SEGMENTATION OF JAPANESE SOURCE LANGUAGE DISCOURSE IN SIMULTANEOUS INTERPRETATION

By

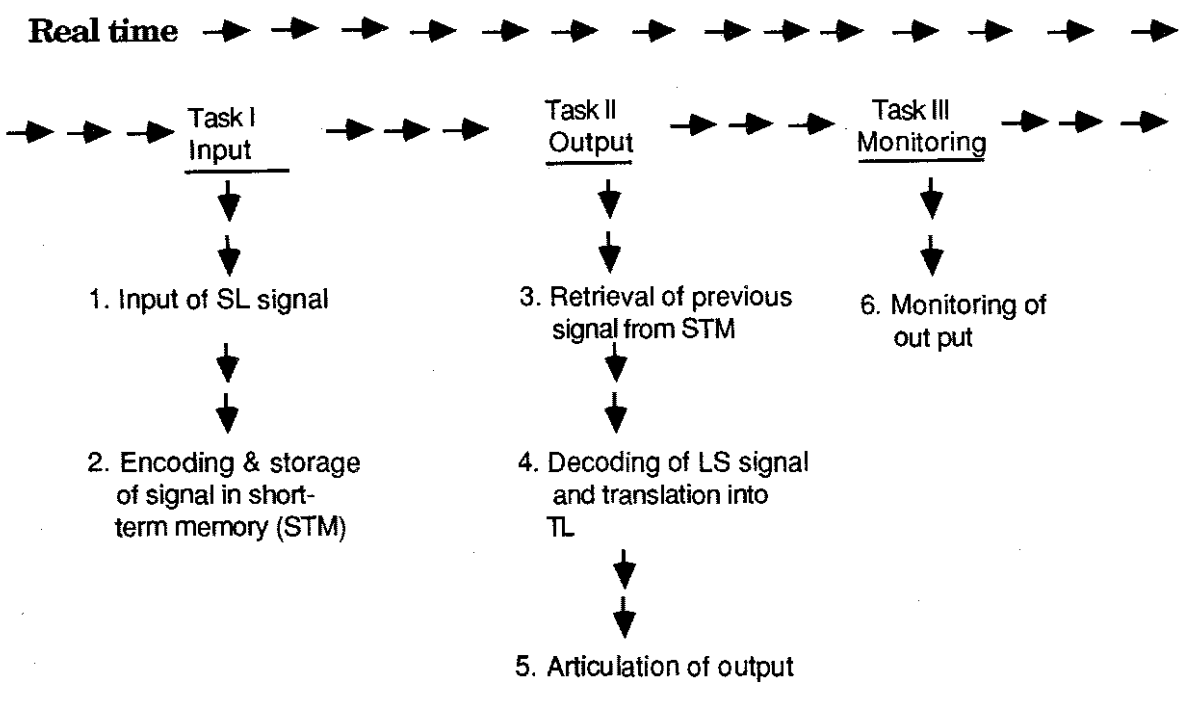
Peter M. Davidson

Department of Japanese & Chinese Studies
University of Queensland

"Some languages, such as Japanese, do not permit simultaneous interpreting, due to the complexity of their grammatical structure."
(Bower, 1959)

Introduction

Simultaneous interpretation is the process by which a message is conveyed into another language for the benefit of listeners who do not understand the first language. The interpretation takes place at virtually the same moment in time as the speaker is expressing his idea in the source language. While it is generally agreed that the process requires at least message reception, message decoding, target language encoding, and delivery into the target language (Barik, 1973, but see also Gerver, 1970), I would suggest that it could be seen in the following way:



There are some three distinct processes taking place simultaneously in real time. While the interpreter is inputting the next portion of the speaker's source language (SL) signal (Task I), he is in Task II outputting the previous portion into the target language (TL), and while he is undertaking these complex tasks, he is also simultaneously monitoring his output for inaccuracies and infelicities (Task III). Tasks I and II are further subdivided as shown above.

Barik raises some interesting questions in regard to this highly complex cognitive process. How is the interpreter able to engage in all of these activities simultaneously? Does the interpreter make use of certain strategies to effectuate some sort of input-output equilibrium? How does the interpreter segment the incoming message, and how does his speech pattern relate to that of the speaker? What difficulties, linguistic and otherwise, arise in simultaneous interpretation? How does the nature of the material being interpreted affect performance? Researchers have attempted to answer many of these questions within what has tended to be looked upon as a European language preserve: Asian languages have been largely ignored, despite the fact that Chinese is one of the six official languages of the United Nations, and that Japanese is the language of some 120 million people whose nation is an economic superpower.

This paper takes up the third of the questions posed by Barik, and examines the way in which student simultaneous interpreters, at different levels of competence, handle the segmentation of an incoming Japanese language message in preparation for interpretation into English.

Ear-Voice Span

Much interest has been shown by researchers in regard to the "time-lag" between the simultaneous interpreter's rendering in the target language and the original speaker's utterance in the source language (Barik, 1969; 1973; Gerver, 1970, 1971b, 1976; Goldman-Eisler, 1967, 1972; Kade & Cartellieri,1971; Lederer, 1978 and Seleskovitch, 1976, 1978). This is a phenomenon commonly referred to as *ear-voice span*. It can be measured by the number of words (or seconds) that the simultaneous interpreter lags behind, or occasionally, runs ahead of the speaker. It is not possible for any simultaneous interpreter (unless privy in advance to the text of the speaker) to commence interpreting at the same moment that the speaker starts to talk. Inevitably, the interpreter has to listen for a variable number of seconds to the speaker before being able to input, encode, store, retrieve, decode and articulate a segment of the speaker's ongoing speech (e.g. to complete Tasks I and II referred to in Diagram 1 above). From the listener's point of view, there will be a short delay after the speaker has started speaking (EVS) before the interpreter is able to commence articulation of the beginning of the speaker's speech, a pause of variable length (EVS), articulation of the next portion of the speaker's message, a variable pause again (EVS), articulation, EVS pause, and so on until the speaker has completed his speech. Unless the interpreter is particularly experienced and skilful, his interpretation will be delivered in "chunks".

In a classic study undertaken by Oléron and Nanpon in 1965, those writers found that the EVS of interpreters interpreting spoken passages of varying lengths into a number of different European languages ranged between 2-10 seconds. The EVS will naturally depend upon a number of variables - the level and complexity of the speaker's discourse, the interpreter's technical knowledge of the speaker's subject matter, the speed of the SL speech, the skills and professional experience of the interpreter, extraneous sounds etc. Measurements this writer has made in the classroom situation indicate a similar range, although students will occasionally lag further behind (mainly when they have failed to comprehend a segment, or when the speed of speech is too fast), and they will catch up only by discarding, uninterpreted, the segment or segments they either forgot or did not comprehend. This is supported by both Gerver's findings (1970) and those of Barik (1973).

Segmentation

While Barik (1969) and Gerver (1976) have both suggested that the speaker's natural pauses might delineate units of meaning for the interpreter which might then assist with segmentation of the almost continuous stream of the speaker's input, Lederer (1978) speaks of 'chunks of sense' appearing in an interpretation whenever the interpreter has a clear understanding of the speaker's intended meaning. She suggests that these 'units of meaning' are the synthesis of a number of words present in short-term memory associating with previous cognitive experiences or recollections.

Goldman-Eisler (1972) went further. She had noticed that *competent* interpreters working from German into English or French had longer EVS's than those working from French or English, and questioned whether this might be due to the fact that the predicate in German comes at the end of the proposition, with objects and other modifiers preceding it, which can be represented simply as:

S (Subject) + O (object) + V (verb)

as opposed to the simpler SVO framework of English and French. She suggested that it was only when the interpreter had decoded the predicate that he could start interpreting. Experiments she then conducted in these same languages indicated that postponement of the verb in a clause resulted in the interpreter electing to store more input, thus lengthening the segment. As she points out, this supports the results obtained by Healy & Miller (1970) who found the verb to be the 'main determinant of meaning'. The verb, they said, "*defines the plot; the subject merely indicates one of the actors*". This naturally has implications for Japanese.

Despite, then, the unfortunate comment made by Bower and quoted at the head of this paper (from his *International Manual of Linguists and Translators* published back in 1959), we all know that Japanese can be and is, interpreted simultaneously into English, and vice versa. If it could not be done, say, because of the delays in comprehension caused by the fact that the verb comes at the end of the predicate, it would have to mean that the Japanese themselves would equally not be able to converse with each other effectively.

How, then, do interpreters working from Japanese into English choose to segment the input from the speaker? What criteria do they consciously or unconsciously adopt? Do Goldman-Eisler's findings for German have relevance to Japanese? What strategies and techniques does this suggest for training simultaneous interpreters working in Japanese and English to be effective?

The Experiment

Two different groups of student interpreters were given the same Japanese speech to interpret simultaneously into English. Group A consisted of students towards the end of their first year of training, with very little formal training in simultaneous interpreting. Group B students were in the last six weeks of their second and final year. The experiment took place in timetabled practical classroom sessions of the Master's course in Japanese conference interpreting run by The University of Queensland.

This paper presents a preliminary analysis of the students' performance in a normal, everyday classroom situation. Students were not made aware that their performance was to provide the material for this experiment. Of the two groups, A and B, which represent students at two distinct stages in their training as conference interpreters, neither is as yet at the professional level. Group B had, however, received considerably more training and practice in simultaneous interpreting than had the novices of Group A, and consequently, Group B can be reasonably considered as being generally competent.

The level of discourse has considerable influence on the performance of the interpreter (see Bowen, 1970), and the "off-the-cuff" discourse selected for this experiment (a recording of a speech made to a luncheon meeting by a former Governor of the Bank of Japan) was of the type Bowen would describe as "spontaneous free speech", but which could

perhaps be better described as "impromptu" speech. *Impromptu* speeches generally prove to be easier for trainee interpreters to handle than the types of discourse found in prepared speeches. *Impromptu* speech is characterised by being less lexically dense (Halliday, 1985), although often more grammatically intricate, than written language. In Japanese they frequently tend to be poorly organised and somewhat fragmented.

The interpretations of the two groups, Group A and Group B, were transcribed from the tapes that had been used in the classroom teaching sessions over a period of several days. Each of the 99 sentences of the speech was transcribed so as to indicate where the student interpreter's segmentation took place. It should be borne in mind that the student interpretations are not necessarily always accurate. No attempt has been made to correct or modify them, however, since the sole purpose of the experiment was to analyse segmentation of incoming source language text, and not accuracy.

Analysis of the Results

As mentioned, the Japanese source language passage interpreted by the two groups consisted of 99 sentences of varying length and complexity. What immediately becomes apparent is that Group A (the novice group) broke the passage up into 535 segments, while Group B (the more competent group) needed to divide it into only 209 segments, less than half the number used by Group A. Expressed in another way, the average number of segments in Group A was *5.4* while for Group B it was only *2.1*. It must be remembered that the segment represents what the student interpreter considers to be the minimum amount of information necessary to start interpreting. While in Group A there were numerous occasions where the segment was lexical rather than syntactic, the generally more competent interpretation of Group B clearly indicates that segments tend to form syntactical entities.

Goldman-Eisler classified segments (which she here refers to as EVS units) into 7 distinct types. To suit the requirements of Japanese syntax, this experiment identifies six syntactical types, of which four have been retained from the Goldman-Eisler typology, two have been modified, and one added. They are listed in Table 1 below:

Type 1	Adverbial expressions or NP only (lexical)
Type 2	S + (O) + V (complete clause)
Type 3	(S1 + (O1) + V1 +) + (S2 + (O2) + V2)..... (two or more clauses)
Type 4	(S1 + (O1) or ..... (O1)..... (beginning/middle of clause)
Type 5	..... (V1) end of clause)
Type 6	(S1 + (O1) + V1) S2/ O2 (subject or object modified by clause)

Table 1

Each identified segment in the interpretations of Groups A and B were classified as one of the six types appearing in Table 1 above. The results appear in Table 2 on the following page. Whereas the novice Group A devoted almost 30% of their segmentation to Type 1 adverbial or noun phrases, the more competent Group B did so on only one occasion. Such phrases are lexical in nature, and novice interpreters, ever anxious to deal with the speaker's output as soon as possible after utterance, tend therefore to translate them immediately, before sufficient

information has been collected. The more competent interpreter waits to make a longer segment, at the end of which he may, or may not translate the adverbial phrase of conjunction that started the speaker's utterance.

What does become particularly clear from these figures is that the competent group (B) segmented virtually 90% of the speaker's discourse into complete clauses (63% single clauses and 27% as segments of two or more clauses), each ending with a verb. This strongly suggests, and reinforces Goldman-Eisler's findings for German, that interpreters working from Japanese into English, will more often need to input a clausal predicate before commencing interpretation than otherwise.

The segments of Group A were something of a smorgasbord. The fact that this group seized upon Type 1 segments so frequently (this type of segment was the most common) suggests that the temptation to provide an equivalent meaning on a lexical basis (as opposed to an idea-for-idea basis), wherever this could be done with ease, was irresistible. A similar, although not quite so lexical, temptation seems to have presented itself in the handling of longer clauses which were not infrequently split up into two or more parts (Types 4 and 5). Even so, Type 5 (end of clause) and Type 2 (complete clauses) both happen to be characterised by a final VP, and account for the second and third largest collection of segments (29% and 22% respectively).

Again this tendency towards shorter segments can be explained as a panic reaction by novice interpreters who are worried about not being able to retain longer segments in STM. Instead of waiting for the final verb to convey the full totality of the meaning, the only solution that they can see in their inexperience is to start interpreting sooner rather than later. This, in turn, leads to a variety of target-language problems with segmental connectives and links.

	GROUP A		GROUP B	
		%age		%age
1. Adverbial expressions or NP only (lexical)	157	20.39	1	0.48
2. Complete clause [S + (O) + V]	119	22.24	131	62.68
3. More than 1 clause [(S1 + (O1) + V1) + (S2 + (O2) + V2+ .....]	4	0.75	57	27.27
4. Beginning or middle of clause [S1 + (O1).....] or [..... (O1) .....]	63	11.78	4	1.91
5. End of clause [..... V1]	155	28.97	9	4.3
6. Noun clause [(S1 + (O1) + V1) S2 / O2]	37	6.92	7	3.35
TOTALS	535		209	

Table 2. Number of segments in text by type.

In Table 3 is shown the degree of segmentation per sentence. It will be immediately noticed that Group B interpreters saw no need to segment 34 of the 99 sentences in the selected passage, but were able to leave them intact. A further 41 sentences were split into only 2 segments. These 75 sentences account for about 76% of the whole speech. If one includes the 15 sentences split into 3 segments, then there only remain nine sentences (out of the total 99) needing greater segmentation (4-8 segments per sentence). Group A is very much the reverse. Over-segmentation is plainly apparent, with only 28% of the sentences in the passage being segmented into 3 segments or less. This offers further support for the comments made in the previous paragraph.

Number of segments into which Sentence split	GROUP A Sentences		GROUP B Sentences	
	No.	%age	No.	%age
1	3	3.03	34	34.34
2	14	14.14	41	41.41
3	11	11.11	15	15.15
4	11	11.11	4	4.04
5	22	22.22	5	2.02
6	11	11.11	1	1.01
7	10	10.10	1	1.01
8	5	5.05	1	1.01
9	3	3.03	0	0
10	4	4.04	0	0
11	2	2.02	0	0
15	1	1.01	0	0
16	1	1.01	0	0
17	1	1.01	0	0

Table 3. Number of segments per sentence.

The writer has suggested above that novice interpreters as exemplified by group A are worried about their ability to remember a segment of the incoming message if they allow it to become too long. They are worried that the inexorability of the incoming signal (which they need to store for later retrieval. Tasks I and II) will wipe out some of the earlier stored segments unless they can be dealt with more promptly.

Interpreters are required to interpret ideas, not words (Seleskovitch, 1977a, 1978b &c; Lederer, 1978 & 1981), and thus it is to the extraction of the ideaš contained in the incoming signal that they have to give their full attention. It thus becomes necessary to devise some simple method of gauging the length of these ideas, and the frequency of "elements"

(Goldman-Eisler's "EVS units") - a variable number of words that express a whole, but single, idea. Examples of "elements" that could be cited from the text of the interpreted transcripts would include "at about that time, "in regard to such points", "trade friction problems", "various other surplus counter-measures", "have been written", "desperately trying to prevent", etc.

	Groups A & B
Total number of sentences	99
Total number of elements (units of meaning)	1006
Average number of elements per sentence	10.1

Table 4. Number of elements in selected text.

	GROUP A		GROUP B	
		%age		%age
1. Adverbial expressions or NP only (lexical)	227	22.56	1	0.10
2. Complete clause [S + (O) + V]	271	26.94	574	57.06
3. More than 1 clause [(S1 + (O1) + V1) + (S2 + (O2) + V2+ .....]	15	1.49	353	35.09
4. Beginning or middle of clause [S1 + (O1).....] or [..... (O1) .....]	141	14.02	15	1.49
5. End of clause [..... V1]	249	24.75	29	2.88
6. Noun clause [(S1 + (O1) + V1) S2 / O2]	103	10.24	34	3.38
TOTALS	1006		1006	

Table 5. Number of elements in text by segment type.

Tables *4* and *5* show respectively the number of elements in the text, and the number of elements that were found in each of the six segment types for both Groups A and B.

Some useful information as to the average number of elements per segment type can be obtained by dividing the number of elements per segment type (Table 5) by the number of segments per segment type (Table 2). This information appears in Table 6.

The results indicate quite clearly that not only do the more competent simultaneous interpreters of Group B segment more than 90% of the incoming message by clause or clausal groups (types 2 and 3), but that the *elemental* content of their segments is considerably richer (*4.38* and *6.19* respectively) than that for the same segment types in the A group (*2.28* and *3.75*). Although it remains to be proved, this would suggest that there could be some relationship between length of segment (the longer the segment the greater the demands placed upon STM) and the number of elements contained therein (the richer the segment is in meaning, the easier it may be to retain in memory). Certainly a more accurate and felicitous interpretation emerges under these conditions.

Segment type	Elements per segment	
	GROUP A	GROUP B
1	1.45	1.00
2	2.28	4.38
3	3.75	6.19
4	2.24	3.75
5	1.61	3.22
6	2.78	4.86
Average	1.88	4.81

Table 6. Average number of elements per segment type.

Conclusions

The more experienced trainee interpreters of Group B showed considerable skill at segmenting the incoming message into larger syntactic chunks containing a predicate with a final verb. That they were able to handle such long segments would also seem to indicate that their STM capacity has grown somewhat greater than that of novice Group A. It would have been helpful to have had the services of a professional interpreter during this experiment, so as to have been able to measure both these two group against a professional yardstick, and this will be remedied in due course. At the moment, however, it does appear that while emphasis has always been placed in the simultaneous interpreting classroom on the need to extract and to work with the speaker's ideas, the student interpreters have consciously or, more likely, subconsciously discovered for themselves from the experience and confidence gained during their long period of intensive training that the Japanese ideas they isolate and work with while interpreting need to be "shaped" by taking cognisance of the structure of the Japanese language. With the final verbs brought into the segments, ideas in their totality can be interpreted. Since the ideas are complete, finding the appropriate linkages to succeeding segments becomes an easier process

for the interpreter, as well as being easier on the ear of the listener. The results thus coincide quite closely with the findings of Goldman-Eisler for German.

These findings naturally have ramifications for teaching methodology. Effective clausal segmentation obviously requires a high STM capacity. Although some time is already devoted to memory training (storage and retrieval) and simultaneous attention-sharing activities, it would be useful to see whether the provision of more training in these areas sooner in courses would bring about an earlier improvement in simultaneous interpreting segmentation skills and techniques.

Then, again, in the Japanese/English or English/Japanese simultaneous interpreting context, one needs to consider the speaker's pauses during delivery of speech - a factor ignored in the current experiment. Do they have any influence upon the way in which the interpreter "shapes" his segmentation process, and if so, how? Or again, on a slightly different track, to what extent is accuracy and felicity of interpretation dependent upon good segmentation techniques? What effect would other types of discourse have upon the simultaneous interpreter's ability to extract ideas? What strategies will the interpreter need to adopt when handling lexically dense, written-style speeches?

Certainly, the results of this preliminary experiment suggest many questions that remain to be answered, and many avenues that remain to be explored, and it is to be hoped that the answers will not be too long in coming forward.

Bibliography

Barik, Henri Charles. *A Study of Simultaneous Interpretation* (Unpublished PhD). Chapel Hill, University of North Carolina, 1969.

Barik, Henri Charles. "Simultaneous Interpretation: Temporal and Quantitative Data". *Language and Speech.* 1973, 16, 237-270.

Bowen, David & Margareta. *Steps to Consecutive Interpretation.* Washington, D.C., Pen & Booth, 1984.

Bower, William A. *International Manual of Linguists and Translators.* New York, Scarecrow Press, 1959.

Gerver, David. *Aspects of Simultaneous Interpretation and Human Information Processing* (Unpublished PhD). Oxford, University of Oxford, 1970.

Gerver, David. *Simultaneous and Consecutive Interpretation and Human Information Processing* (Report HR566/1). London, Social Science Research Council, 1972.

Gerver, David. "Empirical Studies of Simultaneous Interpretation: A Review and a Model". In Richard M. Brislin (ed.), *Translation: Applications and Research.* New York, Gardner Press, 1976, pp. 165-207.

Goldman-Eisler, Frieda. "Segmentation of Input in Simultaneous Translation". *Journal of Psycholinguistic Research.* 1972, 1-2, 127-140.

Halliday, M.A.K. *Spoken and Written Language.* Victoria, Deakin University Press, 1985.

Kade, D. & C. Cartellieri. "Some Methodological Aspects of Simultaneous Interpretation". *Babel.* 1971, 17, 12-16.

Lederer, Marianne. "Simultaneous Interpretation: Units of Meaning and Other Features". In Gerver, David & W.H. Sinaiko (eds.), *Language Interpretation and Communication.* New York, Plenum Press, 1978, pp. 323-332

Lederer, Marianne. "La Pédagogie de la Traduction Simultanée". In Jean Delisle (red.), *L'Enseignement de l'Interprétation et de la Traduction: de la Théorie à la Pédagogie.* Ottawa, University of Ottawa Press, 1981, pp. 75-98.

Oléron, Pierre & Hubert Nanpon. "Recherches sur la traduction simultanée". *Journal de Psychologie Normale et Pathologique.* 1965, 62, 73-94.

Seleskovitch, Danica. "Traduire: de l'expérience aux concepts". *Études de Linguistique Appliquée.* 1976, 24, 64-91.

Seleskovitch, Danica. "Why Interpreting is not Tantamount to Translating Languages". *The Incorporated Linguist.* 1977, 16-2, 27-33.

Seleskovitch, Danica. "Language and Cognition." In Gerver, David & W.H. Sinaiko (eds.), *Language Interpretation and Communication.* New York, Plenum Press, 1978, pp. 333-341.

Seleskovitch, Danica. "Traduction et mécanismes du langage". *Parallèles* (Université de Genève, Cahiers de l'École de Traduction et d'Interpretation), 1979, 2.