

Interpretation Booths for the Third Millennium

Interpreter booths have not changed significantly since the introduction of simultaneous interpretation five decades ago. They have, of course, become more comfortable, their soundproofing has improved, and there has been continuous evolution in console design and sound quality. Nevertheless, one can argue that none of these changes has constituted a radical departure from the original conception of interpreter booths as far as their function and ergonomics are concerned.

By Panayotis MOUZOURAKIS (PMouzourakis@europarl.eu.int)

The democratisation of information technologies (IT) in the 1980s and 90s has led to a revolution in the ways humans communicate and in particular in the ways that information is disseminated and accessed. Does this development not create a new situation for conference interpretation and the functional needs of interpreter booths? And what modifications are necessary to enable booths to profit from the advantages IT offers to interpreters?

The question seems to have been cogently posed for the first time in March 1995 by J. B. Quicheron of the European Commission's Joint Interpreting & Conference Service (JICS) in an (unfortunately) unpublished note titled "*La cabine de l'an 2000, quelques réflexions*". His statement of the problem and most of his conclusions remain valid today.

1. "La cabine de l'an 2000, quelques réflexions" - a summary

In his note, J. B. Quicheron states the following objective:

To have the maximum amount of information available in the booth by electronic means.

The following are identified as expected benefits from the use of IT in the booth:

- Interpreters can deal more effectively with the growing complexity and variability of subject matter.
- Unnecessary document printing can be avoided while ensuring access to all meeting information.
- Interpreters can exchange data through PCs.

Identified interpreter needs consist of:

- Specific thematic information;
- Visualisation of conference material (slides, presentations);
- Specific terminology related to the meeting;
- General thematic information (encyclopaedias, references);
- General linguistic information.

The technical requirements for the use of IT in the booth are identified as:

- Small footprint in the booth (i.e. space efficiency);
- Full language coverage;
- User-friendliness;
- Real-time response;
- Availability for the whole of the meeting;
- Integration of all system components.

This analysis led Quicheron to propose a number of technical options for the provision of information:

- Autonomous system installed in the meeting room;
- PC connection infrastructure alone; or
- Hybrid system (both of the above).

And for the manner in which this information can be viewed by the interpreter, Quicheron identified:

- Flat screen on or next to the interpreter console;
- Head-up-display (HUD) on the booth's glass window; or
- Personal PCs (sub-notebooks, PDAs).

2. Technical developments since 1995 and their consequences

Since the (non-) publication of Quicheron's note, a number of important developments have taken place:

- The explosive growth of the World Wide Web (WWW);
- Java and thin-client computing;
- The introduction of (intelligent?) digital consoles;
- ISDN Video/teleconferencing.

The growth of the WWW is, without question, the most important of the above developments. Most of the information available in electronic form today is accessible through browsers using the WWW protocols, either within the intranets of large organisations and corporations or on the Internet itself.

Insofar as interpreter booths are concerned, it is clear that access to the Web through a browser represents the most effective information interface. It should be noted that, even for large organisations such as the UN or the EU institutions, it is unlikely that their own intranets will ever provide anything near the wealth of information available on the global Internet. Finally, projected multimedia conference systems will also almost certainly use the Internet as a vehicle for document and data interchange between participants. If such conference systems become the norm, interpreter consoles will have to provide equivalent functionality.

Apart from static information that can be provided through intranets and the Internet, interpreters also require access to executable content, such as glossary database applications. This requirement could be satisfied through the use of their own private notebooks or personal digital assistants (PDAs). Such a solution is, however, not the optimal one, for the following reasons:

- Private notebooks or PDAs can be obtrusive and/or noisy, and thus objectionable to other colleagues in the booth.
- Connecting these notebooks and PDAs to intranets and the Internet in the booth requires either dial-up access through a modem (which must be provided for every interpreter in the booth) or else nontrivial configuration in order to hook up to the local network.
- Most importantly, such a solution will be adequate only for the notebook or PDA owner. His/her colleagues will in general not be able to benefit from this information, not even to provide help during that interpreter's turn in the booth. There is, of course a work-around for this problem: imposing the use of a single make of notebook/PDA for all, but this is unlikely to be accepted voluntarily.

The other issue is interconnectivity and interoperability of interpreter consoles with interpreter notebooks: PDAs and the like are not a very realistic option in this light.

What is needed instead is a platform- (and eventually device-) independent solution for information access. The best solution available today would be platform-independent executable content in java and the use of a thin client (WWW browser capable of downloading java executable content as needed) for accessing it.

It is also clear from the above considerations that personal notebooks/ PDAs cannot cover the requirements for document visualisation by interpreters. This leaves open the possibilities of visualisation by a flat screen on or next to the interpreter console or the option of a head-up-display (HUD). The latter would permit the visualisation of documents and other content but not their manipulation, which requires interpreter input; a tactile screen is probably the best solution in this respect.

Thus we conclude that the best way to provide access to information in the booth in a non-obtrusive way and without any need for configuration is to use the interpreter console and its associated screen. It follows that:

Interpreter consoles must incorporate Internet connection capabilities and WWW browser software, plus a tactile screen adequate for the visualisation of material accessed through the WWW.

3. Considerations of ergonomics and cognitive integration

The extent to which conference interpretation is not a linguistic exercise but rather a complex task crucially dependent upon the interpreter having the necessary information at his/her disposal is probably realised only by interpreters themselves-and the rare conference organiser who will go to the trouble of providing them with adequate documentation in good time.

There is no doubt that IT can potentially make more information available to the interpreter in the booth at much higher access speeds than traditional printed document support. In principle, then, the expected increase in interpretation quality from the use of IT techniques in the booth should be well worth the cost and effort required to provide the necessary infrastructure.

There are a number of caveats that must be taken into consideration, however. First of all, it would be dangerous to believe that access to information in the booth could replace adequate preparation before a meeting. Indeed, for many interpreters access to information is useful only *before* a meeting and not during it. A classic example is the case of the interpreter receiving the text of a long speech he/she is interpreting midway through that speech. Most interpreters would just push the text aside. On the other hand, if the same interpreter has received the text beforehand and has had the time to prepare it, it can be an invaluable aid to interpreting.

The interplay between the normal mode of interpreting (i.e. capturing and transmitting the semantic content of a message) and the transcoding from one language to another of isolated, referentially specific units that must be stored in short-term memory, such as numbers and proper names, is another interesting example. Typically, interpreters have trouble combining the two modes and have to rely on the help of colleagues to keep track of numbers if they crop up too often.

In general, there is always the possibility that two or more information channels (the visual and the auditory channels in the first example), or even parts of the information contained in a single channel (as in the second example), will interfere destructively if the interpreter cannot integrate them, whether because of time pressure, extra stress, or simply because of the inherent limitations of human cognitive and memory structures. It is perhaps worth pointing out that, as brain scans of interpreters working have demonstrated, simultaneous interpreting is one of the most intense cognitive activities in which the human brain can engage.

These considerations will probably be applicable to the potential use of speech recognition in the booth, as this technology is expected to attain maturity during the next decade. It could become possible in the near future to provide the interpreter with even an incomplete transcription of the speech he/she is currently interpreting. Whether this would be of help or just create confusion is, of course, impossible to say as yet. It is likely, however, that a robust number recogniser (which would be much less demanding to implement from a technical point of view) might prove very useful.

4. Videoconferencing and remote interpreting

The problem of cognitive integration is also relevant to interpretation under videoconferencing

conditions and especially to remote interpreting (without a direct view of the meeting room). The latter technique is being envisaged by major organisations employing interpreters as a possible solution to the problem of accommodating an increasing number of booths in meeting rooms and/or as a means of reducing interpreter travel costs.

For the transmission of audio and video signals, both satellite links and ISDN connections have been used to date. It is conceivable that in the future the Internet itself will be used as a transmission medium, augmented by some resource reservation protocol to ensure the availability of the required bandwidth for audio and video transmission. As the cost of satellite links is still considerable, most interest has focused on the possibility of using terrestrial links, most notably ISDN (i.e. digital telephone) connections. However, ISDN videoconferencing in the past has had to contend with inferior image and especially sound quality, the latter being limited by the H320.x family of protocols to a pass band of 0 - 7 kHz, which is completely inadequate for interpreting.

It has since proved possible, at the price of using proprietary encoding protocols for the audio and video signals, to provide almost acceptable image quality with only 384 kbps and, much more importantly, CD-quality audio using MP3 (MPEG-2 Audio Level 3) encoding in a single 64 kbps channel. This was perhaps best demonstrated in the context of the January-February 1999 tests at the UN, during which an entire 2-week meeting was interpreted remotely. Despite the rather leisurely pace of the meeting, the participating volunteers, all UN staff interpreters, reported experiencing unusual fatigue and also complained of eyestrain and insomnia or nightmares.

It is almost certain that these problems did not derive from inadequate sound quality; in fact, some interpreters found the sound quality in the tests to be superior to what they were accustomed to in their usual, antiquated booth installations. Could they be put down to lack of image definition alone?

It is more likely that it is the static, two-dimensional nature of the image that is to blame. It has already been observed that, under videoconferencing conditions, interpreters prefer to obtain visual information by looking at live participants in the meeting room, even if they are offered the option of a larger format projected image. It is possible that working from a screen puts the interpreter in a double bind: On the one hand there is an unavoidable loss of visual information due to the fixed angle and the less than perfect image definition and two-dimensional visual field. But, at the same time, the interpreter is forced not only to perform an unfamiliar cognitive task (obtaining non-verbal information from the speaker's and the participants' body languages through a screen), but also to integrate that task with listening and speaking under the extreme conditions of simultaneous interpretation.

If such is the case, the judicious use of HUD displays, i.e. projection on the glass window in front of the interpreter, could at least partly alleviate the problem by providing the illusion of a three-dimensional environment. One would also intuitively expect that the use of virtual reality (VERY) to permit the interpreter to virtually navigate the meeting room (e.g. by using a joystick), observing it from various angles, could dramatically improve the situation, at least for the younger generation of more "wired" interpreters.

5. Some tentative conclusions: the human factor

It is clear that in the near future conference interpreting will be confronted with a major metamorphosis at least as radical as the past change in the dominant mode of interpreting from consecutive to simultaneous. Faced with the prospect of transition to a new mode of interpreting, which for lack of a better word could be termed "cyber-interpreting", it is important that interpreters and the institutions that employ them invest the time and effort to perform experiments and study this phenomenon.

While in the past it has been the technical challenges (access to information, adequate quality for audio and video streams, etc.) that have received the most attention, it is probable that it is the human factor and most precisely the limits of human-computer interaction and cognitive integration that will provide the boundary conditions for the future, determining what is viable and what is not. In particular, there are a number of issues that are relevant to the use of remote interpreting:

- Interpreter motivation (or rather de-motivation) will probably be as serious a problem as cognitive overload. As one of the interpreters participating in the UN remote interpreting

experiment put it, "I just couldn't get the adrenaline running." Almost all considered the quality of their work under the experimental conditions to be substandard. At least for the current generation of interpreters, remote interpreting is likely to continue to be perceived as degrading, with a consequent further reduction in quality.

- Health issues are not to be overlooked. Very little is known, for example, about the impacts upon interpreters' physical and mental health of working from a screen for a sustained period of time, except that they certainly are not benign. Current working time limits for interpreters will accordingly have to be reduced to safeguard interpreters' health, possibly rendering remote interpretation economically unattractive in the process.
- Interpreter training will have to undergo fundamental changes in order to train interpreters to work in a fundamentally different mode. The criteria for trainee selection will probably have to be adapted, with a premium put on "cyber-interpreting" rather than on traditional communication skills. A radically different interpreter profile will probably emerge in the process. A rift between the older interpreters and these young hotshots will inevitably result.

It is to be hoped that institutions employing large number of interpreters will carefully weigh all the pros and cons before any irreversible decisions in this direction are taken.