On the Processing Units in Simultaneous Interpreting

On the Processing Units in Simultaneous Interpreting

Chuta Funayama

1. INTRODUCTION

Samples of simultaneous interpretation are commonly shown as language conversion by phrases while written translation works are exemplified by sentences or texts. What constitutes the units of conversion in simultaneous interpretation? This paper discusses how we should define those processing units.

First of all we can distinguish two different views of the processing units in simultaneous interpretation. One is taken to explore the way to overcome the word order difference between two languages, especially Japanese and English. This view is concerned with ingenuity in cutting an original speech into pieces appropriate to quick yet close enough translation. Those pieces are intuitively based on the units of information. What then constitutes those units of information? To answer this question we should understand more about another essential ingredient of the linguistic act of simultaneous interpreting, the language understanding process needed for verbal communication in general. In what way does a hearer develop his/her understanding on-line as taking in what a speaker says? Presumably a hearer has a temporally changing accumulation of information at every point in time, enabling the establishment of contexts, the prediction of story development, or the modification of previous predictions. The processing units in simultaneous

interpreting should reflect the basic mechanism of human language understanding.

Although these two views of the processing units in simultaneous interpretation may lead to the identical conclusion as to specific units at work, they should be distinguished from the standpoint of a research target. Strategies for overcoming word order differences, on the one hand, pay particular attention to how to translate after arbitrarily fixing processing units in original speech, while on the other hand a particular unit of information processing is determined not by how to translate but by principles underlying human language understanding. In other words, as a matter of sequence, an utterance comes first, giving information piece by piece to form a minimum set of information, which is then available for further processing and the reverse is not the case; translatability does not come before processing units are detected by some mechanism in the incoming utterance. It is considered therefore that a processing unit for original speech constitutes part of a translation unit at least in terms of time.

In this paper we are going to discuss the processing units in simultaneous interpretation from the viewpoint of the general principles of natural language processing, providing for the basis for designing advantageous translation patterns. Research in parsing, such as Mazuka and Itoh (1995) for example, has been mainly concerned with assigning syntactic structures to incoming strings of words. Such studies do not seem to satisfy those who are interested in on-line holistic interpretation. Mizuno (1995b: 9), for example, states that "interpreters are concerned about when they can start interpreting." This paper intends to fill the gap between what the current studies in parsing show and what is expected of a model of language processing by

moving one step further toward the integration of semantic interpretation into the processing model of natural language.

2. SEMANTIC REPLACEMENT UNITS AS PROCESSING UNITS

Frazier (1985), observing that people can manage to understand very long sentences despite severe restrictions on the immediate memory capacity of humans, claims that a sentence processor is licensed to forget nonsemantic information (e.g., information about the phonological or syntactic representation of some portion of a sentence) at a certain point and proposes what is called "semantic replacement units" to show how long nonsemantic information needs to be retained. In this view the process of language comprehension proceeds on-line, leaving behind the nonsemantic information of each semantic replacement unit. Such conception agrees with our intuition and experience. It, for instance, reminds us of the case that while it seems extremely difficult to remember exact sentence constructions and phrases even for a short time, we can possibly remember what the speaker said for a relatively long time. Such experience tells us that retention is relevant not to the expression level but to some conceptual level. Then we might ask if a semantic replacement unit coincides with a processing unit in simultaneous interpreting. Here in this section we are going to discuss the specific notion of semantic replacement unit defined by Frazier and try to characterize the processing unit in simultaneous interpreting.

Frazier (1985) assumes that "complete minimal governing categories" are semantic replacement units. The component concept of governing category comes from syntax and is defined by Chomsky

(1981:188) as follows:

(1) α is the governing category for β if and only if α is the minimal category containing β and a governor of β, where α = NP or S

The term governing category is employed in the theory of binding to designate the domain in which anaphors must be bound and pronominals must be free. Based on the concept of binding in what is called the GB theory, Frazier (1985:173) defines "minimal governing category" as in (2):

(2) the minimal NP or S containing both a governor and governed material (where tense governs the subject, verbs and prepositions govern their objects, and nouns govern their complements)

The element "complete" in the present term is meant by Frazier (1985:174) as:

(3) a minimal governing category containing no unassigned pronoun, bound anaphor (reciprocal or reflexive) or gap

Assuming that semantic replacement units indicate when a sentence processor may forget nonsemantic information, one finds it natural to expect that the antecedents of anaphors or pronominals be determined inside those units. For example, "themselves" in (4a) is bound by its antecedent, " the girls," inside its complete minimal governing category, which is bracketed in (4a). In this instance we

could regard the complete minimal replacement unit as a semantic replacement unit without any problems caused. In the case of (4b), however, in order for the gap, which is indicated by an underlined space, to be given its antecedent a considerably long complete minimal governing category such as shown by the pair of brackets is needed. [1] This length cannot be appropriate as the processing unit in simultaneous interpreting, as will be discussed more later.

(4) a. Mary was laughing because [the girls tickled themselves].

b. [The elephant trainer who John thought _____ and the zookeeper who Peter was sure had appeared on the Tonight Show last week] were eating lunch at the next table.

Frazier (1985) cites two reasons why complete minimal governing categories serve as semantic replacement units: (i) They are semantically complete; (ii) They are effective in interpreting anaphors and pronominals. We can intuitively accept the idea that clauses and noun phrases are semantically complete. In practical reality many professional interpreters, if speech texts are available beforehand, mark the texts, putting slashes usually after each noun phrase or clause in sentences in preparation for simultaneously interpreting

1 The example (4b) is identical with (48) of Frazier (1985:174) except two things: One is the insertion of brackets for easier explication of the point. The other modification is the deletion of the second underlined space. The reason why Frazier's example contains two gaps is supposed to be related to the syntactic analysis of the Right Node Raising. Frazier appears to assume that the raised element is Chomsky-adjoined to the S node.

those texts. We also find (ii) natural if we consider that the concept of governing category is a theoretical construct to account for the syntactic behaviors of pronominals and anaphors. Frazier further claims that governing categories are relevant throughout language processing regardless of the occurrence of pronominals, because no difference in processing difficulties has been experimentally recognized between (5a) and (5b).

(5) a. Mary was laughing because [*the girls* tickled *them*].
b. Mary was laughing because *the girls* tickled *him.*

If a governing category is something to be added in the course of sentence processing by having an anaphor or pronoun triggering it, then only in (5a), where a decoy antecedent (*the girls*) occurs in the governing category of *them,* complication takes place, whereas no complication is expected in (5b), where the corresponding governing category may not be provoked because *the girls* is not a potential antecedent of *him.* If this assumption is correct, then the processing of (5a) should be more difficult than that of (5b). Experimentally, however, no difference is reported to have been found between the two sentences. Frazier therefore argues that this supports the idea that the processor keeps track of minimal governing categories all the time.

We now examine whether semantic replacement units serve as the processing units in simultaneous interpreting as well. The idea of forgetting nonsemantic information at a certain point in processing agrees quite well with the observations of simultaneous interpretation. It is reported in Funayama (1994) that the process of

simultaneous interpreting includes a certain group of phenomena that suggest some form of conceptualization. For instance, there are cases where word forms are hardly believed to be retained as they are by the time of output. Also observed is the use of output words which are usually not included in dictionaries as candidate equivalents yet conceptually related to the original expression. This phenomenon becomes more conspicuous as the retention time is extended more because of some hindrance to immediate translation. Such observations, which motivate the proposal of "cognitive files" to be discussed in section 4, are compatible with the concept of semantic replacement unit.

The specific idea of semantic replacement unit as proposed by Frazier (1985) poses two basic problems. First, as mentioned above regarding (4), to determine the antecedents of anaphors inside semantic replacement units means that the length of units could be potentially as long as the distance between an anaphor and its antecedent. The problem of length becomes more serious when we consider Frazier's qualification that semantic replacement takes place only if the complete minimal governing category lies off the main projection path of a sentence. [2] Put in a reverse way, it means those governing categories on the main projection path might await the completion of a governing category for a prolonged time.

The second problem with Frazier's semantic replacement units is about how to conceptualize syntactic information. For example, while a transitive verb awaits its grammatical object it does not

---

2 The main projection path is defined by Frazier (1985: 176) as the maximal unique unbroken chain of {*Vn*, *Sn*} which includes the matrix VP.

undergo semantic replacement because it still remains within the semantic replacement unit. But it is questionable whether the information about transitivity can be retained in the form of the verb itself if the retention time is extended. The following translation examples show that only the morphological information on transitivity is lost.

(6) a. . . . will *enhance* the status, the power, and the responsibility ⌈of ⌈countries with relatively greater economic capability, most notably Japan ⌈and it will reduce the position ⌈of nations with primarily military power such as the Soviet Union. [3]

b. . . . juuyousei ga *takamattekuru* to omoimasu
. . . sekinin ga *mashiteiru* no desu
. . . chii mo sekinin mo *agatteiku* kuni ga arimasu
. . . yakuwari ga *takamaru* toiukoto de arimasu

The above example (6) is taken from Appendix 2 to Mizuno (1995a:19). We want to pay particular attention to the fact that all four interpreters under training use intransitive verbs in the output language of Japanese, as shown in (6b), while the original speech in English uses the transitive verb *enhance*. It is beside the question whether this result reflects intentional technique or not. To be noted is the grammatical relation preserved through the translation work : The relation between the transitive verb and its object in the source language can be identified with the relation between the

3 The mark " ⌈" in (6a) indicates where each of the four translation fragments in (6b) is reported to begin.

intransitive verb and its subject in the target language. This phenomenon tells us that the syntactic information about transitivity may be forgotten, as suggested by Frazier's idea of semantic replacement unit, yet the relational part of that information appears to be kept by the processor. What type of conceptualization is relevant in this case? It is neither purely semantic nor purely syntactic. We will come back to this issue later.

Although Frazier's concept could be applied to the problem of what constitutes the processing units in simultaneous interpreting, there are some shortcomings as we discussed above. We are going to explore the way to overcome these shortcomings in later sections.

3. A FOCUSSING ALGORITHM

Discussing the antecedents of pronominals and anaphors, Sidner (1983) makes an interesting proposal concerning movements of focus in a discourse. Movements of focus are exemplified in (7) and (8) in terms of what pronouns refer to.

(7) 1. I want to schedule *a meeting* with Harry, Willie and Edwina.
2. We can use *my office*.
3. *It* won't take very long,
4. so we could have *it* in the conference room.

(8) 1. I want to schedule *a meeting* with Harry, Willie and Edwina.
2. We can use *my office*.
3. *It*'s kind of small,

4. but *the meeting* won't last long anyway.

In the discourse (7) the meeting referred to in line 1 remains focused throughout the discourse, whereas in (8) the discourse focus begins on the meeting referred to in line 1 as is the case of (7) but it moves to *my office* because the pronoun *it* in line 3 refers to *my office* according to the way Sidner determines the focused element, and then comes back to the meeting at the end of (8). Sidner (1983) proposes an algorithm for the process from prediction to confirmation and then to modification with regard to the movements of discourse focus. What enables modification in this process is the potential focus list. For example, at the end of line 1 in (7) *Harry, Willie and Edwina* are all included in the potential focus list, though they are less expected to become the discourse focus than a *meeting.*

The potential focus list in Sidner's focusing algorithm suggests a possibility that this kind of list underlies language processing regardless of the occurrence of anaphors. In fact the basic concept of the cognitive files proposed here is similar to the potential focus list in that the structured record of entering pieces of information is maintained or otherwise renewed as the discourse develops.

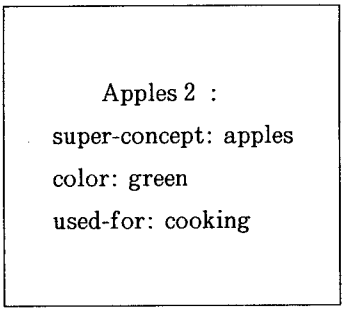
We pointed out in the previous section that the semantic replacement units in Frazier's conception have certain weakness because they are designed to follow the relationship between anaphors and their antecedents. Sidner's focusing algorithm keeps track of a potential focus list in addition to the actually focused elements and therefore the potential focus list could be less rigid as a representation unit in language comprehension. In the framework based on cognitive files a single page containes such a potential focus list

among other elements.

Sidner (1983) puts forth another interesting idea in grasping the relationship between anaphoric expressions and their antecedents. According to his theory, pronouns do not "refer" to other expressions nor "co-refer" to some objects together with their antecedents, but pronouns and their antecedents "co-specify" the identical cognitive elements. For example, *they* in (9) specifies the same cognitive element that *green apples* does. The cognitive elements to be co-specified are on the database and represented as in (10).

(9) I think *green apples* taste best and *they* make the best cooking apples too.

(10) cognitive representation:

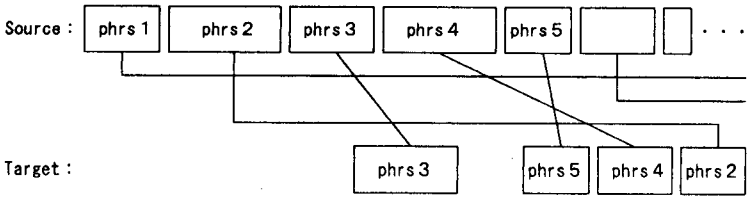


Sidner considers that to interpret sentences syntactically and semantically means specifying such cognitive elements. The kind of representation such as (10) enables us to represent those objects which do not exist in this world or some background information underlying linguistic expressions. The information to be written into cognitive files are all viewed as cognitive elements.

4. COGNITIVE FILES

Now, how can we represent the processing units in simultaneous interpreting with the help of the concept "cognitive file", which was characterized in the two preceding sections in a preliminary way? We will explore the idea further, comparing it with the conventional device which allows us to analyze the input and output of simultaneous interpretation along the time flow.

We begin with Fig. 1, which schematically shows the flow of both source and target languages in parallel with the phrase-by-phrase meaning equivalence indicated by connecting lines between the two layers, and thus it is here called "Two-layer Flow Diagram."

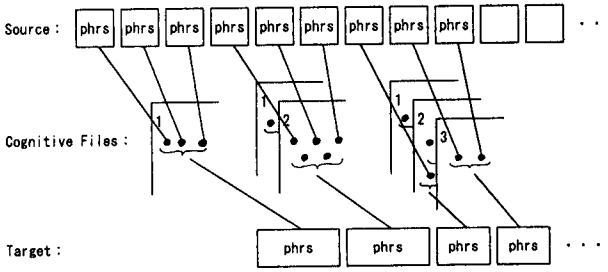


(Fig. 1) Two-layer Flow Diagram

Linguistic units flow in Fig. 1 from left to right temporally. In the case of simultaneous interpreting from English into Japanese, for example, the source language is English and "phrs 1", "phrs 2", • • • represent phrases in English. As was discussed in section 1, these phrases should be determined in accordance with the principles of language understanding, not translatability. The "phrs 3" , "phrs 5", • • • in the target language are meant to correspond to the phrases with the same number in the source language in terms of

translation. There is room, of course, for ingenuity to work in coupling the phrases in the two languages.

In contrast to the two-layer flow diagram, Fig. 2 has an additional layer called the cognitive files and thus dubbed "Multi-layer Flow Diagram."



(Fig. 2) Multi-layer Flow Diagram

This model incorporates an unobservable level, which is composed of cognitive files, between the two observable levels of the source and target languages and makes it possible to analyze the output one level before the final one. The phrasal correspondences between the source and the target are indicated by the lines via cognitive files in the diagram.

The cognitive files try to simulate how a processor-interpreter arranges the conceptual elements. In Fig. 2 a fragment with the number on the upper left corner represent a page in the cognitive file. As the original speech proceeds, new pages are created one over another. The cognitive elements arranged on a page are basically obtained from the input, though the processor's background knowledge may also provide for some cognitive elements on a page. Provisionally, the latter elements are shown in Fig. 2 as ones not

related to the input with lines. In terms of the Frazier model each cognitive page constitues a semantic replacement unit. The cognitive elements arranged on a page are now the sources of the target sentences. It requires supplementary efforts to make full use of the cognitive elements in producing good translation as we pointed out in section 1. In Fig. 2 the results of translation ingenuity are shown with the mark "⌒". In this way the source and the target are indirectly connected by way of the cognitive files in this model. The temporal overlapping of the source and target languages is shown here in the same way as in the two-layer flow diagram. Such an approach is abstract in nature in the sense that it involves unobservable theoretical constructs, yet demonstrative in that the final output can be examined by observable translations.

An example cognitive file is shown in (12), which is supposed to be created by some interpreter-processor as (s)he processes the input (11) [4]:

(11) . . . In light of the growing centrality of economic concerns, however, the United States will have to pursue and defend its economic interests more aggressively and systematically than in the past for overall foreign policy as well as purely economic reasons. . . .

4 The original speech is taken from Mizuno (1995a).

PAGE 1
BACKGROUND = CENTRALITY OF ECONOMY
: MORE

PAGE 2
U.S.
→ PURSUE, DEFEND [FUTURE; MUST]
→ ECONOMIC INTERESTS
: AGGRESSIVELY, SYSTEMATICALLY > PAST

PAGE 3
REASON = FOREIGN POLICY, ECONOMY

The cognitive file (12) keeps track of the cognitive elements which are gained in the on-line processing of (11), giving them structure page by page. The cognitive elements are represented by English phrases in (12) for convenience, but they are to be understood as concepts, not expressions themselves. The cognitive element "BACKGROUND" on page 1, for example, represents one possible way of conceptualizing the input phrase "in light of. " The bracketed indication "[FUTURE; MUST]" represents tense and modality. The symbols "=", " : ", "→", and " > " represent the cognitive elements which might be rendered "its content is", "modified", "part of argument structure ", and "comparatively small" respectively in ordinary language. In (12) the cognitive elements taken from (11) are divided into three groups, each of which composes a separate page, according to the provisional criterion that the thematic structure of a matrix predicate plays a central role in arranging the incoming cognitive elements. Although governing categories or pronoun-antecedent relations alone have difficulties in determining how cognitive elements are grouped into pages as we discussed in sections 2 and 3 and thematic structure seems a promising yardstick, the exact

formulation of page composition is yet to be studied furthermore.

The information represented in a cognitive file could serve as the basis for comparing different versions of translation. For example, one could put the Japanese equivalent of "PURSUE" before that of "ECONOMIC INTERESTS" in simultaneously translating (11) into Japanese or in the reversed order. The merit and demerit of such a choice should be judged dynamically because one choice affects other factors such as working memory burdens involving the following input. The cognitive file (12) also shows that the information on tense and modality such as "[FUTURE; MUST]" should be kept until the end of the translation of page 2, as would be demonstrated by the plural number of possible Japanese translations.

A processing unit in simultaneous interpreting in the framework of cognitive filing corresponds to each cognitive element and is finer than a generally recognized unit. The analysis based on cognitive files, therefore, is more detailed and makes it possible to incorporate many factors.

5. CONCLUDING REMARKS

A simultaneous interpreter is observed to process the original speech unit by unit. This is forced partly by the simultaneity of listening and speaking on the interpreter's part and the word order difference between the source and target languages. The principles of unit formation, however, should be considered to reflect the way humans process natural langauage. This is why we examined the current research on language processing in general in studying the processing units in simultaneous interpreting.

To postulate a hypothetical and unobservable level of representation, on the one hand, provides for a very powerful tool of analysis, but on the other hand it requires a lot of work to justify the hypothesis. There is a limit, however, to the analysis solely depending on the observable input and output. To be exact, the output units are not necessarily the same as the input units. For instance, an interpreter, in the process of producing translation, may postpone or leave out some units of information because (s)he has failed to come up with appropriate translation or because of time constraint, even though (s)he understands the input conceptually. It may also be said that the observation of translation results is not enough to identify processing units when the source and target languages agree in terms of the word order. Because of these factors the output and input units are not necessarily identical. Translations would only present indirect evidence to support the hypotheses on the process of interpreting.

In order to optimize the start of a translation fragment in simultaneous interpreting it is needed to know the accumulated information at each point of the process of accessing the input. Cognitive files could represent such information on-line, unit by unit. One of the future tasks on this line of research would be to work up a more complete theory on how to organize cognitive elements on a single page of the cognitive file.

REFERENCES

Chomsky, Noam (1981) *Lectures on Government and Binding,* Dordrecht: Foris.

Frazier, Lyn (1985) Syntactic Complexity. In: D. Dowty, L. Karttunen, and A. Zwicky (eds.) *Natural Language Parsing: Psychological, Computational, and Theoretical Perspectives,* Cambridge: Cambridge University Press. 129-189.

Frazier, Lyn and Janet D. Fodor (1978) The Sausage Machine: A New Two-Stage Parsing Model. *Cognition,* 6, 291-325.

Funayama, Chuta (1994) The Cognitive Files in Simultaneous Interpreting. *Current English Studies,* XXXIII, 1-12.

Mazuka, Reiko and Kenji Itoh (1995) Can Japanese Speakers Be Led Down the Garden Path? In: R. Mazuka and N. Nagai (eds.) *Japanese Sentence Processing.* Hillsdale: Lawrence Erlbaum Associates. 295-329.

Mizuno, Akira (1995a) The Dynamic Model of Simultaneous Interpretation. *Interpreting Research,* 8, 9-26.

Mizuno, Akira (1995b) On Simultaneous Interpretation from Japanese into English. *Interpreting Research,* 9, 4-21.

Sidner, Candace L. (1983) Focusing in the Comprehension of Definite Anaphora. In: M. Brody, and R. C. Berwick (eds.), *Computational Models of Discourse,* Cambridge, MA: The MIT Press. 267-330.