Cognitive studies in simultaneous interpreting
Christoffels, I.K.

Citation for published version (APA):

General rights
It is not permitted to download or to forward/distribute the text or part of it without the consent of the author(s) and/or copyright holder(s), other than for strictly personal, individual use, unless the work is under an open content license (like Creative Commons).

Disclaimer/Complaints regulations
If you believe that digital publication of certain material infringes any of your rights or (privacy) interests, please let the Library know, stating your reasons. In case of a legitimate complaint, the Library will make the material inaccessible and/or remove it from the website. Please Ask the Library: http://uba.uva.nl/en/contact, or a letter to: Library of the University of Amsterdam, Secretariat, Singel 425, 1012 WP Amsterdam, The Netherlands. You will be contacted as soon as possible.
Summary and Discussion

8.1 Introduction

In this dissertation three main issues concerning simultaneous interpretation were addressed experimentally: Which task components are important in determining task complexity, which skills are relevant for successful SI, and how does retention differ in SI from retention in other tasks. In this final chapter, some conclusions drawn in Chapter 2 will be mentioned first. Then, the main findings of this dissertation will be discussed, organized around the above three issues: components, skills, and retention in SI. Finally, some general remarks concerning the merits and difficulties of SI research will be made.

The review in Chapter 2 makes clear that although more research is needed in order to have confidence in the conclusions on some of the topics discussed, a number of interesting findings and issues emerged, some of which will be highlighted next.

First of all, SI is characterized by simultaneous comprehension and production. Real simultaneity of listening and speaking, taking pauses in speech into account, occurs about 70% of the time (Chernov, 1994). Another characteristic concerns the time lag, or ear-voice span, between input and output. The average time lag appears to be 4 to 5 words, which is similar to what we found in Chapter 3 (Gerver, 1976; Goldman-Eisler, 1972; Treisman, 1965). Finally, there is some evidence to indicate that the processing unit in SI is the clause.

An important aspect of SI processing is how the process of recoding a message from the source to the target language may be best described. No strong conclusion is warranted but two possible translation strategies, or ‘routes’ through the language system(s), are discussed: meaning-based interpreting and transcribing (e.g., Anderson, 1994; De Groot, 1997; Massaro & Shlesinger, 1997; Paradis, 1994). It is often assumed that these strategies coexist. According to the notion of meaning-based interpreting, source input is first fully comprehended and represented in a non-verbal manner before reformulation and production in the target language takes place. In the more extreme version of this account the wording of the input is assumed to be completely discarded even before the output is produced (i.e., deverbalization). According to the notion of transcribing, interpreting proceeds through the transposition of words or multi-word units. In Chapter 2 we argued that these strategies may be difficult to separate experimentally, but there is some evidence that transcribing takes place, and little evidence to suggest that the input is completely deverbalized (Daró & Fabbro, 1994; Isham, 1994).
One factor that clearly influences SI performance is the listening condition. Interpreting output is influenced by the speaking rate, the information density, and the sound quality of the input (Gerver, 1974a, 1976; Treisman, 1965). Direction of translation, whether to interpret from or into the L1, is often considered an important determinant of SI performance. The sparse experimental data suggest, however, that the specific language combination involved may be more relevant for SI performance than whether those languages are the interpreter's L1 or L2. Finally, characteristics of the source text are likely to influence SI performance as well, but as yet, it is not clear which characteristics are most important. Next, the main findings of the subsequent chapters will be discussed; some relevant conclusions of Chapter 2 will be mentioned where appropriate.

8.2 Components

The question in Chapter 3 was which components of the SI task are mainly responsible for its cognitively demanding nature. In particular, we assessed the relative importance of the simultaneity of comprehension and production, on the one hand, and of transformation, on the other. The latter component entails the reformulation of discourse from the source into the target language. At least theoretically, this component consists of two distinguishable aspects: reformulation, and switching language. We tried to distinguish the effects of each of these components on performance by comparing shadowing, interpreting, and paraphrasing in a simultaneous condition and in a delayed condition. Shadowing involves simultaneity of input and output, but no transformation of the input is involved. In contrast, in both paraphrasing and interpreting reformulation of the input required, but only the latter also involves a language switch. The latter comparison was therefore included to address the question whether the demands of reformulation and of language-switching can be distinguished.

Using different measures of online performance (i.e., the quality and amount of output, and the ear-voice span) we drew the following conclusions. First, the paraphrasing task failed to live up to expectations: Paraphrasing performance was as least as difficult as SI, and could therefore not be used to distinguish reformulation from language-switching in the transformation component.

More importantly, we did establish that both the simultaneity of comprehension and production and the transformation component are sources of cognitive complexity in SI, but we also found that the role of each of these components separately is limited. Excellent simultaneous shadowing performance suggested that the participants handle simultaneous comprehension and production of language quite well. Offline interpreting performance suggested that transformation is the more demanding component of the two. However, it is especially the combined demands of simultaneity and transformation that make simultaneous interpreting the complex task that it is.
As mentioned earlier, the paraphrasing task was not suitable for comparison as originally intended in the design. One may wonder why we reported on this task at all. The reason was that paraphrasing makes an interesting comparison to SI, if not for the intended reason. After all, the task is referred to as within-language interpreting, has previously been used in research, and is used in the training of SI (Anderson, 1994; Green et al., 1994; Malakoff & Hakuta, 1991; Moser-Mercer, 1994). However, we had some concerns on its comparability with SI and we therefore wanted to test it in an experimental setting. The conclusion is clear: The paraphrasing task is interesting in its own right but should not be considered the monolingual equivalent of SI.

To conclude, in Chapter 2 we established that interpreting is a more complex and demanding task than shadowing is. The results of Chapter 3 supported this conclusion. However, they also suggest that this complexity is not due solely to the transformation component that is present in SI but not in shadowing. Again, only the combination of simultaneity and transformation seems to exhaust the limited resources.

### 8.3 Skills

In Chapter 2 a number of cognitive skills are reviewed that appear to be relevant for professional interpreting. There is some evidence to suggest that interpreters excel in verbal fluency and other basic language skills (e.g., Bajo et al., 2000; Fabbro & Darò, 1995). The studies reported on in Chapters 4 and 5 and summarized below suggest that memory skills are very important for SI. In these chapters we focused on the roles of memory and lexical retrieval. One straightforward conclusion that follows from both chapters is that it is indeed possible to identify cognitive subskills in SI, and that, therefore, this line of research deserves further attention.

In Chapter 4, we assessed individual differences in lexical retrieval, memory, and SI in participants who had no previous experience in SI. We administered a reading span task in two languages and a verbal digit span in the native language to assess memory capacity. Furthermore, picture naming and word translation tasks were administered to determine the retrieval time of lexical items in two languages. The results showed that word translation and picture naming latencies correlate with interpreting performance. Also digit span and reading span were associated with SI performance, only less strongly so. In other words, we found evidence that better lexical retrieval and working memory performance are related to better SI performance, which suggests that the former are relevant cognitive skills for SI in inexperienced bilinguals.

Furthermore, a graphical model was derived from the data to describe the relations between the different variables. From the model two main observations were made. First, in the model only reading span English and translation Dutch-English had a direct effect on SI performance. The effects of all other variables were mediated by those two. In other words, of the subskills we studied, the efficiency of working memory and the retrieval of translation
equivalents were the most relevant. Second, there were two (conditionally) independent paths to SI performance, involving working memory and word translation skills. These subskills therefore contributed to SI performance separately, which means that word translation efficiency and working memory each form independent subskills of SI in untrained bilinguals.

In Chapter 5 we examined differences between groups of participants with different backgrounds and experience in SI and second language use. As in Chapter 4, the picture naming task and word translation task were administered in two languages. Furthermore, word span, reading span, and speech span tasks were administered in two languages to indicate memory skill. Finally, two control tasks measuring vocabulary and basic reaction time were presented.

Two experiments were conducted in which different groups of participants were tested. In the first experiment, we compared performance of professional interpreters and bilingual university students. It was found that on all tasks professional interpreters showed better performance than the bilinguals. In accordance with the findings of Chapter 4, we concluded that lexical retrieval and working memory are relevant subskills of SI.

Although the superior performance of the interpreters holds as well for a number of LI tasks (except for picture naming) it was deemed possible that different L2 language proficiency in interpreters and bilinguals, rather than SI experience, accounted for (some of) the differences in performance between groups.

To assess the role of language proficiency on task performance we compared in the second experiment professional interpreters with English teachers. The latter group is more similar to the interpreters in language background and proficiency than the university students are. The comparability of the two groups was supported by their similar performance on the control tasks. We found that the teachers and the interpreters could not be distinguished from each other on the lexical retrieval tasks. This qualified the conclusions on the role of lexical retrieval in SI, as based on Experiment 1 and the results of Chapter 4. We concluded therefore that lexical retrieval skill is not uniquely related to SI experience. Although it probably is a prerequisite for successful SI, it appears that this subskill is not ‘boosted’ beyond the level obtained by the teachers, who professionally also need a high proficiency in the L2. This finding illustrates that, when searching for specific subskills possessed by professional interpreters, it is important to consider whom they should be compared to.

The results for the memory tasks were quite different. We found not only that interpreters in general had a higher (working) memory capacity than the other two groups, but also that their capacity was the same for the L1 and the L2. In contrast, the bilinguals and the teachers performed similarly to each other, and were better at the L1 version than at the L2 version of these tasks. We concluded that professional interpreting experience is associated with elevated verbal memory capacity, and this is therefore a relevant subskill of SI.
In sum, the main findings of Chapter 5 are that working memory is an important subskill for SI whereas the role of lexical retrieval in SI appears to be mediated by general language proficiency.

Finally, in Chapter 6 (for details see the next section), some evidence was obtained for another relevant subskill of SI. In Experiment 2 of this chapter, we related individual differences in recall under conditions of (meaningful) articulatory suppression to SI performance. We found that retention under these conditions is associated to SI performance, suggesting that the ability to retain information while being engaged in articulation during listening might be a relevant subskill of SI. The results are discussed in terms of the working memory account including the recently added episodic buffer component (Baddeley, 2000, 2003). It is suggested that individuals may differ with respect to how efficient they are in building an episodic structure of the input. If this happens quickly the lack of articulatory rehearsal in the phonological loop is less problematic. This seems to be advantageous in SI.

8.4 Recall

In Chapters 3, 6, and 7, the issue of how retention of input is affected by the concurrent articulation task at hand is explored. According to the working memory account of Baddeley and colleagues (e.g., Baddeley & Logie, 1999) articulatory suppression (AS), the concurrent articulation of irrelevant sounds, prevents rehearsal in the articulatory loop, which, in turn, has a disruptive effect on recall. Usually in this paradigm recall of lists of unrelated words is measured. Still, the effect of concurrent articulation on recall may be one reason why recall following simultaneous interpreting of a text tends to be poorer than following listening to the same text (Daro & Fabbro, 1994; Gerver, 1974a; Isham, 2000; Lambert, 1988) because SI is a task where people routinely comprehend and maintain speech while articulating at the same time.

In Chapter 3 (for details see earlier section), we reported on recall following shadowing, paraphrasing, and interpreting. In contrast to some previous studies (Daro & Fabbro, 1994; Gerver, 1974b; Lambert, 1988), we found no effect of task on recall in the simultaneous condition. General lower recall in the simultaneous condition than in the delayed condition suggested that articulation of speech indeed interfered with memory in SI.

In Chapter 6, the question was asked whether AS also affects retention when, as in interpreting, coherent text is to be remembered or meaningful material is articulated (Experiment 1). In three conditions, subjects just listened to a set of stories (no suppression), articulated 'de, de, de' ('traditional' AS), or articulated 'hond, kat, muis' (dog, cat, mouse; meaningful AS) continuously. In each condition, one story was coherent while another was made incoherent by randomizing the sentences. We found that AS affected retention of stories, that incoherence influenced recall negatively, and that AS and coherence interacted.
We concluded that the effect of AS generalizes to stories and that its effect is stronger for coherent stories.

In Chapter 7, we asked again how recall following interpreting compares to recall following shadowing. A reason to conduct this study was that previous studies provided a mixed answer to this question (Darò & Fabbro, 1994; Gerver, 1974b; Lambert, 1988). Moreover, no differences in recall between tasks were obtained in the simultaneous conditions in Chapter 3, whereas in Chapter 6 the type of articulatory suppression clearly did affect recall. The goal of this study was to replicate the findings on recall of Chapter 3 and to include articulatory suppression as well, in a study that focused on measuring retention of input. Furthermore, to gain more insight in the factors that contribute to the different results in previous research we assessed the role of language of testing (native and second language) and used two different recall measures, concerning recall of gist (semantic recall) and exact wording of the input (form recall). Moreover, apart from AS, we included a random letter generation condition. These two conditions both involve concurrent articulation, but are known to differ in the amount of central resources they consume. In general, the results for shadowing and interpreting were similar, replicating the results reported in Chapter 3, although type of recall did influence the pattern of results. There was no main effect of language of testing in recall in both semantic and form recall but in form recall language did interact with listening condition. Interestingly, recall following AS was not better than following shadowing and interpreting. A likely explanation for the latter finding is that the articulation in the SI and shadowing conditions involves processing of the input rather than generating irrelevant output, as in AS and random letter generation. This processing is likely to support the building of an episodic structure in the episodic buffer of working memory and compensates for the detrimental effects of AS.

8.5 Final comments

In this final section, we will comment upon a few issues concerning the choice of participants, the ability of novices to perform SI, the need to make theories on language processing and memory compatible with SI, and modeling of the SI process itself.

As discussed in Chapter 2, one of the problems frustrating research in SI is that professional interpreters are not readily available for participation in research. For this reason, in this dissertation in most experiments bilinguals were tested who had no previous experience in simultaneous interpreting. The question (also raised in Chapter 4) is whether conclusions based on studies that do not involve professionals provide valid information on professional interpreting. Since little research is available on this matter, we cannot be sure that findings with novices in SI can always be generalized to professional interpreters. However, there is actually an advantage in using, as we did, novices: SI can be investigated without ‘contamination’ of the idiosyncratic strategies that interpreters may have developed (e.g., Shlesinger, 1994). The question whether and to what extent the interpreting process is
qualitatively different in professionals and novices is a matter of empirical research. The difference in memory performance between interpreters and other groups of participants that we obtained in Chapter 5 suggests that the interpreters themselves are special in at least one respect: they possess exceptional memory skills.

Although the majority of participants tested in this dissertation (the students and the teachers) were novices in SI we asked them to interpret simultaneously (in Chapter 3, 4 and 7). It is noteworthy that all of them, even the participants who were least proficient in the L2 (i.e., the students), produced at least some output and that some of them produced quite intelligible translations. Of course, their performance was not comparable to any professional standard. The point is, however, that at a more modest level SI seems to be within reach of reasonably proficient bilinguals without ever having been trained in this skill.

Consequently, SI can not simply be (dis)regarded as a very specific phenomenon tied to a small group of professionals who may be idiosyncratic in many respects; who may process language differently from other bilinguals and whose language processing abilities therefore have nothing to do with the normal language system. Instead, it must be concluded that models of bilingualism, memory, attention, and of components of language processing such as language comprehension, language production, and discourse processing, should be compatible with the fact that people are capable of SI.

Nevertheless, existing models do not always easily accommodate SI. This is illustrated in the section on language monitoring in Chapter 2. Examples of self-correction in speech production by interpreters warrant the conclusion that they monitor their speech output in addition to comprehending the input and producing the output (Gerver, 1976). Monitoring in SI is not easy to reconcile with the idea that language monitoring proceeds by engaging the language comprehension system (Levelt, 1989), because this entails that in SI the comprehension system needs to process two speech streams simultaneously: the input in the source language and the output in the target language. Another example is the theoretical explanation of control of languages. It is often assumed that the relative activation of the two languages in general (i.e., by inhibition of the nontarget language or activation of the target language) supports processing in the target language rather than the nontarget language. Such a global activation/inhibition account of language control is not easily reconciled with the fact that in SI two languages are in use simultaneously. This issue is discussed in Chapter 2, where three alternative proposals are discussed that allow correct lexical selection even though both source and target language are simultaneously activated. In conclusion, the fact that most reasonably proficient bilinguals do possess the ability to perform SI can constrain models in psycholinguistics.

The study of SI is not only interesting because it may inform theories of ordinary language processing, we may also want to understand the processing involved in SI itself. To model SI, an integrated account is necessary of the relation between language comprehension and production, the relation between and the control of different languages in a multi-lingual person and the relation between language and working memory. However,
language comprehension and language production, for example, concern separate research fields (Frauenfelder & Schriefers, 1997). Questions concerning the relation between processing in language comprehension and production are unanswered as yet (Schiller & Meyer, 2003), making the integration of models within those fields problematic. In fact, the proposals on lexical selection presented in Chapter 2 mentioned earlier, differ even on the basic question of whether the mental lexicon used for comprehension and production is separated for each function or not. Also, within each research field competing models exist that differ in their basic assumptions on architecture and processing. For example, accounts of working memory differ in whether resources are shared or separate between storage and processing functions, whether working memory is a unitary construct or consists of multiple components that have specialized functions, and whether memory capacity is domain general or specific (e.g., see Bayliss et al., 2003; Duff & Logie, 2001; Just & Carpenter, 1992; Miyake, 2001; Miyake & Shah, 1999b for discussion). In Chapter 2 the issue of mental resources is briefly touched upon in the context of SI, raising the question which subcomponents of the full SI task need to be distinguished and whether or not resources are shared between components.

These examples illustrate that the answers to fundamental questions concerning bilingual language processing still need to be asked or are a matter of debate. These questions have to be addressed in any model of the full SI process. On the positive side, more comprehensive models within and between research domains are being developed (e.g., Miyake & Shah, 1999a; Schiller & Meyer, 2003). For example, the bilingual interactive activation (BIA) model (Dijkstra & Van Heuven, 1998) of bilingual word recognition is recently extended to the BIA+ model (Dijkstra & Van Heuven, 2002). It now includes, in addition to orthographic representations, phonological as well as semantic representations, and a task/decision system analogous to the inhibitory control model for language control (Green, 1986; 1998, see Chapter 2). In the domain of working memory, the additional component of the episodic buffer, makes this account more similar to the long-term working memory account of Ericsson and colleagues (Ericsson & Delaney, 1999; Ericsson & Kintsch, 1995) and extends the range of data that can be accounted for.

To conclude, in this dissertation a number of studies are reported that together reveal some of the intricacies of simultaneous interpreting. It is hoped that future research will provide enough consensus on the basic architecture and processing in the bilingual mind that an integrated account of the SI process can be developed.