PRESS REVIEW 2020

CONTINUATION OF COLLECTIVE BRAINSTORMING ABOUT THE JITSI PROJECT AIIC SWITZERLAND

RSI: Defending against unwanted intermediaries

by Maxwell Crisp | 25.11.20 | Version française par Aymeric de Poyen Bellisle

Interpreters are professional intermediaries. One corollary of our professionalism is that we are vigilant in ensuring there are no third parties between ourselves and our clients. This has in the past been the case with agencies, but today the problem has resurfaced with a vengeance in the form of companies providing interpreters over video-conference (VC) platforms.

Although it is pure economic arguments that motivate agencies, it is public health and technology that is behind this new threat. For some years now, audio and later audio-visual remote participation to meetings has been greatly facilitated by the expansion of broadband internet infrastructure. To an outsider, it is perhaps natural to assume that interpretation could also be provided by this means. Indeed, companies providing remote simultaneous interpreting (RSI) platforms have been emerging on the scene for some time now in an inept attempt to stand out in the saturated VC market place. They had captured a niche market, servicing such public relations stunts as panels held on mountain tops and press conferences in otherwise confined spaces, by proposing the 'formula' of platform plus interpreters.

Fast-forward to 2020, the Covid crisis and the impossibility of meeting in person. Our profession is flipped on its head and what was once the exception now becomes the rule. All participants must join the meeting remotely, and VC platforms quickly became in raging demand, dragging with them the RSI platforms.

It was hardly a consolation to know that everyone was affected more or less equally, as the threat to individual livelihoods was as great as it was to the profession itself. What were we then to make of the idea that meetings were being held on the internet? Without us? Over Zoom? Was it a glimmer of hope or the coffin about to be nailed shut?

It came as no great surprise, with panic-mode engaged, that all new actors were seen as potential threats. But, as the dust now begins to settle, it has become clear that the greatest particular threat didn't come from the Covid vedette Zoom, but from those RSI outfits providing the formula. It has also become clear that virtual meetings don't come cheap, and clearly hybrid meetings only increase costs. But this places further downward pressure on prices, especially where the added value of a platform is that it permits simultaneous interpretation. Service providers proposing the formula benefit from the possibility of offsetting the costs of providing the IT infrastructure by recruiting cut-price interpreters. The mathematics are simple: if interpreting used to cost 100, but a platform costs 50, then by paying interpreters 50, the package can be offered without any increase in the price... BOGOF!

On a serious note, however, the question arises as to what to do when asked by a client to provide services over a platform. And perhaps even also to provide a platform itself. Of course, it will depend on a number of factors. But broadly speaking, we are reduced to calling or emailing and requesting a quote. This allows us to retain control of our prices, but it comes with the baggage of being responsible for the choice and quality of platform. And how do we prevent the platform from poaching the client? A contract might work, but not forever. At any rate, the cost is to the client is 150.

The proof of the pudding is, however, in the tasting, and we are all aware that poor quality audio, and to a lesser extent video, leads to poor quality interpretation. In other words, the client is not getting value for money. But boil that down and you find that it exerts yet more downward pressure on prices.

Even if we set aside the economics and take up purely technical standpoint, the thorny issue of ensuring high quality audio in virtual meetings remains. There is no quick fix. Each participant is in a unique acoustic and technical environment and

each signal is taking a unique route across an uncontrollable and highly complex global computer network. Enforcing standards for microphones and internet connection is a start, but it does nothing to tackle echo and background noise. A broadband internet connection is good, but the benefit is lost if the client's computer is outdated, has a noisy cooling fan or is part of a battery of machines all sharing that same connection.

At least one RSI platform enforces microphone and to a lesser extent headset standards.... but only for interpreters, and not for participants. Elsewhere, international organisations that have been relying on RSI platforms have reportedly taken to addressing this issue and working with participants to ensure better quality audio, which is after all, to the benefit of all concerned. But it's indicative of the business model of RSI platforms that their focus is limited to the quality of sound coming from interpreters and not the sound going to them.

Then there's the cogs and wheels of the platform itself. Some are optimised for free flowing conversation at the expense of close lip-sync and of high bit rate audio and video, which are often downsampled precisely so as to allow a discussion to continue in spite of unfavourable network conditions. Others prefer to opt for a streaming approach which does allow for high bit rate audio, without guaranteeing it, not to mention the issues of latency and lost connections.

But we cannot lay all the blame at the doors of the RSI platforms. They are, in the main, reliant on existing technological development, which is after all a bottom up process. Boding well for the future, quality has been rising with the improvements in infrastructure and in the algorithms that convert/compress sound into computer signals ready to be routed across the internet. And this is really the crux of the argument, as there are two types of compression algorithm (known as 'codecs', a portmanteau of coder-decoder): lossy and lossless, similar to the difference between compact discs and mp3s. Streaming-based RSI platforms can opt for lossless compression, which will only ever be as good as the sound going into it, but at the expense of free-flowing dialogue. In more formal settings, this becomes a negligible problem, especially if one can afford the extra staff it takes to ensure good management of up-coming speakers. Furthermore, as the weakest links in infrastructure begin to meet minimum standards, then it will be

possible to shift permanently to lossless compression. Improvements in neural networks are also beginning to offer possibilities to reconstruct audio that has been through a lossy compression process, leaving just the acoustic environment to be tackled.

As an interim conclusion, we can state that although serving the interests of interpreters may be to the benefit of all concerned, it does not seem to be high on the list of priorities of the other parties, despite the length of the crisis. This alone is a call to action.

There are are number of reasons why we as a profession should… perhaps not embrace, but, at least, greet warmly this new technology, at least in the context of this crisis and potentially of the climate one. The most obvious reason is well known to those lucky few of us who have been able to work at all this year.

We are well placed to improve it, should it ever be needed again. Our high levels of professionalism would bring an objective voice for quality at (virtually) all costs. Indeed, we are the only party in a position to identify the quality criteria, thanks in no lesser part to our hands-on knowledge. There are also the opportunities that it would bring, particularly in the current setting. First, it is an opportunity to build upon client relationships, to participate in design and development processes and thereby learn the ins and outs of the systems, which may prove invaluable when it comes to, for example, on-the-spot troubleshooting (I'm reminded of a hot mic in the technicians' booth broadcasting the sound of cooling fans).

Thinking more broadly, it is also possible that the general adoption of virtual/hybrid meetings leads to a market expansion, bucking the trend of several decades of market contraction. In this particular case, the opportunity is to set the gold standard and lead by example. But perhaps the most compelling argument is being prepared for the next crisis, the identity and nature of which will be until the very last minute a closely guarded secret: a cloud of ash, a global conflict, nuclear winter?! Your guess is as good as mine.

On that sombre note, it behooves us to address the reasons why we should stay out. Our weaknesses include a general lack of technical knowledge, although it

doesn't have to be this way, especially considering the 'language' aspect of computer programming. There is also the long-standing position taken against RSI. A volte-face would perhaps result in us being put in the back seat, where our arguments of the past are used effectively against us. On a more practical note, there are conflicts of interest that arise. How can an individual react when they are concurrently responsible for the quality of the sound and the quality of the interpretation? Likewise, there are questions of cost. It would constitute a sizeable investment and would be a white elephant in the happiest of scenarios.

Getting involved comes also with its threats. Admittedly the threat from RSI platforms already exists, but it would exacerbated were we to put our expertise to work and thereby expose it to potential appropriation. As with the conflicts of interest, there is a threat to quality and thereby client satisfaction, and we would be complicit in it. Of course, the way things stand, we are free to step back and wipe our hands clean of all responsibility for the technical hitches with which we are all now so familiar. But isn't the desire to stick our head in the sand perhaps the greatest threat? It's all very well to say, 'hybrid meetings are expensive and glitchy so they'll go away', but what if they don't?

Adding further to the overall cost would be the on-going, indeed never-ending, development, updating and adapting to changes in technology, patching security flaws and so on, all of which constitute a long-haul commitment. This might be summed up as a risk of mission creep: what starts out as a desire to guarantee high quality audio may end up as the micro-management of conference participants or the pursuit of outlawry for certain codecs.

In brief, as a platform is intangible, unlike a conference room or a portable booth, the capital costs are low. It can be deployed on cloud infrastructure allowing individual professionals, fully fledged organisers or otherwise, to offer a solution to their clients on an ad hoc basis, much in the same way they might provide booths or bidules. International organisations in Geneva and Paris that opened informal channels during the crisis were clear that turn-key solutions were the only service in which they were interested. But the current market offer is two broad categories, as alluded to above: those that work well with poor quality sound, and those that work poorly with sometimes slightly improved sound quality. In times of crisis there is a certain level of patience that will evaporate with time, at which

point the tables will turn and we will either return to how things were or turn to an even greater unknown.

In the meantime, we need to defend the space between the profession and its clients: be it through simple outreach regarding microphone discipline and ambient acoustics or be it through the development of virtual conference technology that might one day achieve the levels of quality to which we were accustomed at the end of 2019.