over the past decades. It is considered an important determinant of cognitive operations and is now sometimes referred to as 'working memory' (though 'short-term memory' is a fairly generic concept while 'working memory' is more specific). In 1974, Baddeley and Hitch developed a model of working memory with a specific structure and operational components, including a 'Central Executive', a modality-free cognitive mechanism which coordinates the operation of the other entities in the model, namely a 'phonological loop' which holds information in phonological form and a 'visuo-spatial sketch pad' specialized in spatial and/or visual information coding. Baddeley and Hitch's model is described in most introductory books on cognitive psychology – see for example Eysenck and Keane 1990. Further research in the field naturally led to tests of its ability to explain and predict cognitive phenomena and then to other models with further ideas about its components, including specialized verbal working memory (Caplan & Waters 1998), about its operation, about is relationship to long-term memory. According to Miyake and Shah (1999), several ideas and theories about the components and operation of working memory compete in cognitive psychology circles (for a recent review, see Timarová 2008), and some authors even doubt the usefulness of the concepts of working memory as a separate entity, but a consensus can be found with respect to the following points:

1. Working memory is a set of mechanisms or processes involved in the control, regulation and active maintenance of task-relevant information in the service of complex cognition; it operates primarily on currently 'activated' information from long-term memory.
2. Working memory requires processing capacity
3. Working memory has a small storage capacity

As explained in more detail in Chapter 9, working memory is necessarily part of the language-comprehension process and of the speech-production process. It is obviously part of the Memory Effort and perhaps conceptually very close to it, but the Memory Effort is explained here in such a way as to be intuitively recognizable by students and professional interpreters as a step in the interpreting process which involves memory and memory operations, not as a conceptual entity from cognitive psychology. Readers may consider the Memory Effort as corresponding to working memory if they wish, but for reasons explained in Section 13, I prefer to talk about short-term memory and about the Memory Effort when referring to the mechanisms of interpreting in general and to invoke working memory only in more technical considerations.

4. An Effort Model of simultaneous interpreting

4.1 A first view of the model

Using these definitions, simultaneous interpreting (SI) can be modelled as a process consisting of the three core Efforts described above, namely the Listening and Analysis

Effort L, the Short-term memory Effort M and the Speech production Effort P, plus a Coordination Effort C which corresponds to resources required to coordinate the three other Efforts (Eysenck & Keane 1990):

(1) SI = L + P + M + C

(In this formula, the 'equal' sign should be interpreted as meaning 'consists of', not as an equality in the usual mathematical sense, and the 'plus' sign as some kind of 'addition' in a very general sense, not as the usual arithmetic addition).

In a somewhat oversimplified form (see below), this model depicts simultaneous interpreting as a process which involves a set of operations on successive speech segments. Each of them is heard and analyzed (L), then stored in memory for a short while (M), and finally reformulated in the target language (P).

4.2 Simplifications in the model

Before going into further analysis of the Model, two major simplifications which it incorporates need to be highlighted.

4.2.1 *The sequential linearity simplification*

It is convenient to think of the Listening, Production and Memory Efforts as handling sequentially Translation Units or 'speech segments' (or 'chunks' as they would be called in the psycholinguistics literature), which can vary in length from one to several words forming a clause or even a sentence (see the discussion of Translation Units in written translation in Chapter 5), in the order in which they were uttered by the speaker: if the source speech consists of successive segments A, B, C, D, E, F etc., Production could focus on segment A while segment B has been analyzed and is waiting in short-term memory for its turn to be reformulated, and segment C is being analyzed by the Listening and Analysis Effort. Generally, with the exception of anticipated segments, source-speech segments can be reproduced in the Target Text (Effort P) only after they have been understood (Effort L).

Reality is more complex, if only because of syntactic differences between the source- and target language which naturally lead to information-order changes in the target speech. Other phenomena can lead to similar results. For instance, when the initial words in a speaker's sentence are not clear, the interpreter may need to keep them in memory until they are well understood. By that time, more than one Translation Unit is stored in short-term memory and it is not clear which will be rendered first. Finally, linguistic and semantic anticipation as alluded to earlier occur frequently in speech comprehension, and interpreters sometimes find themselves voicing in their target speech ideas or information which the speaker has not expressed verbally yet, at least not fully.

Nevertheless, the linearity assumption remains a useful simplification for the purpose of explaining strategies and tactics in the daily practice of interpreting, as discussed in Chapter 8.