Session II

'Going simul?' Technology-Assisted Consecutive Interpreting

Franz Pöchhacker
University of Vienna

Franz Pöchhacker is Associate Professor of Interpreting Studies at the Center of Translation Studies in Vienna, Austria. Having trained as a conference interpreter at the University of Vienna and the Monterey Institute of International Studies, he has worked freelance in conference and media settings since the late 1980s and has published articles and monographs on various topics of interpreting research. He is the author of Introducing Interpreting Studies (Routledge, 2004) and co-editor, with Miriam Shlesinger, of The Interpreting Studies Reader (Routledge, 2002) and of Interpreting: International Journal of Research and Practice in Interpreting.

Introduction

Consecutive interpreting with notes, as developed for extended stretches of discourse by the first generation of conference interpreters at the League of Nations, was crucial to the professionalization of interpreting in the twentieth century but has long lost its predominance to the technology-assisted simul¬taneous mode. With the availability of high-performance digital recording devices for mobile use, the classic technique of consecutive interpreting with notes may face yet another fundamental challenge from the simultaneous mode. The new interpreting technique discussed in this paper is a form of consecutive interpreting in which storage of the original message in the interpreter's long-term memory and notes is replaced by a digital recording of the original speech, which the interpreter then plays back into earphones and renders in the simultaneous mode.

The paper describes the origins of this innovation and touches on some of the associated conceptual issues. Based on some of the claims made for this technique by its pioneering users, I will consider various practical implications and research questions raised by this interpreting mode and briefly describe an ongoing small-scale study conducted at the Center of Translation Studies of the University of Vienna.

A new mode of interpreting

The idea for this innovative mode of interpreting is claimed by several professional interpreters. Its first successful use in an authentic assignment, however, is clearly credited to SCIC interpreter Michele Ferrari, who reported in late 2001 that he had used his PDA at a high-level press conference in Rome in 1999 to record the speaker's statements for a subsequent simultaneous rendition which he termed "consecutive simultaneous". Further discussion and tests with different recording equipment and interpreters in the context of SCIC followed, but the new technique, also referred to as "simultaneous consecutive" and "digitally (re)mastered consecutive" by its inventor, appears to have met with surprisingly little attention and acceptance.

Roughly in parallel, two court interpreters in the US proposed using a digital voice recorder for the (consecutive simultaneous) rendition of witness statements in court. Their technique, labeled "digital (voice) recorder assisted consecutive", became viable in late 2002 and was tested in comparison to traditional consecutive with notes in an experimental study by Erik Camayd-Freixas at Florida International University.

Claims and research issues

Based on trials by its pioneers and the few tests that have been reported, a variety of claims have been made for this hybrid mode of interpreting, described even as 'the ideal interpretation' because it combines the best of both worlds (prior audition of the entire original, as in consecutive interpreting, and complete reproduction of the original including paralinguistic features, as in the simultaneous mode).

The essential advantage claimed for 'simultaneous consecutive' (here referred to as 'SimConsec') is its 'accuracy' – the complete rendition of content even for fast and dense statements and speeches, regardless of their length. This superior faithfulness is said to apply also to delivery features such as rhythm and intonation, as any changes of mood, tone or speed can be reproduced closely from short-term memory. Not having to rely on written notes also permits the interpreter to maintain eye contact with the audience at all times, and leaves more processing capacity for listening and analysis (and even partial pre-formulation) in the reception phase of the consecutive interpreting process. The new technique is thus said to be of advantage not only for the clients (by ensuring maximum fidelity to the original) but also for the interpreter, who is freed from the need to take notes and rely on well-practiced memory skills and can give a more confident and relaxed rendition upon the second hearing of the original speech.

As it turns out, the main asset of the technique, i.e. the close reproduction of the original speaker's content and delivery, is also its chief liability, at least for some types of speeches and speakers. Whereas its benefits are evident for fast speeches with high information content, SimConsec impedes the compression and streamlining of extemporaneous source material that is one of the hallmarks of consecutive interpreting. SimConsec obviously inherits also the drawbacks of the simultaneous mode, in which the interpreter can do little to enhance the effectiveness of a slow and repetitive speech delivered with hesitation, possibly in the speaker's non-native language. This problem, encountered during Michele Ferrari's first comparative test of the method in SCIC, can be overcome by accelerated playback, i.e. simultaneously rendering the recorded source speech at a higher speed. Even so, the interpreter's room for manoeuver is limited, which makes delivery speed one of the critical variables requiring further study. Others include the language combination involved and the genre of the original speech, aside from more practical questions pertaining to the equipment (e.g. different types of hardware, software-based sound-editing).

One potentially problematic feature of SimConsec that emerged from trials in SCIC is the interpreter's less than natural delivery style compared to consecutive interpreting with notes. Contrary to its inventor's initial assumption that simultaneous consecutive would permit him to work "with the right intonation at all times", always "looking … the audience straight in the eye" (Ferrari 2001), observers in the trials found the interpreter's rhythm in the accelerated-playback condition unpleasantly awkward and noted that his gaze had been fixed on an indefinite point in the ceiling. This issue, among others, has been addressed in a study of SimConsec at the Center of Translation Studies of the University of Vienna.

Vienna SimConsec Study

Among the many research questions raised by the new hybrid mode of interpreting, the crucial one is its relative merit compared to traditional consecutive interpreting with notes, which it might supposedly replace. In our small-scale experimental study, we therefore asked professional conference interpreters with university training and years of experience in the use of either technique to perform a traditional consecutive and a simultaneous consecutive interpretation of two carefully scripted and thus highly comparable French speeches before a small audience. Data collection was completed in early July, and the results of the study are currently being analyzed. Based on transcripts, video-recordings of the interpreters' output and questionnaires filled in by the experimental audience, the comparative evaluation focuses on accuracy and completeness as well as delivery features such as fluency, intonation and eye contact with the audience. Given the many variables involved, the principal finding of the study is already clear: that much further research will be required to establish whether consecutive interpreting is indeed, or should be, 'going simul'.

Directionality in Chinese/English Simultaneous Interpreting: Impact on Performance and Strategy Use

Chia-chien Chang
National Taiwan Normal University

Chia-chien Chang is an Assistant Professor in the Department of English at the National Taiwan Normal University. Holder of a Ph.D. in Applied Linguistics from the University of Texas at Austin, she received her MA degree in Chinese/English Translation and Interpretation from the Monterey Institute of International Studies. Her major research interests include theoretical and pedagogical aspects of interpretation and translation, second language acquisition, and foreign language teaching.

This study explored professional Chinese/English interpreters' experience of simultaneous interpreting in different language directions, focusing specifically on the impact of language direction on performance and strategy use. Ten professional Chinese/English interpreters from Taiwan participated in the study. All ten participants were native speakers of Mandarin Chinese. Seven participants identified Mandarin Chinese as their dominant language and English as their strongest foreign language. The remaining three had substantial experience of living in an English speaking country as children and either identified English as their dominant language or reported equal competence in Mandarin Chinese and English. The participants were asked to interpret two speeches from English into Mandarin Chinese, and two speeches from Mandarin Chinese into English, each followed with a stimulated retrospective interview. At the end of the fourth retrospective interview, the participants were asked to comment on the differences between their experience of interpreting from English to Chinese and interpreting from Chinese to English in general.

The first major research question raised in this study was how language direction would affect Chinese/English interpreters' performance. This question was addressed through a propositional analysis of the semantic content of the participants' interpreting outputs as well as an error analysis of their linguistic quality. The results indicated that for professional Chinese/English interpreters who had learned English as a foreign language, the percentage of propositions rendered was significantly fewer when interpreting from English to Chinese, or in A-to-B interpreting. Moreover, there was a strong correlation between the interpreters' self-perceived gaps in their A and B language capabilities and the gaps in the percentage of propositions they actually rendered when interpreting in different directions. Error analysis of the linguistic quality of the interpretations showed that the participants reporting dominance in Chinese made significantly more language use errors when interpreting from Chinese to English, or in A-to-B interpreting.

The second major research question raised in the current study was how language direction would affect Chinese/English interpreters' strategy use. This question was addressed by analyzing the participants' retrospective interview data along with their performance data. The results showed that, as the comprehension and production activities in SI interact with each other closely, the L1 advantage and L2 disadvantage often cancel each other out. Thus, the interpreters were found to encounter similar problems in their interpreting in different directions and also used similar strategies to address these problems.

However, the differences between interpreting in different directions became more apparent when we consider the role of language proficiency played in many of their decision-makings. As demonstrated by the retrospective interviews, the participants in the study were well aware of the

gap between their A and B language proficiency. This awareness seemed to affect their allocation of resources, both consciously and unconsciously. For example, when selecting information for encoding (Liu, Schallert & Carroll, 2004), the participants appeared not only to take into consideration the relevance of the information to the audience, but also the linguistic resources available to them. As a result, they tended to be more likely to omit messages that they had difficulty expressing in their B language or to resort to meaning-based interpreting, such as generalization or transformation to express the message. They also appeared to focus their effort more on expressing the essential of the source speech in A-to-B interpreting to achieve optimal overall performance.

An intriguing theme emerging from the data that goes beyond the original research questions is how the concept of norms was involved in the performance and strategy use of interpreters working in different language directions. In terms of expectancy norms (Chesterman, 1997), most interpreters mentioned, directly or indirectly, the importance with which they held their productions to be fluent, logical, without prolonged pauses, and to express the sense or main ideas of the original speeches. In terms of professional norms (Chesterman, 1997) that licensed the appropriate use of strategies, these interpreters agreed that it was acceptable to omit redundant or less important aspects of a message in order to allow them to catch up with the speaker, to generalize when they were uncertain about the meaning of the original message, or to adapt the original message in consideration of the background knowledge of the target audience.

These norms were particularly important when the interpreting task became too difficult and the interpreters felt they had to make compromises between the "ideal" quality standard in SI and their actual working conditions (Garzone, 2002). Interpreting from A to B languages seemed to be an example of such a condition that the interpreters reported needed compromises in order to achieve overall optimal performance. As a result, when interpreting from Chinese to English, in order to achieve a fluent and logical interpreting, interpreters reporting Chinese as their A language seemed to be more ready to forgo completeness and concentrate on the essentials, to use meaning-based interpreting to overcome the linguistic gap in their B language instead of engaging in searching for possible equivalents that may exist, and to dedicate more attentional resources to grammaticality and acceptability of their languages. In other words, the fewer propositions produced in their A-to-B interpreting can be considered a compromise they made to maintain the overall quality of their interpreting through appropriate strategy use. As the English A interpreters also emphasized the necessity of eliminating some of the redundancy in the original Chinese speeches, and the importance of having explicit logical cohesions in their interpreting in English, it can be argued that these interpreters' decisions were also made on the basis of the characteristics of the two languages, or the expectancy norms of these languages (Chesterman, 1997).

Results of this study indicate that professional interpreters who must do simultaneous interpreting in both directions regularly may develop strategic approaches to cope with the different demands experienced in A-to-B and B-to-A interpreting. The difference in their performances seems not only to be a result of the asymmetry between their A and B language proficiency, but also a result of their metacognitive awareness of the limits of their language abilities, the strategies available to them, their audience's expectations and other norms they believe apply to their performance, as well as the discourse structures of their working languages.

Are There Working Memory Differences Between Simultaneous Interpreters and Non-interpreter Multilinguals?

Teresa M. Signorelli*1*, Loraine K. Obler, Henk J. Haarmann*2*, Martin R. Gitterman*1*
1 City University of New York Graduate School and University Center
2University of Maryland Center for Advanced Study of Language

Teresa M. Signorelli is a bilingual speech-language pathologist and a doctoral candidate at the CUNY Graduate Center. Her research interests include bilingual language development, bilingual memory, and speech perception and production in foreign languages.

Loraine K. Obler is a Distinguished Professor in Speech and Hearing Sciences and in Linguistics at the CUNY Graduate Center. After The Bilingual Brain: Neuropsychological and Neurolinguistic Aspects of Bilingualism *(1978, with M. Albert), her publications include* Language and the Brain *(1999, with K. Gjerlow).*

Henk Haarmann is area director for Cognitive Neuroscience at the University of Maryland Center for the Advanced Study of Language (CASL). He has studied verbal working memory processes in a variety of populations using behavioral testing, neurophysiological recording, and computer modeling, focusing on the nature of semantic short-term memory and its role in complex cognition.

Martin R. Gitterman is Executive Officer of the Ph.D. Program in Speech and Hearing Sciences at The Graduate Center, CUNY and on the faculty of the Department of Speech-Language-Hearing Sciences at Lehman College, CUNY. His areas of specialization are neurolinguistics, bilingualism, and second language acquisition.

Introduction

There is inconsistent evidence regarding whether or not interpreters have better working memory (WM) than non-interpreters. Some studies show statistical superiority of interpreters to non-interpreters (e.g., Christoffels & De Groot, 2004). Other studies show no statistically significant differences between groups (e.g., Köpke & Nespoulous, 2004). Many of these studies, however, show a pattern toward interpreter superiority.

The present study attempts to understand why contradictory evidence has been reported. We do this by looking at the various capacities within WM, as described by Martin, Shelton and Yaffee (1994), as independently from one another as possible. Martin and colleagues found that following brain-damage, an individual can be relatively more impaired for phonological (speech sound) information than for semantic (meaning) information and vice versa. This dissociation is well supported in the general literature (e.g., Haarmann, Davelaar, & Usher, 2003) and suggests that the two types of information are processed differently. Most of the literature regarding WM in interpreters, however, has not addressed this dissociation.

Working memory can be likened to a mental work-station where information is stored and processed for a limited amount of time. The present study proposes three dimensions of WM that should be considered when determining whether or not interpreters have better WM than non-interpreters. First, we distinguish between storage-only and storage-plus-processing in WM tasks. Simple storage tasks, like cued recall, measure the ability to maintain a limited amount of information in an activated state. In cued recall, an individual is presented with a list of words followed by a cue that prompts recall one of the words in the stimulus list. More complex storage-plus-processing tasks, like reading span, measure an individual's ability to re-activate recently processed information after a brief period of interference. In reading span tasks, an individual reads a set of sentences and then recalls the last word from each sentence. Second, we hold that within storage-only WM functions, phonological and semantic

components must be differentiated. Third, we believe that lexical-semantic contributions should be minimized to assess the phonological component of WM. This can be accomplished via a non-word repetition task. In non-word repetition an individual repeats word-like sound strings that have no meaning. The present study, to date, looks at performance on a storage-only task whereby semantic storage was assessed with semantic category cued recall, while phonological storage was assessed with order cued recall. The two cued recall tasks were closely matched for stimulus properties and response demands.

In attempting to uncover performance differences, we also contend that the effects of experimental factors within tasks indicate differences in WM contribution. Such factors include commonly reported phenomena of WM such as length effects, where smaller amounts of information are better remembered than larger amounts, and serial position effects (e.g., primacy and recency), where information presented at the beginning or end of a list is better remembered than information that is situated medially in a list.

Present Study: Procedures

A cued recall paradigm, similar to Haarmann and Usher's (2001), allows for relatively independent assessment of semantic and phonological WM. Task procedures are designed to dissociate WM from long-term memory and semantic from phonological WM as much as possible. An oral pre-reading phase and fast stimulus presentation rate decrease long-term memory influence on WM. In the task participants see lists of words, presented one at a time, on a computer screen. They then orally recall either **category** or **order** information from the list following a cue. This allows for relative independence in that category cued recall receives greater contributions from semantic relative to phonological WM and order cued recall receives greater contributions from phonological relative to semantic WM.

The participants, to date, include 12 adults who are highly proficient speakers of English as a second language and have at least good working knowledge of a third language. The experimental group consists of six professional simultaneous interpreters with a mean age of 51. The control group, with no interpreting experience, consists of 6 multilinguals with a mean age of 32.

Results and Discussion

As data collection continues, we report descriptive results rather than ones demonstrated via statistical analysis. Preliminary analysis indicates that the dissociation of phonological and semantic WM was realized. Both groups, in general, demonstrated the classic WM phenomena of recency, primacy, and length effects in the order cued phonological condition. These effects were markedly reduced in the category cued semantic condition. The groups also both performed better, overall, on the semantic task relative to the phonological task. Individual performance across both the phonological and the semantic conditions varied quite considerably (see Figures 1 and 2).

There were no group differences, which is not altogether surprising when one considers the high variability, small number of participants currently, and large age gap across groups with the interpreters being about 20 years older on average. The literature shows that, as we age, certain aspects of language performance, like lexical retrieval and semantic short-term memory, deteriorate (e.g., Nicholas, Obler, Albert, & Goodglass, 1985; Haarmann, Ashling, Davelaar, & Usher, 2005). Therefore, it is possible that professional interpreters have an advantage in cued-recall tasks, but this advantage cannot be seen in the current study because the participant groups are not age matched. The likelihood for finding group differences with age-matching is strong when one considers the individual performance of the one interpreter who is in the same age range of the control group. This individual, age 30 years, performed markedly better in the order task and somewhat better in the category task. He also demonstrated less dramatic WM effects (e.g., primacy, recency) relative to all other participants. Increasing the number of participants should also decrease the variability across performances and permit group differences that may be hidden to emerge as well.

The plan for the present study is to continue collecting data on age-matched participants in cued recall as well as other tasks. In the event that group differences do not manifest, we will explore a broad set of demographic and language-learning and -use histories as potential explanations for inter-participant variability.

From Model to Competency: The Role of Interpreting Theory in Learning, Teaching, and Testing

Andrew Clifford
The University of Toronto

Andrew Clifford is a practicing community interpreter and researcher interested in both spoken language and sign language interpreting. In 2003, he obtained a doctorate in Translation Studies from the University of Ottawa, and currently, he teaches at the University of Toronto. His research interests include psychometrics and discourse theory.

Imagine that you are a fly on the wall in the following classroom scenario. It is the beginning of the school year, and several new students in a conference interpreting program are making their first attempts at consecutive interpretation. They have listened to a speech and taken notes. One student begins to interpret the first portion of the speech, but as she works her way through her notes, she sounds like she is reading off items on a grocery list. The instructor stops her, notes that her interpretation "sounds choppy", and suggests that it "needs to flow more." Another student takes over where the first left off, but he too sounds like he is reciting a list. In fact, despite the instructor's coaching, the list reading continues during classroom exercises, and it is several months before there is a noticeable improvement.

These students are having an ongoing problem with intonation, and consequently with cohesion. They are marking tonic prominence in a way that makes their output confusing, and they are using tones that would not likely characterize their own uninterpreted utterances. Their problem – and its duration – is curious, particularly because the consecutive interpretations that their instructor models for them sound remarkably like normal speech. The students do not appear to be reaping the benefit of their instructor's expertise as an interpreter.

This scenario points the way towards some wider questions. In general, how do we understand the knowledge and skills that are required in conference interpreting? How do we take that understanding and use it to enhance student learning? How can it inform teaching practice? Finally, how can it improve interpreter testing?

To answer these questions, this presentation examines the concept of the *competency*, an internal characteristic that is key for success in a professional sphere, and that can be observed indirectly through on-the-job behaviour. To identify and define the individual competencies related to a profession, the usual recommendation is to carry out detailed empirical studies that recruit large numbers of practitioners. However, this kind of process is impractical in conference interpreting, both because the overall population of interpreters is small, and because the bodies that are most interested in it typically have limited financial and human resources.

One alternative approach for defining competencies proposes the use of theoretical evidence as a starting point. In this presentation, I use this kind of evidence, specifically in the form of the theoretical model of the interpreting process outlined by Robin Setton. Setton's model was selected, because it is arguably one of the most comprehensive developed to date.

> The model is complex, but a relatively coherent account of the interpreting process can be constructed from a discussion of its essential components. On the comprehension side of the model, the *parser* recognizes individual words in the SL utterance as belonging to larger strings (propositions) and derives meaning from them, based strictly on linguistic knowledge. The *assembler* combines the linguistic proposition with world knowledge and situational knowledge to develop an understanding of the actors and the actions in the situation at hand. The *executive* uses knowledge of SL pragmatics to judge the speaker's intent, and it assesses the match between that intent and the elaborated proposition fed to it by the assembler. On the production side, the process is essentially the reverse. The executive uses an understanding of the TL audience to judge whether to include, exclude, highlight or tone down SL material in the TL production. The *formulator* begins to plan out actual TL constructions by selecting word order and focus structures. Finally, the parser uses knowledge of syntax and semantics to encode the TL output into phrases and subphrases.

From this thumbnail sketch of the model, it is possible to gain a sense of the skills and knowledge that are required in conference interpreting. For each major component in the model, I propose a corresponding competency. In other words, there are six in total, three on the comprehension side, and three on the production side.

Comprehension

1. Linguistic Parsing
   This is the ability to analyze language at a variety of levels – phonological, morphological, lexical, syntactic, etc – and extract related meaning.
2. Assignment of Reference
   This is the ability to gather clues from the immediate environment and from the preceding discussion and to use them in understanding "who" does "what" to "whom".
3. Interpretation of Inference
   This is the ability to use existing knowledge – in particular the implicit knowledge of the pragmatic rules people follow when they speak – to understand a speaker's intent.

Production

1. Construction of Inference
   This is the ability to determine the communicative intent of the TL production, and to select information that will transmit this intent, based on an understanding of the TL audience and their needs.
2. Deployment of Reference
   This is the ability to choose the actual focus constructions and word order that help achieve the communicative intent.
3. Linguistic Encoding
   This is the ability to retrieve lexical items and assemble them into propositions that comprise the actual TL output.

These definitions are only rudimentary. They lack a level of concrete detail that makes them meaningful for students, teachers, and test developers. For this reason, I present a more elaborated version of the competencies in my presentation. I also re-examine the kinds of difficulties presented by students and consider they ways in which the elaborated competency can inform progress in the classroom.

This work is important because a well defined competency offers a number of benefits. For students, it identifies and describes the actual behaviours that they need to demonstrate in order to succeed in their chosen profession. For the instructor, it highlights the specific information that can shape instruction. Finally, for test developers, the definition points to the actual behaviours that they will need to elicit from test takers when they create a test, and to the actual behaviours that raters will need to watch for when scoring test taker performance. This particular set of competencies also offers assurance to its users, because it is based on empirical investigation: Setton's model incorporates elements of prior, proven research, the model was tested by Setton's own corpus of interpreted discourse, and the production-side competencies based on the model were validated in a recent psychometric study.

Cognitive Tagging in Simultaneous Interpreting

Chuta Funayama
Kobe City University of Foreign Studies

Professor, Kobe City University of Foreign Studies, 2003-
Professor, Osaka Prefecture University, 1997-2003
Associate Professor, Osaka Prefecture University, 1991-1997
Associate Professor, Kyoto Institute of Technology, 1986-1991
MA, Kyoto University, 1976
BA, Osaka University of Foreign Studies, 1974

Trainees in interpreting courses tend to be concerned more about superficial linguistic expressions than the message, or what is conveyed by those expressions. Their concern is not totally illegitimate, for lexical correspondences between SL and TL are generally expected from translation. On the other hand, human understanding of utterances should not be limited to linguistic decoding, as advocated in Relevance Theory as well as by Seleskovitch in the tradition of interpreter training. In view of such seeming dichotomy, it would help trainers to have a theoretical model about the relationship between the incoming verbal expressions and the content they carry. We propose in this study an on-line processing model of simultaneous interpreting, using the notions of conceptual complex and cognitive tagging.

Our approach has the following salient features: (1) Although it purports to present an SI model with pedagogical implication and practitioners' acknowledgement, its theoretical efforts are targeted at exploring the mental processes involved in communication in general, relying on the ongoing research achievements in linguistics, psychology, and other related academic disciplines, as exemplified in the spirit of Setton(1999), for instance. (2) Our model tries to track the incremental development of utterance comprehension without presupposing the completion of sentential parsing. It means that the contextual and all the other relevant information helpful for the hearer's understanding is taken into the modeling each time an informative fragment of discourse occurs. That is to say that the model should account for such a phenomenon that a certain incoming expression gives rise to a possible whole picture immediately at the time of its occurrence, with the help of the previously stored knowledge on the part of the hearer/interpreter, and it in turn helps to parse the following sentential structures and build up their prepositional contents. (3) The proposed model describes a verbal comprehension process in terms of concepts instead of lexical forms. Parsing typically analyzes a string of lexical forms and, in previous modeling the subsequent semantic identification is also carried out based on a lexical unit. In other words, the contextualized or enriched meanings have also been treated like something evolving out of lexical items on the probably unconscious premise that meaning specification could be described most reasonably in terms of lexical forms. In a reversed approach, the proposed model puts concepts in the central place. This enables us to handle contextual and background information in line with the result of linguistic decoding because both of them could be conceptually described while lexical forms do not go across the gap between linguistic and nonlinguistic information sources.

According to the proposed model, each piece of input information undergoes the following steps in the hearer's mind. These steps are supposed to be taken in this order in ordinary cases and some irregularities account for immature performance.

Step 1: To identify lexical information
Step 2: To form a conceptual complex
Step 3: To attach a tag to the formed complex
Step 4: To operate on related complexes for possible integration or construction

Verbal communication is characterized by the use of linguistic expressions and thus Step 1 is indispensable. Step 1 usually leads to Step 2 in our comprehension. What a speaker has said may be repeated by the very same phrase. For example, you may say that she said 'It's dangerous' in the sense that she used the phrase 'It's dangerous.' What a hearer understands, however, defies easy lexical description in actual cases. In this example, what the hearer has kept in his mind may be some idea about poisonous snakes or possible explosives, depending on the specific context. Therefore, Step 2 seems to be a natural sequence after identifying incoming linguistic expressions, though comprehensive modeling at this step is yet to be explored. Since concepts are intrinsically complex and difficult to handle directly, we propose, at Step 3, to attach a tangible tag to the conceptual complex formed in the hearer's mind at Step 2. We assume that tagging is not simply a theoretical maneuver but reflects a cognitive function the hearer performs. Although these steps are applicable to general communication, Step 2 bears a particularly important message to novice translators and interpreters. Literal translation could be described in this model as skipping Step 2.

A tag at Step 3 may be labeled with a lexical item, which represents the conceptual complex at hand, though the lexical term need not be taken from the incoming expressions. For instance, the lexical term 'movie star' could be used for a conceptual complex produced by the SL expression 'Tom Hanks' and other elements in the context. Furthermore, a tag may not have any specific lexical name and be mentally held as an abstract concept or as a general conceptual frame such as 'an act of giving.' The interpreter's comprehension should be based on conceptual complexes rather than lexical tags, though the latter is important in organizing the interpretation. Integration at Step 4 results from gathering conceptual complexes formed separately at the outset, while construction builds up a conceptual complex of higher order, such as summarization and re-categorization. This step also reflects what we do in our daily communication. For instance, one could easily understand that the speaker intends to refer to 'terrorism' by saying 'flying airplanes into buildings.' The word 'terrorism' may function as a tag given to a concept constructed anew yet based on a series of separate concepts triggered by such linguistic expressions as 'flying' and 'airplanes.'

The advantages of the proposed model further include the new classification of various interpretation output among different interpreters. According to the proposed model, it is justifiably assumed that different interpreters produce conceptual complexes in different ways at Step 2. If conceptual complexes are formed in a similar way, tagging may differ among different interpreters. Furthermore, if formed concepts and tagging are similar, the operation on them may differ. By looking into the nature of such differences over multi-stage mental works you may characterize the sources of differences more in detail.