

**Computer-based Information Management for Conference Interpreters  
Or How Will I Make my Computer Act Like an Infallible Information Butler?**

**Anja Ruetten**

Fachhochschule Koeln  
Institut für Translation und Mehrsprachige Kommunikation  
Mainzer Str. 5  
50678 Cologne  
Germany  
ruetten@sprachmanagement.net

## **Introduction**

Interpreters - just like many other professionals - have to deal with an abundance of information from many different sources, and they do so under extreme circumstances: When preparing for the subject of a given conference (usually highly specialised, with subjects varying extremely from one conference to the next), they need this very special and reliable information in the respective languages and well before the conference starts. Interpreters therefore need a system that serves their professional purposes by accessing, categorising and representing the required content and linguistic information quickly and precisely according to individually pre-set criteria.

This helps them memorise relevant information under time pressure, permits intuitive information research while interpreting (in the booth) and supports efficient follow-up and updating of the personal information database afterwards. This paper will suggest a model that could satisfy the special needs of interpreters, being complex enough to satisfy their special and individual needs as well as simple and user-friendly enough to produce the necessary output easily and quickly, as it will only be useful if it does not cause additional complications.

## **Interpreters' Information Management**

As to the methods and tools for information management before, during and after a conference, there are a number of scientific articles, some dissertations and PhD theses and also several terminology management databases written and designed by and for conference interpreters. However, there are many open questions in terms of methods and optimum practical solutions in this field.

In my dissertation (2000) I started to investigate the basis of optimum information management before, during and after a conference – with the main focus on conference preparation as a means to facilitate information handling during interpretation – and then drew up a list of requirements and a basic structure of an appropriate software model. My intention was mainly to leave aside the methods of information management I and my fellow students had developed spontaneously during our studies at university and first practical experience. To do so, I looked into interpretation research findings as well as neighbouring disciplines like terminology, foreign language teaching and the psychology of learning in order to find out about specific needs. As I had expected, much evidence could be found about characteristics and problems of conference preparation for interpretation:

Interpreting is a very knowledge-intensive activity. Preparing for it means self-organised learning. It is important to systematise this process of information retrieval from huge quantities of data (material received from the customer or colleagues, the interpreter's own existing data e. g. from previous conferences and other sources like the internet), which is repeated over and over again with constantly changing, very specific subjects and under great time pressure. The following points are of special relevance:

It is crucial to prepare on the **semantic** level to make sure the meaning of what the speaker is saying is understood – something that should not be taken for granted considering the fact that today there may be a conference about sugar beet harvesting, tomorrow one about anti-aircraft missiles and the day after that a discussion about accounting problems in an international company. In this context, efficient background document research, classification and storage (in order to be able to find the material again when working more than once for the same customer) is essential; on the other hand, when it comes to extracting/finding and managing terminology, instead of recording mere vocabulary, to also include additional information like semantic and conceptual aspects and relations (definitions, hypo-, hypero-, antonyms, mero-, holonyms etc.).

When working through technical texts it will be easier to concentrate on the meaning of the text if relevant **terminology** has already been **extracted** (semi-)automatically. Copying terms from the text onto a list is quite time-consuming. Automatically generated term lists will always have to be revised by the user, but even so, the “rough” work has already been done and the interpreter,

when reading the list, can concentrate on those terms which are relevant and important to remember or that he/she already knows.

Working with multilingual **word lists** is something widely disapproved of at least among translators or terminologists. These simple lists lack any kind of additional information concerning grammar, meaning, reliability etc., they often even seem inconsistent and unstructured or don't correspond to the official, standardised terms. However, several interpreters have already underlined their usefulness, and there are occasions where such lists may very well serve the special purpose of conference interpreters, which simply entails finding the right word at the right moment. Unfortunately, they can not be relied upon to always deliver the needed information immediately ("*Crankshaft flywheel?* - Wait a second, I think have this one on a list from last month's conference."). It should be noted that interpreters' requirements are quite different from those of translators, terminologists or lexicographers and that they need something much less complete, but that, on the other hand, something more complete than what is found currently in practical life might still be useful, e. g. including all entries in one database or at least under one surface.

The existence of **equivalences** between languages is normally denied; however, simultaneous interpretation makes it necessary to automate parts of the transfer, so that in a limited number of cases (names, some technical terms etc.) elements of different languages can be considered equivalent at least for the context of a special conference.

**Memorizing** terms and other language elements could be more systematic and efficient when controlled by means of a program. It could also be helpful not only to train terms visually, but also acoustically – which, apart from addressing a different perception channel and enhancing integration into the memory structure, also corresponds to the real working conditions.

Interpreters will never be machines spitting out word equities; however, in order to retain very technical terminology within (and for) a short period of time, some **automation** of word pairs may be necessary, e. g. in the case of complicated new expressions, so that when talking about things like *Ansaugkrümmerunterdruckumformer* or *hydroxyanthraquinone*, the interpreter still gives a convincing and professional rendering. A distinction can be made between transfers that

can or should be automated between two or more languages – “1 to 1” (or near) correspondences – and transfers that must be done consciously, which take up a considerable part of the attention available. An ideal interpreters’ program should be designed accordingly.

Basically there are two levels of preparation: **long-term** preparation – i. e. the cultivation of the working languages (active and passive, mother tongue and foreign languages), especially in terms of general language – and **short-term**, technical conference preparation (special language).

Findings in learning psychology suggest that the **structuring and classification** of information in different categories enhances retrieval. Especially in the situation of simultaneous interpretation less attention can be dedicated to the task of word retrieval, so it must work smoothly. The information inventory of a database should be structured individually, according to the individual interpreter's working and memory structures into which the new information must be integrated, in order to make it retrievable not only from the computer but from the individual's memory. A well-structured information stock on your hard disk will also aid a well-structured memory. When memorising linguistic and technical knowledge, classification plays an essential role; it can be helpful to work in different categories, but it may also be good to deliberately memorise elements in a disordered way.

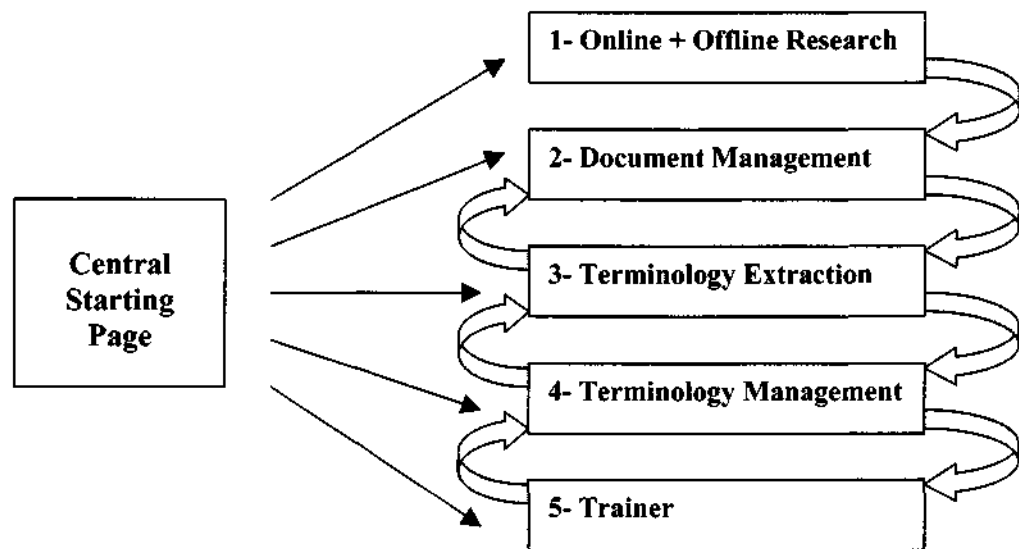
There is a huge amount of potentially useful **additional information** and characteristics for the entries of a terminology database. Thus, space for many additional items should theoretically be made around each entry, even if not every interpreter makes use of the whole range of possible additional items and additional information is not entered for each and every entry. A link between a special term and a corresponding picture can provide illustration, which aids memorisation. Notes like “difficult”, “very important” or “useful/nice formulation” (to improve style) as well as degrees of reliability or information sources are examples of quite a number of useful pieces of information. But I would say that filing more than what is deemed beneficial by the user would be counter-productive and would result in interpreters not using this (or any other) system.

The above roughly summarises the characteristics and difficulties of interpreters’ information management identified in my dissertation and the implications for the required software.

Basically, it can be stated that the computer offers a great diversity of possibilities to make interpreters' work easier, and this entails much more than simple terminology management.

### A Software Model

The following illustration and synoptic description of functions show the software model deduced from the requirements identified in my dissertation. This model consists of five modules and one central starting point.



Module 1 – **Online and Offline Research** – finds internet documents of any kind about a certain subject (text, pictures as well as glossaries and dictionaries) as well as documents already filed in Module 2 using key words. It sorts the results by relevance and language.

Module 2 – **Document Management** – manages documents found in Module 1, archives the interpreter's own documents, the ones received from the customer etc. This can be text, glossaries, graphs, bibliographies, summaries etc. Those documents can be found or sorted by subject, relevance to a certain subject, text type, language, date etc. Furthermore, status information from Module 3 is saved in Module 2, i.e. Module 2 “remembers” from which document terminology has already been extracted (automatically or manually). It is also possible to indicate whether a text has already been read, where it comes from and when it was last used.

Furthermore, a connection between documents and term extraction and management could be useful in the form of information from Modules 3 and 4 popping up when moving the mouse over a word or phrase. Parallel reading of texts in different languages can also be facilitated.

Module 3 – **Terminology Extraction and Analysis** – analyses documents (Module 2), extracts potential technical terms and their equivalent in a different language when parallel texts in the respective languages are available or it is possible to consult electronic dictionaries or encyclopaedias. The result (term list) is then synchronised with Module 4 – terminology management – in order to check which terms are already filed there. If a term is already registered in Module 4 and it is marked as “known to the user”, it is not put on the new list of extracted terms from Module 3. If it is filed in Module 4, but not marked as “known to the user”, it will appear on the list of extracted terms in Module 3, but there will be an indication saying that it is already registered in Module 4. Module 3 can also create word field lists for key subjects of a given text.

Module 4 – **Terminology Management** – manages entries (singles words or whole sentences) coming from Module 3, those entered manually by the user and also entries imported from other databases (from colleagues, customers etc.). Such “external” data will of course be marked accordingly, making sure that it will not be integrated into the terminology database without being checked. Similar or double entries will also be marked and the user will be informed about them the moment they are entered into the database. Additional information can be entered if need be, like subject area (with no limit in the number of structure levels), project (conferences), source, grammatical category, degree of difficulty or importance, style, definitions, descriptions, context and graphs, also abbreviations and acronyms, date of entry, last access. The latter could be registered automatically. Relations can be established between different entries like “1 to 1 translation”, “suggested translation”, possible interference/faux amis, synonyms, antonyms, hyperonyms and hyponyms.

Furthermore, status information from the training module (5) will be saved automatically. This information includes the date when an entry was last tested and whether it is actively or passively known or not known at all. Queries can be made individually (“looking up”, morphological variants will also be recognised) and in the form of lists sorted and structured according to special

criteria (conference, subject, degree of difficulty etc.) It is also possible to have a hierarchical representation of terminology of a certain subject area with the different subject levels.

Module 5 – **the Trainer** – helps systematic memorisation of terminology saved in Module 4. A difference is made between entries that must be “drummed in”, i.e. 1 to 1 equivalents between two languages, and translations that are merely suggestions or elements that are translated differently in different situations or contexts. The latter are not “tested” in the sense that the user must render the exact equivalent but they are only “presented” visually or acoustically. The user’s learning information is saved: Unknown or problematic entries (“tip-of-the-tongue”, tongue twisters or simply unusual or complicated expressions) are marked as such (manually by the user or automatically, if the user fails the test). They will then be presented or tested automatically at regular intervals. This does not happen to known entries: They will remain untouched by the trainer except after a certain (user-defined) period of time. This function is particularly useful for the maintenance of a certain level of general language, a task often neglected in the heat of the moment of concrete conference preparation.

The testing or presenting can be done in an ordered manner (subjects, degree of difficulty, projects etc.) or in a deliberately unsorted way.

Finally, immediately before a conference, it is possible to print out a “last minute” list of very important or complicated (“*Bremsleitungsbördelwerkzeug*”) or still unknown entries of a certain subject.

Another useful function, especially for students, could be exercises aiming at the practice of interpretation-specific skills, e.g. clozing (visually or acoustically presented) or the presentation of word series that must be memorised and then rendered in a structured way.

In order to make it easier to search for a specific item during simultaneous interpretation in the booth, there should be an **overall-quick-search-key** that can be used blindly and independently of the module or function the user is using at that particular moment. When the user strikes this key and types the word or expression he or she is looking for (or part of it) all the modules are searched and the result in the respective (pre-defined) languages pops up in a big window showing all the results in a clearly legible and well-structured way. Different colours for different languages (user-defined) might be helpful to grasp the search result at one glance. With another key strike, the window will close again. So if, for example, the interpreter was reading a

document which is being discussed in the conference in Module 2, he or she can intuitively strike the search key, find the respective word and translation and then strike the close-search key in order to go on reading the text.

The software model described above is the ideal picture of the interpreter's electronic information manager. It is based on scientific findings, which, however, don't always come from interpretation research as such but from neighbouring disciplines. This picture still needs to be completed and refined.

### **Existing Software Solutions**

Online search engines or meta search engines are available and must be tested as to the relevance of the results delivered. There are also programs for document management and web search results management as well as automatic indexing and classification. For term extraction, there are also several products available on the market which will have to be tested in order to find out which one best serves interpreters' purposes. The same holds true for vocabulary trainers which most probably can be adapted to interpreter's needs. The major problem here will be the compatibility and the information interchange between the programs.

As to terminology management tools, some professional interpreters have developed their own terminology databases as opposed to the existing commercial systems like Multiterm or Term Star, which have basically been designed for translators' purposes. This fact shows already that the standard terminology tools do not precisely serve interpreters' purposes. Apart from compatibility issues, the decision between one of the standard tools and a special interpreters' system is one of the major questions in this complex matter.

### **Technical Challenges**

Although the technical aspect is not the subject of this article, it should be briefly addressed. Currently existing technical devices in principle fit the requirements described above. There is, however, room for improvement in some areas:



- It would be good to have mobile scanning devices in order to scan paper documents even if they are provided only some minutes before a conference/speech starts. Scanning results should be reliable and the devices of course affordable and not too bulky.
- Using laptops in the booth is still considered annoying by some colleagues. Much is however being done, flat silent keyboards are being developed. Laptops have already become smaller and less noisy.
- Voice recognition (and recording) could support interpreters in situations where they do not acoustically understand the speaker. So far, voice recognition is still in its infancy.

### **Research Project**

As stated above, much evidence concerning interpreters' information management was derived from neighbouring disciplines. The picture thus still needs to be completed and refined, some basic questions concerning the interpreter's work in the booth (at the conference) and at his desk (preparation and follow-up) need to be examined more closely. In order to do this, the University of Applied Sciences in Cologne and the Institute of the Society for the Promotion of Applied Information Sciences at the University of the Saarland are planning a research project aiming at developing a software model based on findings from relevant theoretical and practical research and subsequently designing a prototype program to be tested in the "real world".

The research programme to develop an interpreters' information management system will be subdivided into four phases:

1. Conceptualisation: The 5-module concept will be refined and examined by further research in some key areas (see Perspectives) in order to obtain a model based on reliable practical and theoretical research results.
2. Evaluation: Existing programs will be evaluated in order to find out which ones best serve interpreters' purposes. New elements will possibly have to be created.
3. Prototype: A prototype will be designed according to the implications of phases 1 and 2.
4. Testing: The prototype will be tested in real life.

### **Perspectives**

Apart from the more technical problems and the question of whether or to what extent existing software solutions serve interpreters' needs, there are some key areas of interest that deserve further investigation:

How should **relations** between terminological entries be organised? There are various kinds of relevant relations in the terminological inventory of an interpreter: Conceptual relations like hypero-/hyponyms and mero-/holonyms, semantic relations between denominations like synonyms and antonyms, also more subjective relations like classification in different subject areas or projects, faux amis and many more. Here, the question of multiple classification is also relevant.

A good example to illustrate the complexity of this aspect is a part of the organisation of the Eurokorps in Strasbourg: Here, a "branch" is called "Abteilung, sección, bureau" in German, Spanish, French. A "section" (subordinate to a "branch") however is called "Dezernat, negociado, section" in German, Spanish, French. The Spanish "sección" can also have a different meaning (Führungsgrundgebiet) in the same environment. In this context, it also has a synonym, "sector de actividades", which again could also be used in an economic area. The relations between concepts (hierarchy) are clearly important here in order to understand the system; relations between denominations are also important, like synonymy relations within the same language (Spanish "sección" and Spanish "sector de actividades"), relations between denominations with different meanings (concepts) and their different equivalents in other languages (Spanish "sección" – German "Abteilung"/German "Führungsgrundgebiet") and similar denominations of different languages with different meanings (Spanish "sección" – English "section") which, when highlighted or at least marked in the database, can help avoid interference especially in a situation where there is limited mental capacity to spare, as in the case of simultaneous interpretation. Another question arising from this concept-denomination item issue is where to situate the classification into different subject areas and projects or customers. The concept is usually decisive when it comes to dividing into subject areas, but when it comes to customer-specificity, denominations can be different for the same concept (German "Armaturenbrett" or "Instrumententafel" for English "Instrument Panel").

Further investigation is necessary here in order to find out which relations an interpreter actually needs or wishes to be registered in his database and how this can best be implemented in an adequate structure and how such a structure will then be compatible with the term extraction and

training module. For practical acceptance, however, the system must be easy to use, with comfortable input, search and list generation functions. This aspect is often criticised when it comes to evaluating existing terminology management systems.

What are the **learning methods** of conference interpreters when preparing for a technical conference? Are there general “patterns”? Could they be optimised? Do interpreters ever dedicate some of their time to long-term preparation and maintenance of a high language level in their mother tongue and foreign language? Is self-discipline a problem here and if yes, how could it be tackled?

Where is the limit between **individual** databases and **general** (public) ones? Does an interpreter have to save every detail that comes up in the context of a conference in his personal database, or is it sometimes enough to know where to find it (dictionaries, internet, other people's glossaries)? Or might a kind of “book marking” be more useful? What kind of information is likely to be useful later on, what is not? For example, in a conference about agricultural statistics, a list of 20 different types of fruits and vegetables are discussed. The interpreter, eager to know all the 20 words in all his working languages (two, three, four or more), puts all the terms into his database. The following week, he works at a conference about fishery – 20 different fish names occur. Will all this very special terminology (have to) be memorised? Does it all have to figure in the interpreter's database if he can also find the terms in his electronic dictionaries? Does the structure of the database have any impact on these questions? Isn't it naïve to believe that we will remember a term just because we have it in our database? Should everything be in one single file? And do I need to save organisation-specific “insider” information like names, acronyms, terms “invented” for a special thing?

What does the optimum **follow-up** of a conference look like? What impact does it have on the quality of the interpretation in another conference on the same subject? What impact does it have on the quality of the interpretation in another conference on a different subject?

Would it be reasonable to save **note-taking symbols** in a terminology database?

Could compatibility with a **Personal Digital Assistant** be useful?

Does any special terminology **format**, e. g. MARTIF (machine-readable terminology interchange format), serve interpreters' purposes? Does ISO 12620:1995? Are they too rigid or too vast to be successfully applied by interpreters?

### **Summary**

IT offers promising possibilities, so there is a chance that we might be able to make our computer act like an infallible, perfectly organised, nearly all-knowing, but still discreet butler who reminds us not to forget our keys (key words) when leaving the house, finds the shoes that match our dress (the word we are looking for) in the remotest locations, keeps our place tidy and teaches us where to put our stuff in order to be able to find it when we need it. This could help us to concentrate more on the task of interpretation as such, also in the training phase, whilst being more efficient, less stressed and getting an even better grasp of the (technical) language and underlying knowledge.

After all, interpreters and butlers have one thing in common: They are at their best if you don't even notice they are around. So they should be a perfect team.

### **Bibliography**

Anja Rütten (2000): "Informationsmanagement für Dolmetscher – Anforderungen an spezielle Software zur Konferenzvorbereitung". In: Saarbrücker Studien zu Sprachdatenverarbeitung und Übersetzen, Band 17