REMOTE SIMULTANEOUS INTERPRETATION (RSI) PLATFORMS AUDIO SCENARIOS

AIIC-CANADA, OTTAWA, 4 APRIL 2020
JEAN MARC LARIVIÈRE

- Remote simultaneous interpretation or RSI has many technical components: hardware, software, communication networks, compression codecs, ISO standards, etc.
- Along side, there are also numerous human components : human biology, intelligibility index, ergonomics, health and safety, etc.
- All these apply to both AUDIO and VIDEO components of RSI and things get complicated rather quickly.
- This presentation is meant as a primer ONLY ABOUT THE AUDIO COMPONENTS OF RSI platforms and STRICTLY ABOUT THE INPUT TO THE INTERPRETERS, not the return signal from interpreter to audience.
- This is NOT a presentation about individual platforms, nor about RSI guidelines and recommendations, AIIC's or others.
- WHY ? I feel it is paramount that interpreters must become familiar with the basic audio configurations of RSI before making decisions about it and its role in our future.
- PUT SIMPLY :   we can't interpret what we don't hear clearly.
- For the sake of clarity, I have simplified and made generalizations that I will attempt to qualify.
- To those of you who are more technically savvy, I beg your indulgence.

A CHAIN IS
ONLY AS STRONG
AS ITS WEAKEST LINK

- If you remember just ONE thing from my presentation, it is THIS : a chain is only as strong as its weakest link.

BASIC AUDIO RSI CONFIGURATION

PARTICIPANTS' DEVICES → **LINK** → **PLATFORM** → **LINK** → **INTERPRETERS' DEVICES**
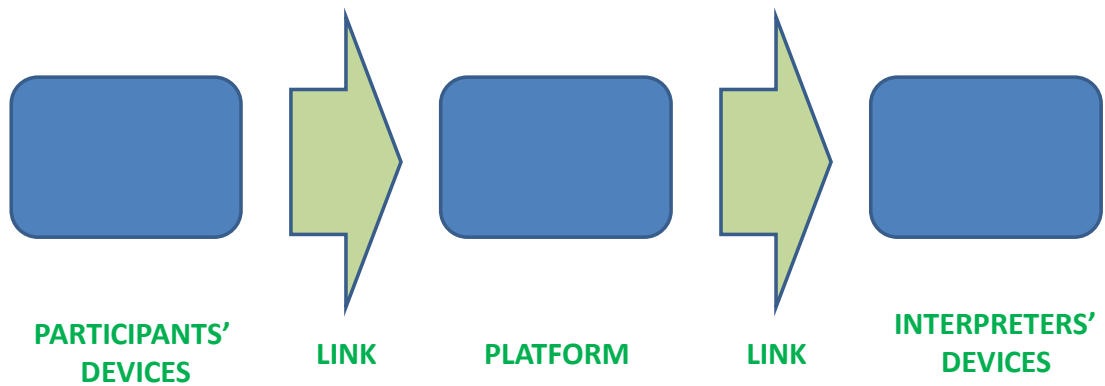
- At the heart of all RSI systems is a PLATFORM.
- The platform is intermediary between the participants' devices and the interpreters' devices.
- Participants and interpreters connect to the platform via a telecommunication network.
- So far, so good, I'm willing to bet everyone is still with me.

PLATFORMS

- **2 TYPES OF PLATFORMS**
  - **TELEPHONE CONFERENCING**
  - **VOICE OVER INTERNET PROTOCOL (VoIP)**

- There are basically 2 TYPES OF PLATFORMS:

  - telephone conferencing, such as : Bell Conferencing (in Canada)
  - voice over internet protocol (VoIP), such as : WebSwitcher, Interprefy, Kudo, etc.

LINKS

- **3 TYPES OF LINKS**
  - **WIRED LEGACY PHONE SYSTEMS**
  - **CELLULAR NETWORKS**
  - **DATA NETWORKS**

- There are basically 3 TYPES OF LINKS:

  - wired telephone systems (land lines)
  - cellular networks
  - data networks (the Internet is just a kind of data network)

- We are all familiar with the good old copper wire phones...
- ... well, at least some of us are familiar with it...
- ... and some may have already jettisoned their wired service.
- Younger colleagues may never have had a wired phone.
- Furthermore, a growing number of workplaces have trade wired phones for smart phones or VoIP systems altogether.

- PLEASE NOTE : Smart phones use both the cell network AND the data network.
- It is critical to know these are not the same.
- As a matter of fact, we get charged differently depending on which one we use.

- Whenever we dial a phone number on a smart phone, we are connecting to the cell network.
- Whereby when we use the internet on a smart phone, we are connecting to the data network.

- PLEASE NOTE : both can carry voice signals.

AUDIO FREQUENCY RESPONSE

- **TELEPHONE NETWORKS   =   4 KHz**
- **CELL PHONE NETWORKS   =   4 KHz**
- **DATA NETWORKS (in theory) = 15 KHz +**

- SO WHAT'S THE DIFFERENCE ?
- In two words : frequency response, that is, the range of sounds a device or network is capable of reproducing or transmitting.
- The human ear's frequency range is 20 Hz to 20 KHz.
- Just a couple of years ago, the ISO standard for all interpretation equipment was12 KHz.
- Recently, because of the masking effect, i.e. the sound of the interpreter's voice interfering with their capacity to ear clearly, this was increased to 15 KHz.
- When wired phones were invented, it was thought that 4 KHz was enough to make conversations understandable.
- To this day, phone and cell phone network providers have stuck to the 4 KHz frequency range …
- … this, despite the fact that your smart phone's microphone and audio circuitry is capable of much higher fidelity sound.
- One only needs to plug a good quality headset in a smart phone and listen to music to confirm that.
- On the other hand, if you use a voice app with your smart phone, you'll be connecting to the data network which is capable of a much higher frequency response.

ISO COMPLIANT RSI
Frequency response = 15 KHz

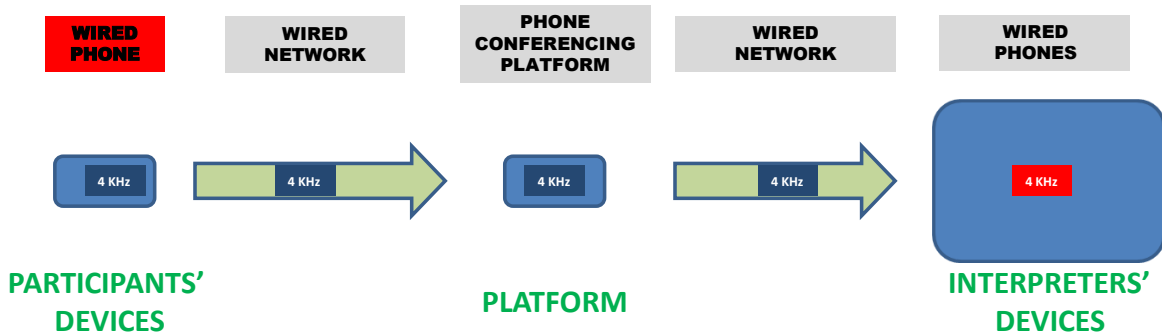PARTICIPANTS' DEVICES — PLATFORM — INTERPRETERS' DEVICES
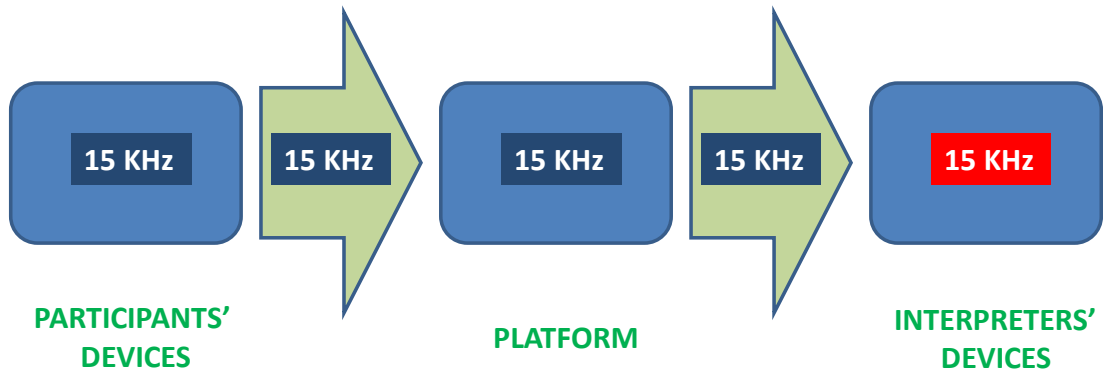
- As I pointed out earlier, ISO standards for sound is 15 KHz, both for traditional and remote interpretation
- 15 KHz throughout the chain, from participants' devices to the interpreters' headset.

PLEASE NOTE : In this and the following graphics, the size of the boxes and arrows graphically illustrates the frequency range capability of the devices, network or platforms.

The effective frequency range for each depending on the scenario is indicated in the inside box.
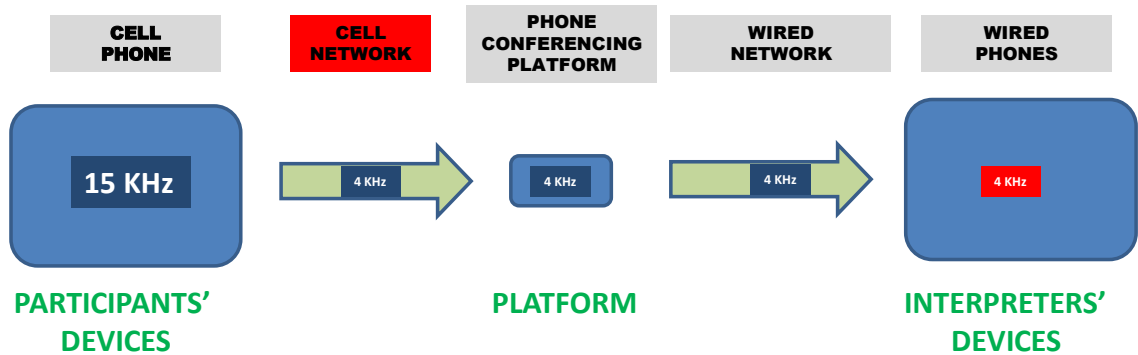
OVER THE PHONE INTERPRETATION (OTI)
WIRED PHONE SCENARIO

WIRED PHONE	WIRED NETWORK	PHONE CONFERENCING PLATFORM	WIRED NETWORK	WIRED PHONES
4 KHz	4 KHz	4 KHz	4 KHz	4 KHz

PARTICIPANTS' DEVICES　　　　**PLATFORM**　　　　**INTERPRETERS' DEVICES**

OTI SCENARIO 1 : participants connect to conferencing platform with wired phones (land lines)

- Wired phone frequency response = 4 KHz, hence weakest link.
- Interpreters received lo-fi 4 KHz sound, even though they may be using high quality headsets.

OVER THE PHONE INTERPRETATION (OTI)
CELL PHONE SCENARIO

CELL PHONE	CELL NETWORK	PHONE CONFERENCING PLATFORM	WIRED NETWORK	WIRED PHONES
15 KHz	4 KHz →	4 KHz	4 KHz →	4 KHz

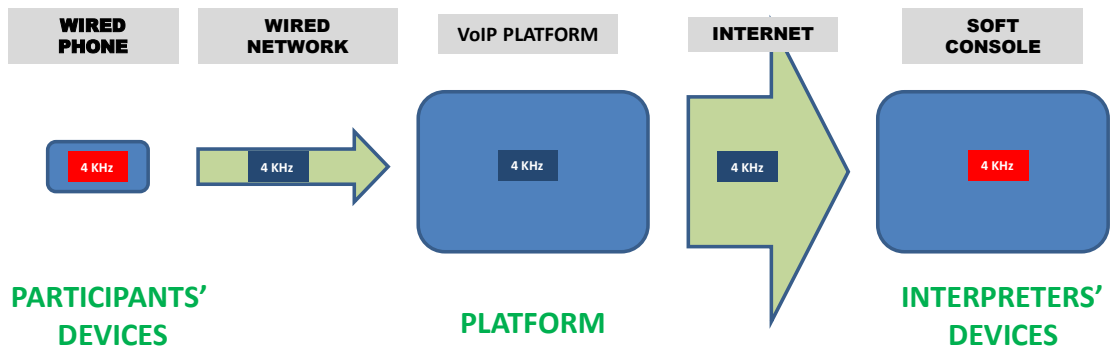**PARTICIPANTS' DEVICES**   **PLATFORM**   **INTERPRETERS' DEVICES**

OTI SCENARIO 2 : participants connect to conferencing platform with cell phones

- Though the cell phone's microphone and circuitry produces a sound whose frequency range is much higher than 4 KHz, the cell network (weakest link) throttles it down to 4 KHz and it's 4 KHz from thereon right to the interpreter, AGAIN notwithstanding the quality of the interpreters' device.

VoIP RSI
WIRED PHONE OVER PHONE NETWORK SCENARIO

WIRED PHONE	WIRED NETWORK	VoIP PLATFORM	INTERNET	SOFT CONSOLE
4 KHz	4 KHz	4 KHz	4 KHz	4 KHz

PARTICIPANTS' DEVICES

PLATFORM
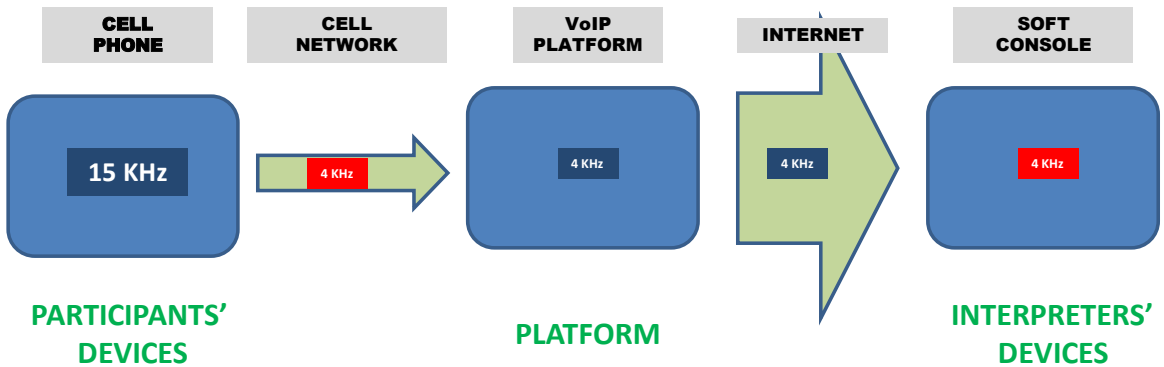
INTERPRETERS' DEVICES

VoIP RSI SCENARIO 1 : participants connect to VoIP platform with wire phones (land lines)

- In theory, VoIP platforms are capable of a frequency response of 15 KHz or higher…
- …but sound quality to interpreters depends entirely on the devices and networks participants use to connect to VoIP platform
- In the scenario above, participants use land lines and therefore the sound quality will be lo-fi.

VoIP RSI
CELL PHONE OVER CELL NETWORK SCENARIO

CELL PHONE	CELL NETWORK	VoIP PLATFORM	INTERNET	SOFT CONSOLE

15 KHz → 4 KHz → 4 KHz → 4 KHz → 4 KHz

PARTICIPANTS' DEVICES   PLATFORM   INTERPRETERS' DEVICES
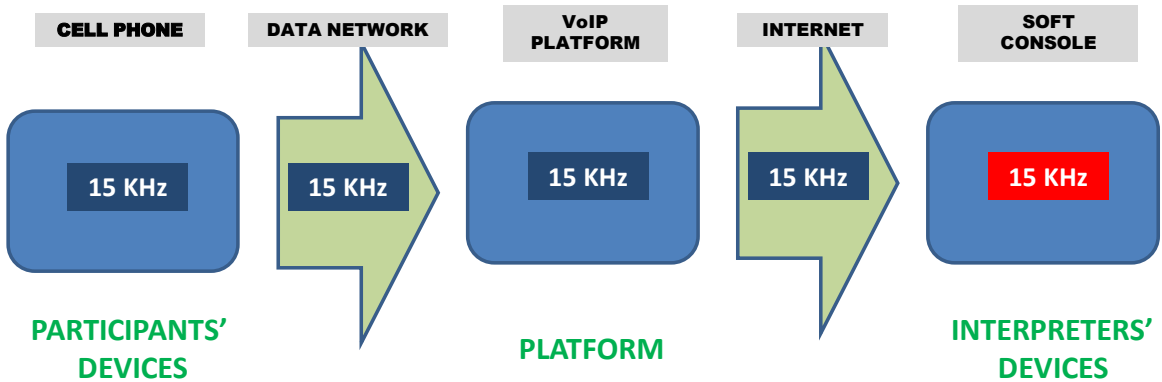
VoIP RSI SCENARIO 2 : participants connect to VoIP platform with cell phones

- A smart phone connected to a headset or earbud equipped with a high fidelity microphone is capable of producing a high quality input signal...
- ... BUT, as I pointed out earlier, if you dial into the cell network to connect to the platform, the network will throttle the signal down to 4 KHz or below depending on network traffic.

VoIP RSI
CELL PHONE OVER DATA NETWORK SCENARIO

CELL PHONE	DATA NETWORK	VoIP PLATFORM	INTERNET	SOFT CONSOLE
15 KHz	15 KHz	15 KHz	15 KHz	15 KHz

**PARTICIPANTS' DEVICES** — **PLATFORM** — **INTERPRETERS' DEVICES**

VoIP RSI SCENARIO 3 : participants connect to VoIP platform with cell phones via data network

- On the other hand, a smart phone connected to the platform via the data network is capable of producing a high fidelity sound input that will carry right down to the interpreters...
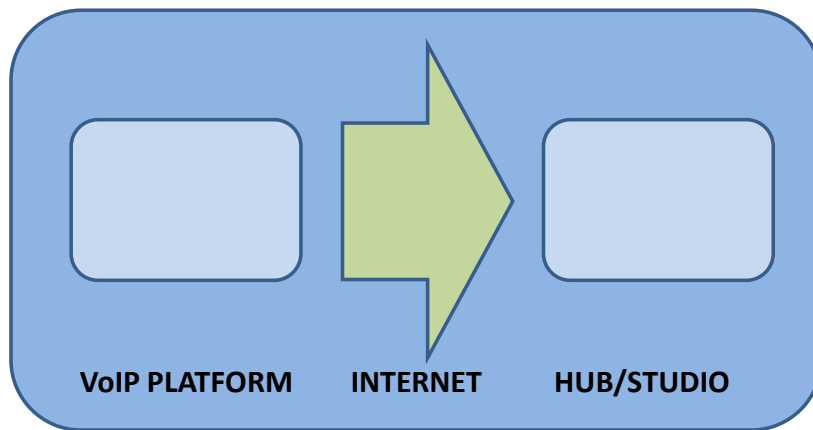- … as long as there is sufficient bandwidth.

VoIP RSI
COMPUTER OVER INTERNET SCENARIO
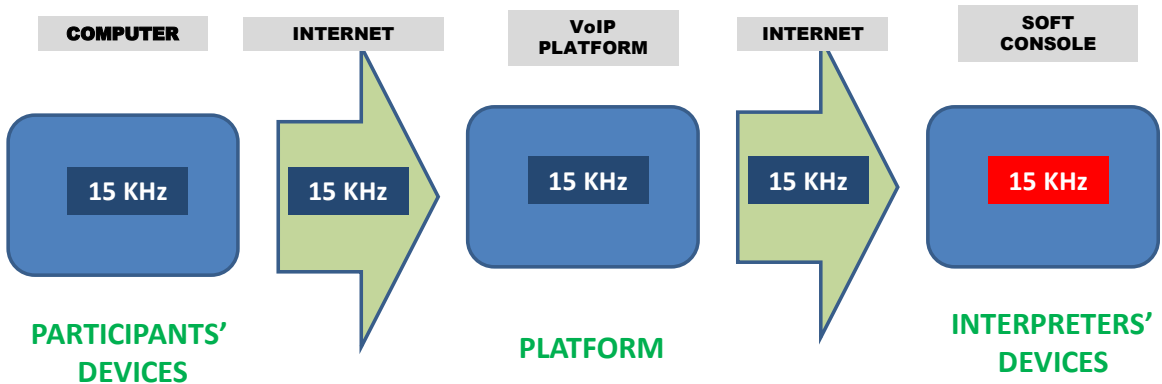
VoIP RSI SCENARIO 4 : participants connect to VoIP platform with computer via internet

- When participants connect to the platform with a computer, they do so via the internet which is capable of carrying high fidelity sound.

PLATFORM – HUB/STUDIO

VoIP PLATFORM   INTERNET   HUB/STUDIO

- In all the preceding OTI and VoIP scenarios, the interpreter is working from home or from an office.
- But what about the RSI hub/studio scenario?
- Unless a hub is hosting its own plaform (which is a rarity), the hub is itself connected to a VoIP platform via the internet.
- Hence, the middle box labeled PLATFORM in previous scenarios can now be thought of as illustrated here.
- In such a scenario, it goes without saying that the quality of the internet connection between the platform and the hub is as critical as quality of the participant/platform and platform/interpreter connections.

- A WORD OF WARNING: In the appropriate scenarios, a hardwired connection between a participant's device and router, and between the interpreter's device and router is a MUST.
- Wi-fi connections are HIGHLY INADVISABLE.
- In cases where this may be physically difficult, a very good and inexpensive solution is to use powerline adapters that use an office or home electrical wiring to hard wire devices directly to the router.

TAKE AWAYS

- SYSTEMS ARE AS STRONG AS THEIR WEAKEST LINK
- WIRED TELEPHONE NETWORK IS LOW FIDELITY
- CELL NETWORK IS LOW FIDELITY
- DATA NETWORK (WITH DEDICATED BANDWIDTH) IS HIGH FIDELITY
- VoIP PLATFORMs OR HUBs ARE NO GUARANTEE OF HIGH QUALITY SOUND
- SOUND QUALITY PROVIDED BY VoIP PLATFORMS DEPENDS ON THE QUALITY OF THE QUALITY OF INPUT SIGNAL
- DIAL IN ACCESS TO PLATFORMS = LOW QULITY SOUND
- DATA NETWORK OR WEB ACCESS = (POTENTIALLY) HIGH QUALITY SOUND