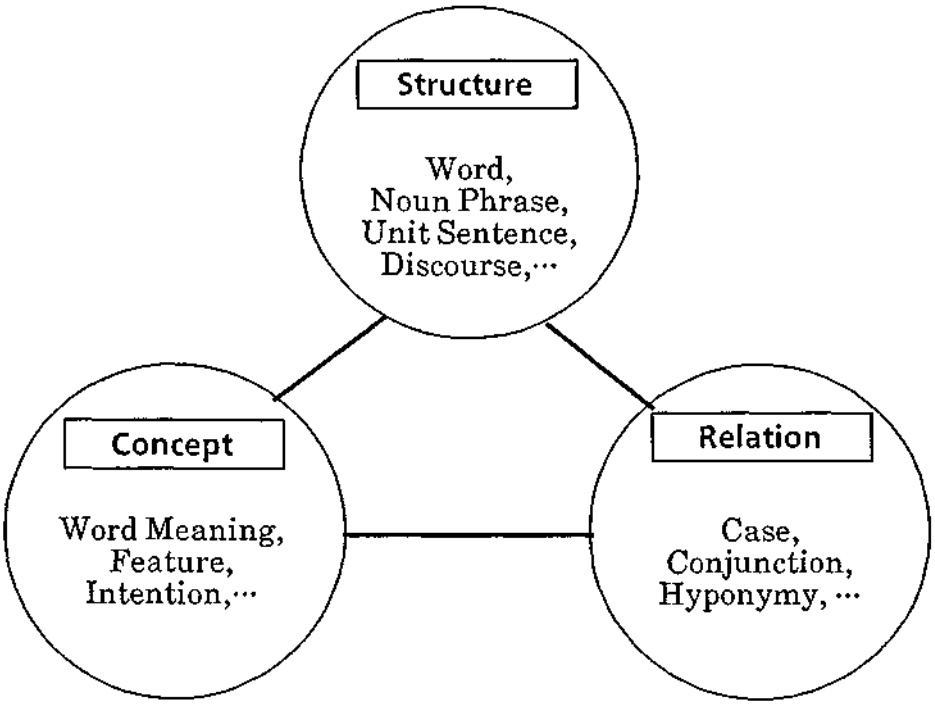

Fig. 7 Semantic Structure

1) Two-types:
   Primitive Structure: cannot be divided into substructures
   e.g. Single-word structure

   Composite Structure: integration of semantic structures with semantic relations
   e.g. Phrase or Sentence Structure

2) Language-dependent: J-ECS for Japanese language
   E-ECS for English language

3) ECS incorporates Linguistic structure and Conceptual structure in a unified framework called Frame-Network

Fig. 8 Semantic Relation

1) Noun relation: between nouns
   e.g. Whole-part, Upper-lower, Possession, Material

2) Case relation: between a case element and a predicate
   e.g. Object, Agent, Instrument, Place, Time

3) Embedded relation: between an embedded sentence and a noun phrase
   a) Case relation between a modified noun phrase and the predicate in the embedded sentence
   b) Noun relation between a modified noun phrase and a noun phrase in the embedded sentence
   c) An appositive or subsidiary relation between a modified noun phrase and the embedded sentence

4) Conjunctive relation: between sentences
   e.g. Cause-result, Time-advance, Assumption, Parallel

Fig. 9 Concept

1) Associated with Structures
   Word structure → Word meaning

2) Word Meaning:
   Principal and supplementary concept → Semantic category
   Semantic dependency → Semantic relation frame, and
   Semantic structure frame

3) Characteristics of Concept:

   a) Two Ttypes: Prototype → Selectional constraint
   Instance → Assimilated structure

   b) Semantic commonness or analogy between structures
   → Information sharing and paraphrase facility

   c) Hierarchical structure
   → Inheritance facility and information sharing

Fig. 10 Non-linguistic Knowledge

1) Concept Relation: Hyponymy, Synonymy, Antonymy, Whole-part, Possession
   e.g. "whole-part" relation between train and window

2) Event State Relation: between two events or between an event and a state
   e.g. "subsidiary situation" relation between 'grill' and 'smell'
   "秋刀魚を焼くにおい"

3) Meta knowledge for reasoning:
   Traversing the concept networks, and
   Checking semantic consistency.