Note that the distinction between 'linguistic' and 'extralinguistic' anticipation is made for didactic purposes, because of the practical implications, just as the distinction was made between 'linguistic' and 'extralinguistic' knowledge (Chapter 4), but depending on the type of analysis performed, one could well challenge the existence of a fundamental difference between the two.

~

Basically, the more anticipation reduces uncertainty, the more cognitive relief can be expected from it. This is where extralinguistic knowledge and conference preparation (see Chapter 6) become important. By using documents and preparatory briefings in order to acquire knowledge about a conference, that is, about relevant facts, names, ideas, terms and expressions likely to be referred to or used respectively during the conference, interpreters increase their ability to anticipate and therefore decrease processing capacity requirements for their Listening and Analysis Effort, and sometimes for their Production Effort. More capacity is left for tasks which require it, and risks of saturation can be reduced.

6. An Effort Model of consecutive interpreting

The initial Effort Model was developed for simultaneous interpreting. Using the same principle, a similar Model was developed for consecutive interpreting. Note that this section deals with what AIIC members call 'true consecutive', in which the speakers' uninterrupted utterances are at least a few sentences long, as opposed to sentence-by-sentence consecutive in which there is no systematic note-taking.

Consecutive interpreting is performed in two phases, the comprehension phase (or listening and note-taking phase), and the speech production (or reformulation) phase.

Phase one: listening and note-taking

(8) Interpreting = L + N + M + C

- L Listening and Analysis
- N Note-taking
- M Short-term Memory operations
- C Coordination

During this phase, L is the same Listening and Analysis Effort as in the simultaneous mode, and M is similar to the Memory Effort in simultaneous interpreting. In simultaneous, it is associated with the time which elapses between the moment a speech segment is heard and the moment its content is reformulated in the target language, deliberately omitted or lost from memory. In consecutive, it is associated with the time between the moment it is heard and the moment it is written down (if it is written

down) or processed mentally and sent on to (long-term) memory. As to the Production Effort, during the first phase of consecutive interpreting, it is devoted to the production of notes, in contrast to the production of structured natural language as in simultaneous interpreting. The implications of this difference are discussed later in this section.

Phase two: target-speech production

(9) Interpreting = Rem + Read + P + C

Rem Remembering
Read Note-reading
P Production

In phase two, the Rem component is the set of mental operations devoted to recalling the successive parts of the original speech from long-term memory and is therefore different from the short-term M component. At first sight, phase two may seem more difficult than phase one, with its long-term memory (Rem) and note-reading (Read) Efforts. However, if notes are good, they help perform Rem operations and may actually reduce Rem processing capacity requirements rather than increase them. Interpreters occasionally mention the role visual memory plays in recalling the speech: indeed, when notes are taken according to a few simple layout rules (see Rozan 1956), the layout itself can be hypothesized to act as a visual stimulator of memory regarding the logical structure of the speech. Furthermore, while phase one is paced by the speaker, in phase two the interpreter is free to perform the three Efforts and allocate processing capacity to each at his/her own pace, which also reduces pressure on the Coordination component.

In the reformulation phase, unlike the situation in the comprehension phase or in simultaneous interpreting, the interpreter does not have to share processing capacity between tasks under high cognitive load, and for competent interpreters with adequate mastery of their working languages, there are no problems arising from an accumulation of tasks under the pressure of time resulting in capacity requirement peaks. This is why, in terms of processing capacity, only phase one seems to generate potential threats of saturation. Operational requirements are defined as inequalities (10) to (14) as they were for simultaneous interpreting through inequalities (3) to (7) in Section 5.1. In consecutive, they apply to the comprehension phase, not to the reformulation phase.

Similarly to the case of the simultaneous interpreting model, the following conditions must be met at all times in order for consecutive interpreting to proceed smoothly:

(10) $LR + NR + MR + CR \leq TA$

(11) $LR \leq LA$

(12) $NR \leq NA$

(13) $MR \leq MA$

(14) $CR \leq CA$

When inequality (10) is not true, saturation may occur. When inequalities (11), (12), (13) or (14) are not true, failure may result in spite of the possibility of total available capacity being larger than total requirements.

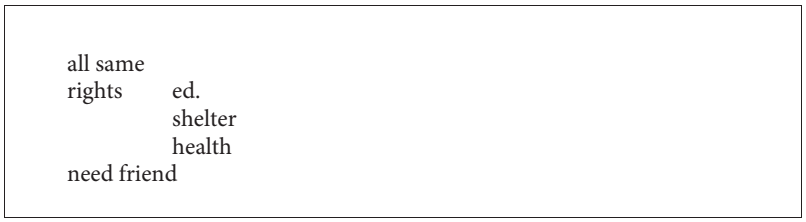
Consecutive interpreting, processing capacity and note-taking

The fact that in consecutive, speech comprehension and speech production are separated in time is a major difference which distinguishes it from simultaneous. It lowers markedly the pressure on target speech production as well as short-term memory load arising from syntactic differences between the source language and the target language: interpreters can take down information in notes as it arrives, as opposed to simultaneous, where they have to keep it in short-term memory until it can be put together and reformulated into a succession of natural-sounding target language sentences. The fact that during the second phase of consecutive interpreting, more capacity and time are available for speech production than in simultaneous may explain why some interpreters who refuse to work into B in simultaneous are willing to do so in consecutive.

As for the first phase, which is paced by the speaker, it does not necessarily generate the same constraints as simultaneous, because note-taking allows more freedom than speech production. Note-taking is not governed by rules of linguistic acceptability – lexical, syntactic, stylistic, or otherwise. As explained below, notes can be taken with much latitude. When processing capacity requirements for the Listening and Analysis Effort become high, some resources can be freed by reducing the quantity of notes being taken. In simultaneous, slowing down speech production results in some lag, which has a cost in Memory Effort load and may have to be made up later at the expense of increased capacity requirements for Production. In consecutive, reducing the amount of information that is written down in notes does not necessarily result in much increase in information to be stored in short-term memory and reformulated in the target language at a later stage. Below are a segment from a speech and an approximate reproduction of one professional interpreter's notes (from an experiment reported in Gile 2001):
Source-language speech

> "… because every child has the same needs, and the right to a basic education, the right to food, the right to shelter, and the right to basic health and every child needs a friend …"

Consecutive notes



As can be seen, words such as 'because', 'every child', 'needs' and 'basic', as well as grammatical words, were not written down by the interpreter, who nevertheless rendered faithfully the whole segment. This illustrates the latitude in note-taking mentioned above. Notes do not reproduce the speech; they are only written indications to help the interpreter remember it.

During the first phase of consecutive interpreting, problem triggers are similar to those found in simultaneous interpreting. The main difference between the two situations is associated with the manual nature of note-taking: it requires more time than speech production because hand movements are slow, and therefore lead to some lag, hence to a higher pressure on short-term memory, which may in turn reduce the capacity available for the Listening and Analysis Effort. This has implications. For instance, single-word names which are recognized by the interpreter need not be problematic in simultaneous, but if they need to be written down and are long, they can trigger problems in consecutive.

It could be argued that note-taking requires less time than speech production in simultaneous, because notes can take the form of single words, abbreviations, drawings and symbols, whereas speech production in simultaneous requires the construction of full sentences. Furthermore, as highlighted above, only part of the information is taken down while the rest is committed to (long-term) memory. This is true enough. However, for whatever information is written down, the time factor becomes important, especially when the interpreter does not have a readily available symbol or abbreviation for the information, and as shown by a number of empirical studies (see Section 5.3), lag at local level can impair the interpreter's ability to render the source speech successfully further down the road.

It follows that in terms of processing capacity, note-taking is critical, which explains and justifies the large volume of literature it has generated, from Rozan (1956) to Matyssek (1989) to a number of recent empirical studies (including Andres 2000; Gile 2001; Dam 2004a, b; Albl-Mikasa 2006 and several MA theses from China reported in various issues of the *CIRIN Bulletin* at www.cirinandgile.com). The critical nature of note-taking is the focus of a classroom experiment (Gile 1991a) which demonstrates that students who have not yet mastered its principles and techniques tend to miss more names (used as an indicator – see Gile 1984b) when interpreting after taking notes than when they do not take notes (see a description of the experiment in Section 12).

In the interpreting community, there are differing opinions as to the desirable quantity of notes, the language in which they should be taken, the use of symbols, etc. Note-taking is an area in which the concept of processing capacity can be useful but has only recently been used extensively in performance analysis. The basic question is how to reduce processing capacity and time requirements of note-taking while maintaining the efficiency of notes as memory reinforcers.

Viewed from this angle, symbols and abbreviations are an attractive tool, provided they have been fully mastered by the interpreter – otherwise, retrieving them from one's memory when they are needed for writing may take up too much time and processing capacity. The much-decried idea of learning a large set of symbols rather than making them up when needed may therefore have distinct advantages.

The same line of reasoning can be applied when considering whether it is preferable to take notes in the source language or in the target language. Some interpreting instructors recommend taking notes in the target language, arguing that this fosters analysis during the comprehension phase and does away with the need to 'translate' in the reformulation phrase. On the other hand:

- Thinking of target-language 'equivalents' in the source language while listening takes up extra processing capacity,
- This is done during the listening phase, which is critical in terms of processing capacity, and may therefore increase risks of saturation,
- The extra time and processing capacity required in order to 'translate' source-language notes during the reformulation phase do not jeopardize the interpreter's performance, since reformulation is self-paced, with little risk of saturation.

A reasonable alternative would be taking notes in the target language when cognitive pressure is not too high and reverting to source-language notes when close to saturation. In an experiment, Dam (2004a,b) found that students tended to take notes in their A language when the going got rough irrespective of whether it was the source language or target language, but these data which reflect the performance of students are not enough to generalize. Until enough empirical evidence is available, it is difficult to say which position is of more practical value, but instructors do seem to agree on the need to save both time and processing capacity in note-taking.

7. Efforts in sight translation

Sight translation is less frequent in conference interpreting than simultaneous or consecutive (but is very common in signed-language interpreting – according to a personal communication by Carol Patrie). It consists in 'reading' a source-language text aloud in the target language. It occurs when delegates receive a text and want to have it translated orally on the spot, or when a speech segment has been read from a text which is then handed over to the interpreter who is asked to translate it orally. It can be modelled as follows:

(15) Sight translation = Reading Effort + Memory Effort + Speech Production Effort + Coordination

In sight translation, what was the Listening and Analysis Effort in simultaneous becomes a Reading Effort, and the Production Effort remains, but while short-term memory is necessary to identify and understand speech segments as it is in simultaneous or consecutive, longer-term retention of words and clauses until they are enlisted for target-speech production is less of a problem because of the continued availability of the information in the text.

Pressure on short-term memory therefore seems lighter in sight translation than in simultaneous or consecutive (as it is assumed by cognitive scientists to be lighter when reading than when listening to spoken discourse – see Michael *et al*. 2001: 240). Moreover, sight translation is self-paced, not paced by the source-language speaker. It would therefore be tempting to consider cognitive load less heavy in this particular translation modality than in interpreting.

On the other hand, in the Reading Effort, sight translators are not helped by vocal indications such as the speaker's intonation, hesitations or other pauses which are found in simultaneous and consecutive interpreting and which help them segment the text into Translation Units. Moreover, in contrast to consecutive, while reading, they cannot devote all their efforts to understanding the meaning of the text but must think of its translation as well in order for their target-language rendition to be smooth enough. This may pose little difficulty when the two languages are syntactically similar and when the source text is written in a style easy to segment, in particular when sentences are short and made up of independent clauses. When sentences are long and/or include embedded clauses, it may be necessary for the sight-translator to read much more than one Translation Unit before reformulating it, which involves more time and more effort, both in the reading component and in short-term Memory during production. This occurs inter alia when translating from languages such as German or Japanese into French or English. The same difficulty arises in consecutive and simultaneous interpreting, but, as noted above, in sight translation there is no help for segmentation from the speaker's voice.

When the text is given to interpreters in advance, they can take care of this particular difficulty partly or fully by reading sentences and preparing their segmentation mentally before sight translation, or by inserting handwritten slashes in the appropriate places. It is meaningful that many interpreters seem to perform this operation before they do anything else, including writing down target-language equivalents of source-language terms. Sight translators may also write numbers above words in a segment which will have to be reordered in the target language, as in:

2	4	3	1
Interamerican	Tropical	Tuna	Commission

To be translated into French as:

Commission (1) *interaméricaine* (2) *du thon* (3) *tropical* (4)

When prior preparation is possible, Effort load can be reduced considerably.

Another major difficulty in sight translation is important in the initial training context: in interpreting, and especially in consecutive, the sounds of the source-language speech disappear rapidly from the interpreter's memory, permitting the reconstruction of the speech from its semantic content with little interference from source-language words and linguistic structures; in sight translation, these remain present before the practitioner's eyes throughout the operation (though the reader's gaze focuses only on a small text segment each time). This increases the risk of interference between the two languages and calls for more intense anti-interference efforts than in interpreting, making it a difficult exercise for beginners in spite of the facilitating factors described above.

8. Simultaneous interpreting with text

Simultaneous interpreting with text is a very common interpreting modality, inter alia in speeches at international conferences, when speakers read a text which has also been given to interpreters.

Simultaneous interpreting with text is associated with the following Efforts:

(16) Simultaneous Interpreting with text = Reading Effort + Listening Effort + Memory Effort + Production Effort + Coordination Effort

This combination of simultaneous interpreting and sight translation has the following features:

On the facilitating side:

- The existence of vocal indications from the speaker, though these may not be as helpful as in ad-libbed speeches, because pause and intonation patterns when reading are not the same as when planning and producing speech online in ad-libbed statements (see an interesting analysis in Déjean Le Féal 1978).
- The visual presence of all the information, which reduces memory problems and the effect of acoustic difficulties and unusual accents as well as the probability of failures due to insufficient processing capacity in the Listening and Analysis Effort. In other words, the Reading Effort and Listening Effort cooperate to a large extent – but they also compete, as explained below.

On the negative side:

- The high density and peculiar linguistic constructions of written texts as opposed to oral discourse, which require more processing capacity in the analysis component (see Halliday 1985; Brown & Yule 1983).
- Increased risks of linguistic interference, as explained above for sight translation.