



Interpreting Platforms

Consolidated test results and analysis

Table of Contents

1.	SUMMARY	2
2.	BACKGROUND	2
3.	DESCRIPTION	3
4.	PLATFORMS	3
4.1.	Choice of the platforms to be tested.....	3
4.2.	Security and data protection.....	3
5.	TESTS	4
5.1.	Limitations	4
5.2.	Equipment.....	4
5.2.1.	Laptops.....	4
5.2.2.	Headsets	5
5.2.3.	Built-in laptop cameras	5
5.2.4.	BYOD	5
5.2.5.	Network connections.....	6
5.3.	Settings	6
5.4.	Overall analysis of results.....	7
5.4.1.	General.....	7
5.4.2.	Network connection	7
5.4.3.	Sound	7
5.4.4.	Image.....	8
5.4.5.	Lip sync.....	8
5.4.6.	Interfaces.....	8
5.4.7.	Reactivity of controls	9
5.4.8.	Switch/Handover.....	9
5.4.9.	Chat function.....	9
5.4.10.	Working as if with portable interpreting equipment	9
5.4.11.	Working from home	10
5.4.12.	Impact on interpreters in tested conditions.....	10
5.4.13.	Nice-to-haves	10
5.4.14.	Speakers	11
5.4.15.	Participant's Apps	11
6.	INFORMED OPINION	11

1. SUMMARY

For the past two years, software-based interpreting solution providers have made their presence strongly felt on the market. International institutions are being approached by developers presenting their products as state of the art, flexible, and cost effective. This rapidly emerging technology is also the subject of a new ISO document for which there is strong demand from the industry. The relevant ISO Working Group will discuss the new ISO draft at its next meeting. The ISO working group is chaired by a SCIC representative with the participation of representatives from the European Parliament, the Court of Justice and the Council of Europe.

In order to help SCIC prepare an informed opinion for the ISO meeting, SCIC interpreters tested four interpreting platforms in April and May 2019. The tests mainly focused on their ease of use, subjectively perceived audio and video quality, lip sync, features, interpreter's wellbeing, and use conditions in different scenarios. The tests were conducted exclusively from the interpreters' perspective. The user's or participant's perspective was not addressed.

Testing the four platforms enabled DG SCIC to acquire first-hand experience of how these new products work in real life. It also enabled DG SCIC to identify features and requirements, which would be beneficial for users. Several parameters of one or two selected platforms will have to be researched further; use cases and the interpreters' working environment will need to be defined.

The testing showed that, in principle, interpreting platforms can be used to provide interpretation services. Should technical issues, such as the stability of the network, be adequately addressed, the platforms might be usefully deployed in short meetings with reduced language regimes, such as meetings with two languages, and when the meeting participants are not located at the same place, and/or the interpreters do not have to or cannot be on site. In well-defined cases and circumstances, this technology might enlarge SCIC's palette of services – beyond the core business of conference interpreting in multilingual meetings - perhaps satisfying a hitherto hidden demand.

2. BACKGROUND

Technologies used for distance interpreting are as varied as the current modes of face-to-face interpreting. Distance interpreting is replacing and will continue to replace some face-to-face interpreting in medical, court, business and conference settings. Face-to-face interpreting at conferences, in courtrooms, and health clinics will not disappear; it will evolve.

Distance interpreting is deemed ideal for new forms of communication, such as audioconferences, videoconferences, web-meetings, webinars, virtual press conferences, and expert network interviews. In these examples, none of the participants is in the same place as the others. All of them are distant, just like the interpreters. Distance interpreting allows interpreters to provide interpretation services in situations where in the past such service would have been either unaffordable or technically impossible.

In December 2018, ISO TC37/SC5/WG3¹ started the development of an ISO deliverable on Simultaneous Interpreting Delivery Platforms. The future ISO document is intended to give guidance to providers and users about the minimum requirements for these platforms. Since DG SCIC had no knowledge or experience of the platforms' features and possibilities it decided to develop use scenarios and test software-based interpreting platforms in order to obtain an informed opinion on these platforms. The results of these tests will be the basis for the input of DG SCIC's experts involved in the development of said ISO deliverable.

¹ ISO Technical Committee 37 - *Language and Terminology*, Sub Committee 5 - *Translation, Interpreting and related technology*, Working Group 3 - *Facilities and equipment for interpreting services*

3. DESCRIPTION

Interpreting platforms consist of web interfaces and mobile applications that enable interpreters and participants (speakers and/or listeners/viewers) to be connected to one another, whether they are local or distant.

Interpreters are located either together in interpreting booths, in the meeting room itself, or in another location, or on their own anywhere in the world.

Interpreters work in front of a computer screen, part of which serves as a software-based interpreter console. Moreover, the active speaker and/or a presentation and/or a shared screen are displayed on the computer screen. The interface also contains a chat function, allowing users to communicate with each other. A second screen can be used to display visual content and/or meeting documents.

Listeners can follow the meeting via audio and sometimes video in a language of their choice, on their computer or smart device, wherever they are. Although conventional interpreting and language distribution equipment is not required, it can integrate seamlessly with software-based systems.

4. PLATFORMS

4.1. Choice of the platforms to be tested

Publicly available information of twenty-one platforms was analysed. Platforms providing only sound but no image to interpreters, only offering bidirectional interpreting, excluding multiple language channels and relay, or only offering sign language interpreting, were disregarded.

On the basis of this analysis, twelve manufacturers of platforms received a detailed questionnaire, enquiring into the platform's technical specifications. Nine replied, some of them stating that they were not interested in participating in the tests.

Based on the answers received, four platforms were selected for the tests. These platforms operate on the providers' servers, but they can also be installed locally (i.e. on the Commission servers), claimed to provide sufficient sound and image quality, sufficiently low latency (delay), good lip sync, a sufficient number of simultaneously active language channels, encryption of communication, and availability of a technical support desk.

4.2. Security and data protection

Given that platforms operate in the cloud, security and data protection (including GDPR) need to be carefully considered and further investigated.

5. TESTS

5.1. Limitations

The tests were fully focused on user experience, and did not contain any technical measurements on sound and image quality, latency and lip sync.

The SCIC Technical Compliance Team drew up the questionnaire aiming to collect information about the interpreters' experience. Some questions might have been ambiguous or understood in a different way than intended.

The platforms provide features, which were not included in the questionnaire, as they were not known in advance. However, for example the handover function, which was not by itself covered in the questionnaire, was the object of many spontaneous remarks.

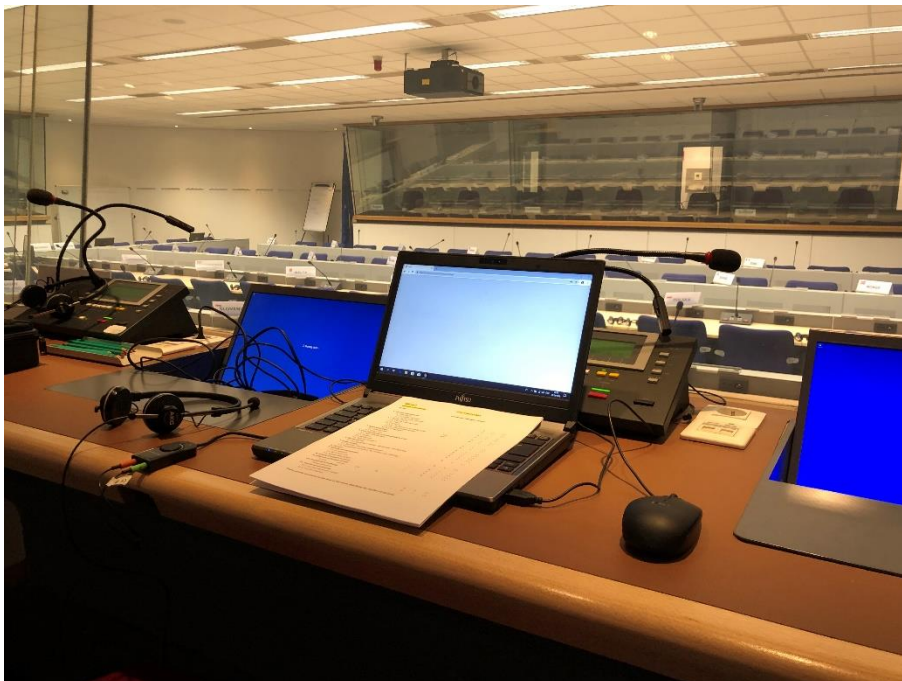
All tests were conducted with laptops with a Wi-Fi connection. This is not the providers' preferred option, but was the solution SCIC used for testing due to self-imposed criteria.

Many respondents only answered some of the questions. Totals therefore do not add up.

5.2. Equipment

5.2.1. Laptops

Eight laptops provided by SCIC C.5 were used. These laptops did not contain the usual European Commission security features and could therefore not be connected to the internal cabled network, but only to the EC_Guest Wi-Fi. In this way, the possible blocking of the websites or the webcam by DIGIT was avoided. All platform providers favour a cabled over a Wi-Fi connection, for the sake of stability.



However, it turned out that even with a fully secured Commission laptop it was possible to access all the platforms' websites, both with a cabled and a Wi-Fi connection.

The laptops were of an older type, and their processing speed relatively low (Intel Core I3). Platform providers favour a faster processor, but this was not reported as an issue.

The laptops had an unsuitable cooling system. Interpreters mentioned that the frequently activated fan was noisy and disturbed their attention.

5.2.2. Headsets

AKG HSC15 headsets were used by interpreters and speakers. These headsets comply with ISO 20109:2016 and are used daily by SCIC interpreters.

For the tests, headsets were not connected via the customary TRSS 3.5 mm jack, but a USB adapter.

In fact, platforms providers consider headsets with a built-in USB connector safer, because a single connector is necessary, and the ADC-DAC conversions, bypassing the device's built-in sound card, are faster.

On the other hand, the chosen adapter provided a physical rotary volume control, which the interpreters found very useful.



5.2.3. Built-in laptop cameras

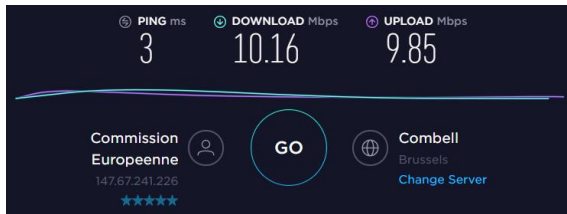
Speakers only used the built-in laptop cameras.

5.2.4. BYOD

Listeners were invited to bring their own device, be it a laptop, smartphone or tablet, to connect to the platforms. They also used their own headphones or earbuds. Specifications are unknown.

5.2.5. Network connections

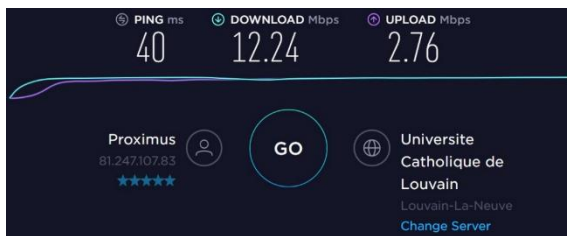
EC_Guest Wi-Fi:



Best private cabled connections:

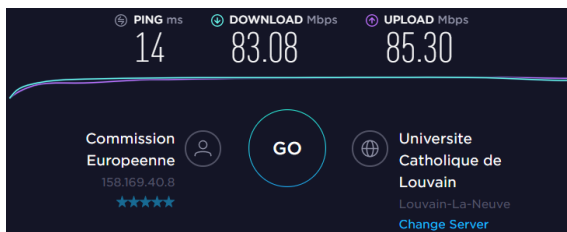


Worst private Wi-Fi connection:



In these cases, it was clear during the tests that the upload speed was not enough to ensure a good sound and image transmission.

For completeness' sake and for reference, the EC standard cabled connection:



5.3. Settings

Different settings were tested, from a setting with all actors in the same room, where the only specificity was that interpreters and speakers used a software-based interface, and listeners used their smartphone or tablet, to a setting with all actors at different sites.

- Interpreters
 - In a booth per two interpreters, three booths in the same meeting room
 - In booths in different rooms in the same building
 - Inside the meeting room
 - Elsewhere (at home)
- Speakers and listeners
 - In the room
 - In the same building
 - Elsewhere (at home)

5.4. Overall analysis of results

5.4.1. General

It is very important to note that the testing environment was very different from a standard meeting with interpretation:

- There was no interaction among speakers and participants;
- The speakers read prepared speeches on various unrelated subjects;
- There was no meeting title, no common subject and therefore no fluidity in the discussion;
- Obviously, there was no documentation and no meeting preparation.

Despite these particularities, the replies demonstrate that most interpreters assessed the platforms in the light of their daily work as staff interpreters (i.e. large multilingual meetings / long events / conference settings with participants negotiating and aiming to reach agreement). This mismatch between the "meeting" and the interpretation was a necessary limitation of the particular tests. It will need to be addressed in future tests.

Moreover, interpreters were not encouraged, as it was not part of the tests, to imagine other use cases where the platforms might be the adequate instrument to satisfy a different kind of demand, which might be unknown to them or yet not expressed by clients. It is therefore clear that the question of the use cases needs to be analysed further, ideally with the interpreters' involvement.

The platforms and their features are assessed as being quite different. Two of them are praised for their completeness, with many functions, which are also available on traditional interpreter consoles. The two other ones have fewer features, which can help interpreters focus more on their work. This question needs to be addressed as part of the standardisation work.

5.4.2. Network connection

In view of establishing a stable connection between the platforms and the devices used by interpreters, speakers and listeners, all providers strongly recommend a cabled connection with sufficient upload and download speed. To avoid possible problems with the platforms not being allowed on the EC corporate network, all tests were performed with Wi-Fi, which no doubt had a very clear influence on the quality and stability of the connection, and therefore on the quality of the sound and image transmitted.

It is therefore clear that this part of the test is inconclusive. Future tests should use cabled connections to provide the best possible quality of the input, which the interpreters use. Objective tests (not of perceived audio and video quality) should also take place.

5.4.3. Sound

Sound is the most crucial element for the work of an interpreter. There were therefore many replies concerning the quality of the sound. At the tests, interpreters noticed that the sound was sometimes tinny or unnatural. They also regretted the lack of treble and bass adjustments. In general, they did not consider that the audio was comparable to audio transmitted by a traditional system. The difference in the sound level between the different sources (speakers, other interpreters) is disturbing, both for interpreters and listeners.

On the other hand, interpreters often assessed the audio quality as sufficient. However, they gave very different indications about how long the interpreter would want to work with this audio quality (between 20 minutes and half a day).

It has to be noted that the audio quality can be improved if cabled connections are used. Therefore, more tests are needed using cabled connections while objective technical tests should also take place. Depending on the use cases, the audio quality should be compared also with the quality of telephone connections, the PIE ("bidule") and whispering/doing consecutive without equipment.

5.4.4. Image

The image quality is fluctuating. If the speaker were present in the room where the interpreters are working, the image would seem less important. However, since all controls and indicators are displayed on the same screen, interpreters tend to keep looking at it. This might of course change when interpreters become accustomed to using one platform.

Full screen mode is appreciated, but then the image sometimes hides (part of the) controls, which is a source of discomfort if they are urgently needed.

So, image was deemed to be useable but possible improvements should be considered, in particular if cabled connections are used, if objective tests take place and if interpreters receive training before using a platform.

5.4.5. Lip sync

The delay between sound and image is disturbing if too long. It is less reason for remarks, though. It is clear that working for too long with a clearly perceptible lag between sound and image causes nausea and visual fatigue. However, this is not the case with most of the tested platforms.

5.4.6. Interfaces

Many remarks are made about the design and user-friendliness of the interpreter's interfaces, ranging from the layout of the controls, the size of the microphone control, the highlighting of an active microphone to the brightness of the colours and the size of the fonts.

Complete interfaces resemble a console, but simple interfaces distract less and increase focus.

In terms of accessibility, none of the tested platform interfaces allows visually impaired interpreters or speakers to work autonomously.

The look of one of the platform's interfaces resembles an interpreter's console, offering a sense of familiarity. As it is mostly in black and grey tones, it is less tiring to the eye for some. Others find it confusing that the interface suggests buttons, which then operate quite differently from a hardware control button.

Pop-up screens with their sound indicator are experienced as distracting. Moreover, they might hide (part of) the image of the active speaker.

Interfaces are prone to human errors. It is for instance very easy to accidentally close the web browser window.

One can assume that a lot of the errors and other issues result from the fact that interpreters had to use four different platforms with very different interfaces without having the time to familiarise themselves with them or receiving any prior training.

So, further tests are necessary, in particular if standardisation advances and focuses on the interpreter's interface. Moreover, most problems can be overcome if interpreters are trained to use a platform and become familiar with its layout.

5.4.7. Reactivity of controls

Several times, the slow reactivity of the controls, including the microphone control and the switching between incoming channels (floor & relays) is reported. It is perceived as a source of additional strain on interpreters' attention.

This issue has to be addressed at the standardisation process and requires further testing.

5.4.8. Switch/Handover

This feature is indispensable for interpreters when they are not located together, to hand over the microphone to their booth partner(s). The function was reported to be complicated, long and not intuitive enough. Several interpreters complained because the switch/handover could only be initiated by the active interpreter, and not by the one about to take over.

To facilitate the switch/handover, platforms transmit the interpreting from the active booth partner as a voice over the floor. Not being accustomed to this, interpreters found this feature disturbing when the partner was sitting in the same booth but necessary when located elsewhere.

Please note however, that only two platforms allow interpreter's microphone override.

The switch/handover is an innovation of the platforms, necessary when interpreters are not in the same location. On top of further testing and the improvements, which can come out of the standardisation process, it is an issue, which needs to be addressed when discussing the use cases.

5.4.9. Chat function

Direct communication with the booth partner(s), moderator, speakers and technical support desk is vital, especially if not all actors are at the same location; hence a chat function. It is an additional cognitive load and a great source of distraction. Understandably, interpreters reported that they could not deliver quality interpreting while using the chat. Moreover, it does not replace mutual support of partner(s) in the same booth (documents, numbers, terminology, etc.).

Only one platform offers a predefined list of outgoing messages, which can be used semi-automatically in the chat.

It was not always possible to send a message to all interpreters.

So, the impact of chat function should be discussed when defining the use cases. More work is needed to define how it can be improved to help interpreters instead of hindering their concentration.

5.4.10. Working as if with portable interpreting equipment

One setting which was tested was when interpreters worked inside the same room with the participants (speakers and listeners). This setting mainly gave rise to the following two remarks.

The meeting room used was rather small; noises generated by participants in the room walking around and talking, were found to be disturbing but not related to the use of the platform. On the other hand, the platform generates a delay that was considered too important between the sound coming directly from the speaker and via the headphones. Interpreters therefore preferred to listen to the speaker directly.

So, this setting should be discussed when defining the use cases in particular in comparison with the use of PIE ("bidule") which is used when interpreters and participants are in the same room.

5.4.11. Working from home

Working from home is new for SCIC interpreters. As they normally work in a booth in an EU building, they found themselves feeling alienated and became easily distracted, even by just being in too familiar an environment.

Every little home noise became a distraction and made concentration difficult. Noises like a clock ticking, traffic, the fridge humming seemed magnified and disturbing to concentration. Seeing all one's stuff around distracted; a separate office might be necessary. It was remarked that the table and chair used did not offer the required comfort.

It turned out to be stressful to manage different chat channels, microphone settings, switch/handover requests, etc. at the same time. Moreover, having to focus on/stare at the same screen all the time was assessed as very hard on the eye, greatly increasing visual fatigue compared to a physical booth where the interpreter would be gazing, in turn, at delegates (if present) through the booth window, the interpreting console, a screen showing the speaker or a presentation, and paper documents.

So, this setting needs to be addressed when discussing the use cases. Moreover, the working environment should also be discussed.

5.4.12. Impact on interpreters in tested conditions

Remarks about the interpreter's wellbeing are not bound to a certain platform, whereas the look and feel, and the ease of use, are detailed for each platform individually.

Many interpreters reported screen fatigue, which increased when the day advanced. The many visual inputs (not only the image of the active speaker, but also the controls and indicators and of course the chat function) complicated the interpreter's tasks. Interpreters received many more stimuli than in a booth setting.

Stress and fatigue were also aggravated since four different platforms were tested on one day. Each platform being different, this increased the learning effort.

Getting accustomed to the new tools would certainly decrease stress and fatigue. At this time, the use of the platforms is said to be conceivable for a short duration (going from 20 minutes until half a day at a time).

5.4.13. Nice-to-haves

Interpreters mentioned a few features they felt were missing on some or all platforms.

Examples are:

- Indication of relay status, informing them if interpreters taken on relay work from the floor or are themselves working from relay;
- Audience indicator (available on some platforms or for speakers/moderator only);
- Document, presentation and screen sharing (available on some platforms);
- Request list with speakers' names (available on some platforms) and their languages;
- Layout of the interpreter interface should depend on configuration (e.g. if the booths partners are not at the same location, switch/handover and chat functions should be activated and/or more prominent);
- Image of the booth partner would facilitate switching/handling over and problem solving;
- Touchscreen would facilitate operating the interface; and
- Team sheet.

5.4.14. Speakers

Speakers delivered their contribution in front of a laptop computer, where the built-in camera captured their image; and the headset microphone captured their voice. The fixed camera obliged the speakers to sit rather still in front of the laptop; the interpreters did not like speakers to move too much.

Speakers did not like to have to look at the camera, and hence the screen with their own image, during their entire speech. Concentrating on speaking, they often did not see the request to slow down available on one of the platforms coming from the interpreters.

5.4.15. Participant's Apps

Participants installed a free iOS or Android application per platform on their own devices, be it a smartphone or tablet. For one of the platforms, the Android app did not work. Manipulation of the apps was feasible to easy. Participants deplored the absence of video sources in most of the platforms; not seeing the speaker was assessed as alienating for the listeners. Two platforms offer the possibility to participate; only one offers extra features like messaging, accessing documents and voting.

6. INFORMED OPINION

From the testing, it clearly appeared that the following elements need careful attention in document ISO 24019:

- A number of basic features have to be available for the interpreters on all platforms:
 - Microphone control
 - Volume control
 - Treble and bass adjustments
 - Preselection and choice of incoming and outgoing channels
 - Communication tool
 - Switch/handover function
- What images should be available to the interpreters:
 - On one or two screens
 - At interpreter's choice:
 - Speaker
 - Presentation
 - Documents
 - Which controls should remain accessible in full screen mode
- The controls should react immediately, without any delay
- The volume from the floor and from the interpreted channels should be equal (Level consistency)
- The communication tool is indispensable, especially if interpreters are not collocated
- The switch/handover function is indispensable, especially if interpreters are not collocated
- Security and data protection are paramount