OVERVIEW OF LITERATURE ON THE EFFECTS OF COMPRESSED SOUND IN

REMOTE SIMULTANEOUS INTERPRETATION

INTRODUCTORY NOTE

This overview collates information pertaining to the potential adverse health impacts on auditory function in general and interpreters' hearing and cognitive overload in particular. The underlying rationale is that after two years of working essentially on remote platforms, a number of occupational incidents have been documented in the interpretation field, notably in Canada and within some International Organisations. These give cause for concern in a context where many meeting participants are still not using the recommended equipment or following best practice for their interventions.

Indeed, this should be a minimum requirement in a remote meeting setting, all the more so because it alleviates but **does not actually resolve** the problem of compressed sound, as detailed in the articles below. The reader's attention is also drawn to a number of **video links** in the text that demonstrate 'good' versus 'bad' sound quality (cf. articles 7 and 10 by Caniato and Giaducci respectively).

This compendium is not exhaustive but rather presents a wide-ranging overview of the impacts that currently affect ALL participants in zoom meetings, but are especially significant in a simultaneous interpreting context.

Collated by Karen Twidle, staff interpreter at OECD

March 2022.

Table of contents

1. Franceinfo, les effets observés du son compressé, Anne Le Gall, France Info, 13/01/2021

https://www.francetvinfo.fr/replay-radio/le-billet-vert/audition-les-degats-du-son-compresse-mis-en-evidence-par-une-etude-realisee-sur-des-cochons-d-inde_4898493.html

90 cochons d'Inde ont écouté de la musique compressée ou non. Bilan : ceux qui ont écouté de la musique compressée ont subi une fatigue auditive.

Un *"son compressé"* est un son numérique qui a été *"tassé"* électroniquement pour faire remonter les niveaux sonores les plus faibles, afin que ce soit plus audible.
Le son de franceinfo est légèrement compressé. Tout comme la grande majorité des musiques et des sons que l'on écoute sur internet, en DVD ou en visioconférence quand on télétravaille.

Ce format s'est généralisé car le son compressé a l'avantage de se placer au-dessus des bruits d'ambiance, ce qui permet d'avoir un plus grand confort d'écoute. Mais visiblement la compression du son n'a pas que des avantages. L'inconvénient, c'est que les oreilles reçoivent une énergie sonore plus forte et avec moins de nuances qu'avec un son classique, explique Christian Hugonnet, ingénieur acousticien et président de la Semaine du son qui se tiendra la semaine prochaine.

Des chercheurs de l'Inserm et de la faculté de médecine de Clermont-Ferrand ont donc voulu savoir si ces sons compressés représentaient un danger pour les oreilles. Pour cela, ils ont donc fait écouter de la musique pendant plusieurs heures à 90 cochons d'Inde, car ils ont un système auditif proche du nôtre. Certains cobayes avaient droit à de la musique compressée et d'autres à de la musique en format original. Ils ont écouté de la pop, du classique ou encore de l'électro, il y en avait pour tous les goûts.

le système auditif n'a plus de répit

Ces cobayes ont ensuite eu droit à un examen ORL. Bonne nouvelle : aucun n'avait perdu en audition. Mais les cochons d'Inde ayant écouté de la musique compressée ont subi une fatigue auditive pendant plus de 48 heures. Les muscles protecteurs situés à l'intérieur de leurs oreilles étaient fragilisés.

Le son compressé ne contient plus aucun silence, explique le Pr Paul Avan qui a dirigé les recherches. Il n'y a même plus les quelques millisecondes de blanc, qu'on peut trouver dans un son non compressé, ou modérément compressé, donc le système auditif n'a plus de répit, il est comme asphyxié. Certains médecins font un rapprochement entre l'augmentation de certains troubles auditifs chez les jeunes et l'écoute de ces sons, mais si cette fatigue auditive est démontrée chez l'animal, elle reste encore à prouver chez l'homme.

Quelles sont les solutions de prévention, puisque les sons compressés sont déjà partout ? Un comité scientifique est sur le point d'être créé pour voir s’il est possible, à l'avenir, de labelliser des sons moins compressés et garantis "sans danger" : ils contiendraient les quelques millisecondes de silence nécessaires pour laisser les oreilles "respirer".

2. ESSENTIEL SANTE MAGAZINE, PUBLIE LE 31/01/2022

HTTPS://WWW.ESSENTIEL-SANTE-MAGAZINE.FR/SANTE/PREVENTION/AUDITION-UNE-ETUDE-ALERTE-SUR-LES-DANGERS-DES-SONS-COMPRESSES

AUDITION : UNE ETUDE ALERTE SUR LES DANGERS DU SON COMPRESSE

À LA RADIO, AU TELEPHONE, OU ENCORE EN VISIOCONFERENCE, LES SONS COMPRESSES SONT PARTOUT. UNE ETUDE VIENT DE METTRE EN EVIDENCE POUR LA PREMIERE FOIS LEURS EFFETS NEFASTES POUR NOS OREILLES. LE PROFESSEUR PAUL AVAN, QUI A DIRIGE L'ETUDE, NOUS REVELE SES PREMIERES CONCLUSIONS.

Difficile d'y échapper ! Dans nos loisirs comme au travail, nous sommes confrontés tout au long de la journée à des musiques ou des sons compressés. De quoi s'agit-il ? Un son « compressé » est un phénomène qui consiste à **réduire les écarts entre les sons forts et les sons faibles, en le « tassant » de manière électronique**.

POURQUOI COMPRESSE-T-ON LA MUSIQUE ?

« C'est un vestige du passé, explique Paul Avan, professeur de biophysique à l'université de Clermont-Auvergne et directeur de cette étude. L'héritage d'une époque où les radios étaient analogiques et le matériel pour diffuser le son de faible qualité ». Cette technique s'est généralisée au début des années 80, parce qu'elle permettait alors d'augmenter la portée des ondes radios. En outre, la compression était un moyen de parer à la mauvaise qualité des écouteurs et des casques audios utilisés par le grand public. « Aujourd'hui, la compression ne se justifie plus » ajoute le professeur.

« Un son naturel, qu'il s'agisse de paroles ou de musique, **alterne des temps forts et des temps faibles**, explique le professeur Avan. Durant les phases de temps faibles, le bruit extérieur peut dominer ce son, voire le rendre inaudible ». Or, au cours des 20 dernières années, la diffusion de sons et de musique s'est multipliée dans les lieux bruyants : voiture, gare, supermarché, transports en commun.

Afin de conserver la qualité d'écoute de leurs clients, les diffuseurs ont eu l'idée de **transformer les sons faibles en sons forts**. Comment ? En remontant de manière électronique tous les sons en dessous d'un certain seuil de décibels. « Cette augmentation leur permet de dépasser le bruit de fond extérieur, expose le professeur Avan. En conséquence, depuis 20 ans, **de plus en plus de musiques sont compressées**. »

Certaines stations de radio utilisent désormais une compression intense, de façon délibérée. Une compression que l'on retrouve aussi lors de **l'usage du téléphone, durant une visioconférence, l'écoute d'un CD ou encore d'un lecteur MP3.**

UNE ETUDE SUR LA MUSIQUE COMPRESSEE LIVRE SES PREMIERES CONCLUSIONS

L'équipe du professeur Avan a voulu mesurer **les effets de cette musique compressée sur nos oreilles**. Cette étude est une première.

Le professeur et son équipe de l'université de Clermont-Auvergne ont donc constitués deux groupes d'une quarantaine de cochons d'Inde, à qui ils ont fait écouter de la musique. L'un des groupes écoutait un son compressé, l'autre un non-compressé. « On a fixé le niveau pour les deux groupes à 102 décibels, qui est celui d'une boîte de nuit », rapporte le scientifique. Les sujets ont ensuite été exposés à la musique pendant quatre heures, soit le temps moyen passé en discothèque.

Les cochons d'Inde ont été choisis grâce à leur système auditif proche du nôtre. « **Lorsqu'un son fort survient, une zone du cerveau appelée tronc cérébral peut déclencher des réflexes protecteurs** », explique le chercheur. En quoi consistent ces réflexes ? De petits muscles, dans l'oreille moyenne, peuvent se contracter de manière automatique et atténuer les sons qui rentrent. « Nous avons décidé d'utiliser ces réflexes comme marqueurs, afin d'étudier comment le cerveau auditif « encaisse » les sons compressés ».

LA MUSIQUE COMPRESSEE ENTRAINE UNE « FATIGUE AUDITIVE »

Autre sujet d'intérêt des chercheurs, les éventuelles pertes auditives suite à une exposition prolongée à la musique compressée. « Nous n'en avons pas observées, souligne le professeur Avan. En revanche, au niveau de la réponse cérébrale, le constat est moins positif. **Dans les deux groupes, une « fatigue » est observée à la sortie, avec une diminution des réflexes auditifs** ».

Le groupe exposé à la musique non compressée recouvre ses facultés en quelques heures. « L'autre groupe, au bout d'une semaine, n'avait toujours pas récupéré à 100 % », constate Paul Avan. Cette altération des réflexes protecteurs est un risque direct pour le circuit neuronal auditif. **En effet, les muscles ne se contractant plus, l'oreille n'est plus protégée**. « Si l'on est exposé à des sons forts la semaine suivante, par exemple en retournant en discothèque, la fatigue auditive va encore s'aggraver », alerte le chercheur.

L'étude conclut qu'une partie des voies cérébrales auditives est affectée anormalement par l'exposition à la musique compressée. « Cela peut cacher d'autres lésions, avertit le professeur en biophysique. L'étude doit se poursuivre, notamment pour étudier **les atteintes précises sur les neurones** ».

Pour Véronique Bazillaud, directrice des affaires publiques d'Ecouter Voir et Déléguée Générale de la Fondation Ecouter Voir, cette étude est importante car elle pourrait permettre au grand public de prendre conscience de ces risques. « **Tout le monde est exposé à ces sons**. Les chercheurs veulent identifier les dangers pour pouvoir ensuite faire de la prévention ». La Fondation Ecouter Voir a financé cette étude suite à un appel à projet en 2019.

COMMENT LES SONS COMPRESSES NOUS AFFECTENT-ILS AU QUOTIDIEN ?

Il n'est pas nécessaire de passer une soirée en discothèque pour ressentir la « fatigue auditive » mise en lumière par l'étude. Au téléphone, à la radio, durant les réunions en ligne, les sons compressés sont partout. **Et avec eux, l'absence de niveaux faibles nécessaires au repos de l'oreille**. « Le son est tout le temps fort », résume le professeur Avan.

Plusieurs situations du quotidien permettent d'illustrer la fatigue auditive liée aux sons compressés. En premier lieu, les longues réunions en visioconférence qui se sont multipliées avec le développement du télétravail. « Après une utilisation prolongée, les gens se plaignent d'un son « très fatigant », donnant mal à la tête », rapporte Paul Avan.

Autre exemple, les publicités à la radio ou à la télévision sont ressenties comme « plus fortes » par les auditeurs. « En fait, c'est faux, remarque le scientifique. Les publicités sont au même niveau que le reste, mais l'absence de sons faibles donne ce sentiment d'être « bombardé » par une pub ».

De plus, on suspecte que ces expositions prolongées aux sons compressés rendent **plus sensibles aux sons forts.** Ils pourraient aussi provoquer **des acouphènes**.

COMMENT LIMITER LES EFFETS DES SONS COMPRESSES ?

La première recommandation du professeur Avan est de **se préserver d'une exposition à des musiques ou des sons compressés**. Et si l'on ne peut les éviter ? « S'éloigner des écrans, diminuer le volume de son casque audio ou de son enceinte est plus sage », recommande le chercheur. D'autant plus que le but initial de la compression est de rendre les sons plus audibles.

En revanche, ménager des pauses régulières n'aurait pas permis de limiter la fatigue auditive. « On imagine que cette fatigue vient au bout de quelques minutes. **Il faut savoir que le circuit auditif travaille bien plus que tout autre circuit neuronal** », explique Paul Avan. « Le problème, c'est que le grand public ne peut pas identifier à l'oreille une musique compressée, reprend Véronique Bazillaud. Il n'a pas non plus la possibilité de choisir d'écouter un son naturel ou modifié ».

Afin de limiter l’exposition à ces sons compressés, **la création d’un label pour les musiques non (ou moins) compressées est envisagée**. « L’idée serait de garantir qu’une musique a été produite sans compression et qu’elle est donc « sûre » », imagine la Déléguée Générale de la Fondation Ecouter Voir. Un groupe de travail sous égide de l’association de la Semaine du Son doit voir le jour prochainement, auquel la fondation Ecouter Voir participera pour en étudier la faisabilité. Des contacts directs avec les producteurs de musique sont d’ores et déjà établis.

**Il s’agit du niveau sonore maximum autorisé par la loi dans les lieux publics en France.*

3.Audiologie Demain, article publié le 03/03/3021

Lien:https://audiologie-demain.com/100-sante-musique-laudition-au-coeur-de-la-semaine-du-son/les-animaux-exposes-a-la-musique-compressee-sont-vulnerables

« LES ANIMAUX EXPOSES **A LA MUSIQUE COMPRESSEE SONT VULNERABLES** »

La musique compressée est omniprésente. De l'avis de Christian Hugonnet, président de La Semaine du son, et de nombreux mélomanes, elle affecte le système auditif. Le Pr Paul Avan et Tamara Dos Santos ont entamé des travaux pour vérifier cette assertion. Premiers éléments de réponse, présentés lors de La Semaine du son.

Propos recueillis par Bruno Scala

Audiologie Demain : Quel est l'effet de la compression sur la source sonore originale ?

Paul Avan : Il y a deux sortes de compression : l'une qui concerne la taille des fichiers et l'autre la dynamique. C'est cette dernière qui nous intéresse et que nous testons. Schématiquement, cela consiste à diminuer les contrastes entre les sons faibles et forts. En pratique, tous les sons pertinents vont ainsi dépasser le bruit de fond. C'est un type de compression qui est utilisé dans les environnements bruyants, notamment à la radio[1] ou encore par les logiciels de visioconférence. Le résultat, c'est que tous les micro-silences que l'on trouve dans la musique sont supprimés.

AD : Qu'est-ce qui a motivé ces travaux ?

PA : Tout vient d'une intuition de Christian Hugonnet (fondateur de La Semaine du son). En tant qu'acousticien, il juge qu'écouter pendant longtemps de la musique ayant subi une forte compression de la dynamique est délétère. Selon lui, le système auditif est sollicité en permanence. Mais on ne sait pas s'il a besoin de se reposer. Pour l'instant, le dogme est le suivant : il ne faut pas dépasser un certain niveau sonore pendant une durée donnée. Mais peu importe le déroulement temporel.

AD : C'est ce dogme que vous remettez en question ?

PA : Notre hypothèse est en effet celle de Christian Hugonnet : la musique compressée affecte le système auditif. Nous avons commencé une étude de faisabilité[2] sur des cobayes (ou cochons d'Inde), qui ont une audition assez proche de la nôtre. Nous leur avons fabriqué une véritable discothèque – ce qui nous a valu des problèmes de voisinage... Et nous les avons exposés à de la musique compressée pendant 4 heures, à 102 dB. Si la musique compressée est fatigante pour le système auditif, alors les réflexes de protection – les muscles de l’oreille moyenne et le système efférent médian – qui sont sollicités en permanence, vont s’épuiser. Nous avons donc testé les réflexes des cobayes juste après exposition, à 24 h, 48 h et 7 jours.

AD : Quels sont les premiers résultats ?

PA : Concernant la sensibilité auditive, nous n’avons pas vu de changement significatif. C’est plutôt une bonne nouvelle dans la mesure où nous avons respecté la réglementation en vigueur. En revanche, chez les animaux exposés à la musique, la sensibilité des réflexes diminuent : ceuxci ne répondent plus pour des sons modérés à intenses, alors qu’ils le devraient. C’est donc symptomatique d’une fatigue. Tandis que les cobayes exposés à la musique normale récupèrent au bout de 48 h, ceux exposés à la musique compressé ont besoin de 7 jours. Ce qui veut dire que si on réexposait ces animaux dans l’intervalle, il y aurait des problèmes. D’un point de vue clinique, tout est normal, mais en réalité, ils sont vulnérables. Reste à identifier le mécanisme.

AD : Comment allez-vous procéder ?

PA : Nous allons étudier l’histologie et tout ce qui est en lien avec les surdités cachées. Car c’est cela dont il s’agit. Il faut mettre le doigt sur l'élément qui ne fait pas son travail : les mitochondries, les peroxysomes, les pompes à glutamate… Il y a finalement assez peu de candidats. Trouver le mécanisme est essentiel, car on découvrira certainement que des personnes sont plus vulnérables que d’autres. Cela nous mettra sur la voie d’un traitement.

En attendant, on peut alerter la communauté du son afin d’éviter les compressions féroces.

4. Handicap.fr : Semaine du Son (UNESCO) : dégâts des sons compressés sur l'audition

Par Cassandre Rogeret, 16 janvier 2022

https://informations.handicap.fr/a-semaine-du-son-degats-sons-compresses-sur-audition-32175.php

RDV du 16 janvier au 1er février 2022 pour la Semaine du son de l'Unesco. 65 évènements pour sensibiliser à l'importance de la qualité de l'environnement sonore et prévenir le déclin auditif. Focus sur la musique surcompressée : attention danger !

THEMES :

Les limites de la perception auditive ne sont pas extensibles compte tenu des capacités humaines d'écoute. L'oreille ne disposant pas de « ***paupière*** », l'être humain est sans cesse exposé à des niveaux sonores de plus en plus élevés, souvent de manière continue. Excessive, cette exposition peut avoir un impact sur le bien-être, la santé, la qualité et même l'espérance de vie et provoquer troubles du sommeil, des apprentissages ou encore anxiodépressifs (article en lien ci-dessous). Du 16 janvier au 1er février 2022, l'Unesco organise la Semaine du son afin de sensibiliser le plus grand nombre à l'importance de la qualité de l'environnement sonore et les éventuels facteurs de perte auditive. Au total, 65 évènements seront organisés dans plus de 30 villes à travers la France.

MUSIQUE SURCOMPRESSEE : DANGER !

En amont, son fondateur, Christian Hugonnet, ingénieur en acoustique, a invité le professeur Paul Avan, directeur du Centre de recherche et d'innovation en audiologie humaine à l'Institut de l'audition, à exposer les résultats de son étude sur les sons compressés, un son numérique qui a été « ***tassé*** » électroniquement pour faire remonter les niveaux sonores les plus faibles et être plus audible.. Durant quatre heures, des cochons d'Inde ont été exposés à de la musique très compressée ou non, au niveau maximum légal de 102 dBA. Leur audition a été évaluée juste avant, juste après, puis 24h, 48h et jusqu'à une semaine plus tard. Conclusion : même si aucun n'a perdu l'audition, « ***les animaux exposés à la musique surcompressée présentaient une fatigue plus importante des voies réflexes protectrices de l'oreille. De plus, le temps de récupération du réflexe était de plus de 48 heures*** », expliquent les scientifiques. « ***Ainsi, l'exposition répétée à la musique surcompressée est potentiellement dangereuse pour la sensibilité auditive car elle rend l'oreille plus vulnérable, même lorsqu'elle ne la menace pas immédiatement*** », poursuivent-ils. Grâce à la mobilisation de chaînes de radio et de musiciens, la mise en place d'un label « ***non compressé*** » pourrait voir le jour. Le 1er février à 20h, un concert démonstration sera organisé pour « ***rendre tangible*** » ce qu'est une musique compressée, au théâtre du Châtelet, avec Thomas Dutronc. Tendez l'oreille... ou pas!

DES CONFERENCES ACCESSIBLES EN LIGNE

Le 17 janvier 2021, la voix de Roberto Alagna, parrain de cette 19e édition, résonnera au sein du siège de l'Unesco (interview en vidéo ci-contre). Pour ce concert d'ouverture, le ténor franco-sicilien a choisi d'interpréter des compositions de Puccini, Verdi, Tchaïkovski... Il sera ensuite suivi de forums, conférences, tables-rondes, ateliers et concerts, accessibles en présentiel mais aussi en distanciel, en français et en anglais sur la web TV de lasemaineduson.org (programme complet en lien ci-dessous). L'occasion d'aborder différentes thématiques liées au sonore dans les domaines de l'environnement, la santé, la société, l'économie, l'industrie et la culture. Tout au long de l'année 2022, vingt pays organiseront à leur tour une semaine du son afin de sensibiliser sur cet enjeu de santé publique mondial.

5. Audiologie Demain (31/01/2022)

https://audiologie-demain.com/la-notion-de-micro-silence-est-fondamentale-pour-la-sante-auditive

« **La notion de micro-silence** est fondamentale pour la santé auditive »

À l'heure où nous mettons sous presse, la 19e Semaine du son de l'Unesco est sur le point de s'achever. La soirée du 19 janvier, qui s'est déroulée au sein de l'Unesco, était consacrée à la santé auditive. Plusieurs thématiques ont été abordées par les orateurs, dont la mise en place d'un partenariat avec l'Institut de l'audition, les résultats des travaux de l'équipe de Paul Avan sur les sons compressés et le projet de faire des audioprothésistes des référents du son. Explications avec Christian Hugonnet, fondateur et président de La Semaine du son.

Propos recueillis par Bruno Scala

Audiologie Demain : En quoi consiste le partenariat que La Semaine du son a noué avec l'Institut de l'audition ?

Christian Hugonnet : La Semaine du son et l'Institut de l'audition ont un intérêt à collaborer ; nous nous apportons beaucoup mutuellement. Notre transversalité, à travers notre relation avec le monde artistique, avec le monde industriel (acoustique, etc.), les intéresse beaucoup car ils sont dans la recherche et le sanitaire. L'étude sur la compression du son (voir plus bas) émane de cette complémentarité : nous leur apportons des connaissances sur la physique du son, qu'ils n'ont pas forcément. Et inversement, cela nous confère une assise et une caution importantes pour tout ce qui touche à la santé auditive. Il y a une dimension essentielle également : nous portons une vision positive de l'audition. Nous l'avons toujours envisagée sous l'angle du cadre de vie, et non de la pénibilité ou de la maladie. Concrètement, ce partenariat se traduira par des rencontres au sein de groupes de travail auxquels nous serons associés, ainsi qu'un échange de logos.

AD : Les travaux sur la compression du son que vous évoquiez ont été présentés par Paul Avan. Quels sont les résultats ?

CH : Tout est parti de cette intuition, qui était la mienne mais aussi celle de nombreux musiciens ou ingénieurs, que la musique surcompressée était nocive. Encore fallait-il le démontrer. Grâce au soutien financier de la Fondation Écouter Voir, nous avons pu monter une étude sur les cochons d'Inde, que nous avons confiée à Paul Avan (Ceriah/Institut de l'audition) et Tamara Dos Santos, postdoctorante dans son équipe. Ces travaux ont montré que la notion de micro-silence est fondamentale pour la santé auditive. La musique surcompressée est tassée sur une plage dynamique de 3 ou 4 dB, et plus rien ne respire ! On remplit l'espace sonore et on supprime ces micro-silences qui sont indispensables au fonctionnement de l'oreille interne, et probablement du cerveau – car, comme l'a expliqué la Pr Christine Petit, il faut arrêter de dissocier l'audition du cerveau. Ainsi, les cochons d'Inde exposés à la musique surcompressée mettaient plusieurs semaines à récupérer (leurs réflexes protecteurs, NDLR – lire l'article que nous avions écrit l'an dernier, lors de la présentation des résultats préliminaires : *Les animaux exposés à la musique compressée sont vulnérables*), là où les animaux exposés à la même musique non compressée, récupèrent presque instantanément. Ce qui veut dire que le dogme consistant à préconiser de « faire des pauses sonores toutes les quatre heures » est caduc.

AD : Quelle suite allez-vous donner à ces travaux ?

CH : Nous allons poursuivre l'étude pour affiner nos résultats. Mais, forts de ces premiers enseignements, nous allons aussi, en partenariat avec Universal et l'Ircam, développer un label, qui sera déposé sur les albums, et qui garantira un faible niveau de compression. C'est un pas en avant très important !

AD : Lors de cette soirée de la santé auditive, Arnaud Coez a présenté un projet, en partenariat avec l'Association des maires de France, consistant à faire des audioprothésistes des référents du sonore. Pouvez-vous nous expliquer ce projet ?

CH : Arnaud Coez a fait une présentation interactive, qui permet de découvrir les différentes notions du son, le décibel, etc. Ce document, avec un micro et un haut-parleur, constitue un kit de sensibilisation que nous mettons à disposition des audioprothésistes. Pour déployer ce projet, nous avions besoin de la collaboration de l'Association des maires de France. Les édiles accueilleront ainsi les audioprothésistes, qui sensibiliseront les populations. Tous les professionnels volontaires peuvent participer à ce projet et devenir les référents son au sein de leurs régions.

6. AUDITION-INFOS : UNE PREMIERE ETUDE SUR LES DEGATS DES SONS COMPRESSES SUR L'AUDITION, 13/01/2022

http://www.audition-infos.org/actualites/1607-une-premiere-etude-sur-les-degats-des-sons-compresses-sur-l-audition.html

SELON UNE ETUDE A PARAITRE, MENEE SUR DES COCHONS D'INDE, LE TRAITEMENT INFORMATIQUE ACTUEL DE LA MUSIQUE NUIRAIT A NOS OREILLES.

Pour des raisons techniques, que ce soit le son diffusé en radio ou la musique enregistrée sur des support numérique, le signal est actuellement compressé informatiquement. Comparée au son naturel, cela revient à aplatir le son vers le bas, puis à remonter le niveau sonore de l'ensemble, pour le rendre plus compréhensible, « Ce qui fait que le son sature. Il n'existe plus de silence, comme le bruit permanent d'une ville » explique Christian Hugonnet, président fondateur de La Semaine du Son et ingénieur acousticien.

Paul Avan, directeur du centre de recherche de l'audition à l'Institut de l'Audition a mené pendant deux ans une étude sur l'effet de ces sons compressés sur des cochons d'Inde, qui ont une ouïe assez proche de celle des humains. « Nous les avons exposés à de la musique pendant quatre heures, ce qui correspond à une soirée en discothèque » explique le chercheur. Plus précisément, les cobayes ont subi une chanson d'Adèle à 102 dB, à une semaine d'intervalle. L'effet sur leur audition et la récupération ont ensuite été évalués.

« L'oreille interne a bien supporté le choc, en revanche, le système cérébral semblait fatigué. Des petits muscles derrière le tympan ne récupéraient pas » poursuit le chercheur. « Ces petits muscles sont comme des paupières sonores. Nous pensons qu'ils dosent la quantité de son qui doit entrer dans l'oreille droite ou gauche, en fonction de ce que l'on souhaite entendre ».

En revanche, l'effet sur nos oreilles est supposé, car il n'existe pas encore d'étude comparable sur l'humain, ni de statistiques établissant le différence entre une exposition au son compressé et au son naturel. Alors tous appellent de leur voeux la poursuite de ces recherches.
"Le vrai risque serait que des jeunes qui ont 18 ans aujourd'hui deviennent des personnes de 80 ans avec des problèmes congnitifs" explique Alain Londero, médecin ORL.
La solution ? "Ne pas tout applatir de manière sauvage" estime Paul Avan. Selon les auteurs de l'étude, les ingénieurs du son maîtrisent parfaitement bien cette compression et ils sont eux-mêmes demandeurs de la diminuer. France Culture serait par exemple peu compressé. Les musiciens seraient également dans la même démarche.
"Thomas Dutronc ou Universal Music nous ont rejoint, car ils pensent également que cette musique devient indigeste" conclut Christiant Hugonnet.
Cette étude, financée par Visaudio (maison mère d'Ecouter Voir), devrait prochainement se prolonger par une recherche sur l'humain, et la mise en place d'un label "non compressé" pourrait voir le jour.

7. ACOUSTIC SHOCKS ARE A RED HERRING. A DIFFERENT, NOT-SO-SILENT THREAT IS SLOWLY POISONING THE INTERPRETER'S EARS.

Caniato, A. 13/05/2022

Andrea Caniato is a voice researcher/consultant and certified voice trainer (Applied Physiology of the Voice) with a background in psychoacoustics and music and is an EU-accredited interpreter.

Abstract

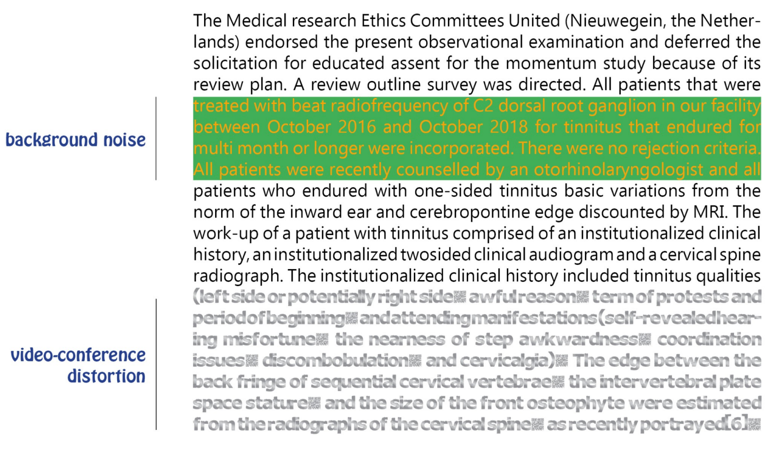
The idea that interpreters might be exposed to acoustic threats causing highly undesirable symptoms like tinnitus, hyperacusis, hearing loss and the like is rapidly gaining ground in the conference interpreting industry. A study is being conducted by AIIC on “acoustic shocks” in the booth and recent, worrying news about the strikingly high incidence of said symptoms with interpreters working for the Canadian Parliament in remote mode indicates that the RSI setting might be particularly conducive to the onset of hearing problems. Similar syndromes have been observed in call centre worker populations. Sudden spikes in sound pressure levels (SPL) unleashing protective reflexes in the middle ear (acoustic reflex) have been provided as the reason for the onset of symptoms, and the syndrome has been named “Acoustic Shock Disorder” (ASD), the idea being that the sudden arrival of a high “quantity” of sound pressure “shocks” an otherwise healthy ear. SPL output limiters are therefore viewed as the solution: by limiting the amount of decibels a headset can produce, no sudden noise will reach an sound pressure levels (SPL) high enough to cause an “acoustic shock”.

This purely “quantitative” approach does unfortunately not consider the quality and composition of the sound feeds both call centre workers and interpreters are typically exposed to, nor the conditions they actually work in. In addition, scientific studies conducted in call centres recognize that “acoustic shock” symptoms can occur in workers who have not experienced any sudden increase of SPL and that the use of output limiters (limiting the amount of decibels a headset can produce) does not prevent the onset of symptoms, which has led some scholars to believe that “acoustic shocks” are psychogenic. A different, qualitative approach focusing on in-depth understanding of middle ear physiology and the analysis of sound feed quality can explain the symptoms and indicate a more effective way of protecting the interpreters’ ears, voices and well being. Having to extract meaning from poor sound with low signal-to-noise ratios (excessive amounts of microphone gain and/or background noise), or distortion, artefacts etc is unfortunately not the exception in the booth and is very often the norm in RSI, no matter what headset is used. This overworks the human ear and forces interpreters to turn up their volume knob to maximise the amount of “useful” signal. The cocktail effect of louder than needed volume (but still well below theoretical “acoustic shock” thresholds) and middle ear muscle fatigue owing to bad acoustics over long periods of time suffices to explain why symptoms can occur without

being caused by sudden SPL peaks (over the 93 dB “dangerous” threshold), or following SPL peaks below this dangerous threshold, which no sensible output limit would ever prevent and which a healthy ear would be perfectly able to withstand. Interpreters therefore need to learn how to tell poor sound feeds from good quality sound feeds and avoid the former. Sound clips exemplifying quality differences are provided in the article.

Key take-home messages:

- The way they are currently defined, acoustic shocks are extremely rare if not impossible occurrences in the booth. They are probably not the real reason behind the high incidence of hearing problems in the conference interpreting community.
- Output limiting headsets have been shown not to solve the problem. Insisting with the current approach might even prove detrimental as it diverts the focus from the real root of the problem.
- Interpreters are frequently exposed to bad quality sound feeds. More so in the RSI setting. Frequent exposition to bad quality sound overworks your ear and in time can generate the symptoms described as “acoustic shock”.
- Interpreters are not trained to recognize poor sound and are seldom bothered by it. As long as they can somehow make out the words, they will be prepared to work and will not complain. If they struggle because of poor sound, they will typically blame it on the speaker or on themselves. This contributes to excessive exposition to ear-poisoning sound quality.
- The ability to tell good sound from bad sound can be developed, and can promote the improvement of acoustic working conditions in the booth. This can prevent the onset of “acoustic shock” symptoms more effectively than any output limiter.



(source: Pulsed radiofrequency of C2 dorsal root ganglion in patients with tinnitus. Authors: Henk M Koning, Bas C Ter Meulen)

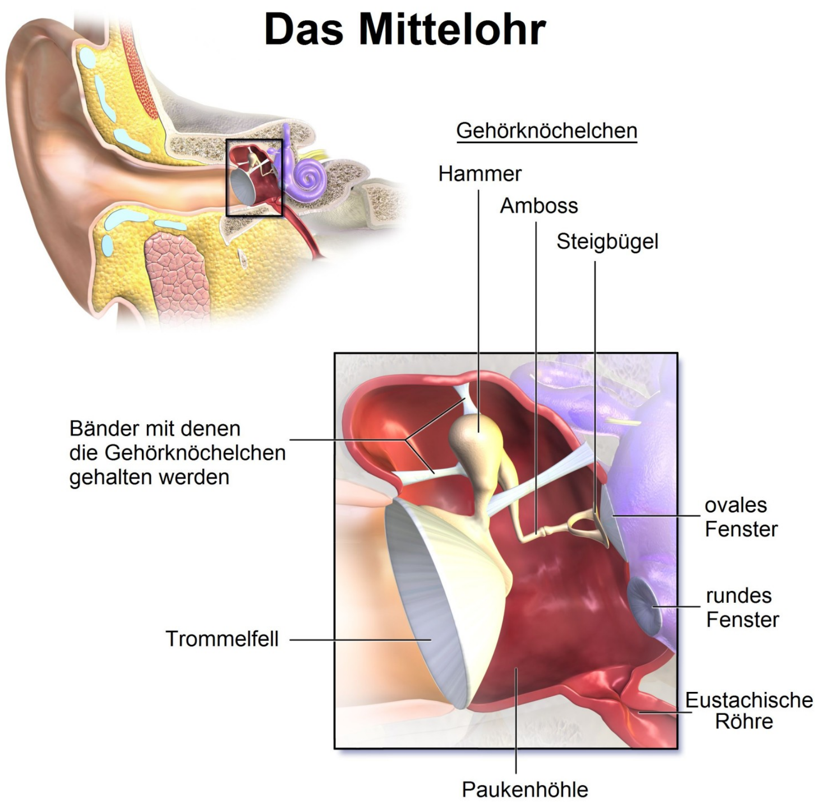
What poor sound would "look like" if your ear could see it

Main article

The Covid19 crisis and lockdown have sparked a lively debate about Remote Simultaneous Interpreting as a valid alternative to simultaneous interpreting from the booth. Dozens of platform reviews have been published and interpreter organisations have issued guidelines and recommendations. Little or no attention is paid here to the quality of sound feeds, which is more or less being taken for granted. Whenever sound is discussed, the main concern seems to be "avoiding acoustic shocks" by using sophisticated headsets or even special devices (limiters) to be attached to the interpreter's computer to cut any loud noise suddenly appearing in the interpreter's ears. In the following, I will argue that the worryingly high incidence of nasty hearing problems is very probably not due to "shocks" but rather to frequent if not constant exposition to poor quality sound, both in the booth and, especially, in the RSI setting, although "acoustic shocks" in the booth predate the age of RSI.

How are "acoustic shocks" defined?

Most definitions of "acoustic shock" refer to an unexpected, loud sound causing a wide range of symptoms mainly including pain, tinnitus, vestibular disturbance and hyperacusis. The term "acoustic shock" does therefore not necessarily refer to a situation where a very loud noise like an aeroplane engine or the like can instantly shatter your eardrums and leave you deaf because of the very high levels of sound pressure (SPL, expressed in decibel, dB) generated. That type of damage can only occur at SPLs equal to or exceeding 120 dB, and given that the SPL output of virtually all of of today's digital equipment is limited by default, (Hooper 2014) the odds that a commercial headset, interpreting console or home PC (in the RSI setting) will ever get even close delivering a similar amount of sound pressure to the interpreter's eardrums are extremely low or equal to zero, unless of course the equipment is being intentionally misused, attached to an external amplifier or has been tampered with.



And indeed the ear canal and tympanic membrane (the eardrums) of subjects reporting "acoustic shocks" appear healthy and normal (Westcott 2008). Moreover, the Acoustic Shock Disorder symptoms are known to occur in people who have never experienced any sudden, painful noises or the like but have long, cumulative exposure to headset use (Westcott 2006). Studies conducted on call centre workers have shown that operators continue to develop ASD symptoms even if noise-limiting headsets are being used to prevent "acoustic shocks" (Westcott 2008). Canadian Parliament interpreters have started reporting "acoustic shocks" in droves just a few weeks after transitioning to online meetings (The Hill Times, May the 6th, 2020). Has this transition from the booth to RSI suddenly increased the amount of dB these interpreters are exposed to? The very existence of such thing as an "acoustic shock" has even been questioned by some scholars in prestigious scientific journals and some researchers believe the problem is psychogenic (Hooper 2014).

Do you really need a sudden increase in sound pressure levels to develop "acoustic shock" symptoms?

I believe the problem exists, and that it is neither psychogenic nor dependent on sudden increases in SPLs. According to the traditional "protection theory", sudden loud noises can indeed trigger the established protective function of the tensor tympani muscle and of the stapedius muscle (the two muscles governing the function of the middle ear). The "protection theory" has recently been the object of scholarly criticism (Huttenbrink, Buytaert et al, Gelfand, Cheng and Gan, Kirikae), but even assuming this theory is correct, the protection function is known to be triggered even in absence of sudden “noise events”. And indeed if sudden “noise events” were the real problem, output limiters would solve it, but unfortunately they do not.

So where do these nasty symptoms come from?

The tensor tympani muscle is in fact activated in response to different stimuli including loud noise, vocalization, articulation (soft palate and jaw movements) and appears to be involved in the discrimination of lower frequencies (Westcott 2006, Rock, 1995). Indeed, by adjusting the tension and therefore the impedance of the tympanic membrane (Westcott 2006, Rock, 1995) and oval window respectively, the tensor tympani and stapedius muscle play a key role in determining both the quantity (amplitude or “how loud”, in dB) and the quality, i.e. the type of frequencies that will filter through to the cochlea (Tomatis 1987). This means they play a crucial role in the discrimination of different sounds, by feeding the cochlea with the right amount of the right type of frequencies, an essential function when it comes to the ability of understanding language.

Back to the protection function, in people with normal hearing the activation of the acoustic reflex begins at around 70 dB. (Campbell 2018). Interestingly, the reflex is also activated when people start vocalizing. Activation begins 10-20 dB below the discomfort threshold, i.e. long before the subject can realize that any given sound is too loud. And even more interestingly, the activation threshold decreases if one or more additional sounds are presented to the ear, especially if their frequency is lower than the sound causing the acoustic reflex (Kawase et al, 1997). This means speaking and listening at the same time activates the acoustic reflex (causing the tensor tympani and stapedius muscles to contract). Listening to a noise-riddled sound feed basically means exposing your ear to additional layers of sounds that will lower your acoustic reflex threshold even further. If sound is compressed or sampled by means of inefficient algorithms, it will present to our ears with missing or misplaced frequencies, weird equalisation and missing or distorted bits that will throw the impedance adjustment mechanisms of the middle ear off their natural balance. Exposition to this type of sound 7 hours a day for many days a month will keep your middle ear muscles constantly contracted and working inefficiently for very long stints, (on the road to chronic hypercontraction) and will increase your sensitivity to events that would normally not pose any threat to your ears. Stress and anxiety, that are no strangers to the interpreting booth, are also believed to lower the acoustic reflex threshold (Patuzzi, Milhinch and Doyle, 2000).

When sound is muffled, distorted and/or full of compression artefacts, words appear to our ears the same way text with poor contrast appears to our eyes. Bad contrast is the visual equivalent of background noise or compression artefacts (RSI). What would happen to our eyes if we had to read and understand 80 pages of text written in a bad contrast day in, day out? Having to extract words from sound with low signal-to-noise ratio or low fidelity owing to processing by compression algorithms (the acoustic equivalent of difficult-to-read fonts) overworks middle ear muscles, as they are constantly striving to adjust the impedance of elastic membranes (tympanic membrane and oval window) to chase crucial, "load bearing" frequencies that are either being masked or disrupted by interfering noise or have been removed, distorted and/or replaced with surrogates during compression (which typically happens in RSI, but not only), while at the same time stiffening the very same membranes (i.e. reducing their elastic potential) to protect the cochlea as the acoustic reflex is being activated by the combined, cocktail effect of our vocalization, stress, fast-paced articulation and the totally artificial frequency structure of muffled audio. It is a bit like wanting to eat and swallow your food while singing opera.

What does all of this means for interpreters?

The only thing you can do when the quality of your sound feed is low and when your own voice is interfering with the listening process, is raise your headset volume to maximize the "useful portion" of your signal. If you need to understand language, there is no other way of compensating for poor sound than turning up the volume, (and raising your voice as a consequence: Yu-Hsiang Wu 2018). This is a vicious circle that will expose your ears to many more decibels than you would actually need, over long stints, on an almost daily basis. And constant poisoning does not need high dosage to cause huge damage over time.

The good news is that you can learn to distinguish good sound from bad sound. On the whole, interpreters do not appear to be aware of sound quality, and typically assume the sound engineers in charge will deliver the best possible output. Unfortunately, this is much too frequently not the case for multiple reasons. But like with good food and good wine, the ability to tell decent from poor is acquired by comparing and contrasting different "products". Below are a number of videoclips that will help you train your ear and understand what type of sound quality you should expect and demand from your clients to make sure you are working in "sound" acoustic conditions. In 2020, decent sound quality is perfectly possible even without expensive equipment. What it requires is proper and accurate management of the chain of transmission and an understanding of what improves and what hampers language intelligibility on part of sound engineers and conference technicians. The sound quality of "good" examples below in the booth is perfectly possible, and it happens, although it is not necessarily the rule. The current structure of the RSI market makes it virtually impossible to ensure consistently decent quality when using videoconferencing platforms.

In listening to the following examples, readers are invited to ask themselves questions like "Which of these sound feeds would be easier to understand and interpret?"; "Would I want to be interpreting this type of sound feed, regardless of the content, 7 hours a day?"; or "Which of the 2 types of feed would I like to be interpreting?". And last but not least, "How are my ears reacting to the different qualities"?

Exemple 1: lecture, 2 clips

VIDEO LINK (1"22): Lecture on alternative financial instruments – good quality

https://www.youtube.com/watch?v=WNIIyoNfN5w

VIDEO LINK(1"22): Lecture on alternative financial instruments – bad quality (conference setting)

https://www.youtube.com/watch?v=YEJy9k96f-4&t=3s

Exemple 2: BBC radio (1 clip)

VIDEO LINK: BBC4 "In our time" – alternating original radio quality and "videoconferencing" (RSI) quality every 30-40 seconds

https://www.youtube.com/watch?v=P9wHWEdq7Mk

Exemple 3: Norway Prime Minister (2 clips)

VIDEO LINK: Norwegian prime minister Erna Solberg good quality feed, conference setting https://www.youtube.com/watch?v=qiB0iS-ssH0 VIDEO LINK: Norwegian prime minister Erna Solberg poor quality feed, RSI setting https://www.youtube.com/watch?v=3M0jHI1YtOw

Example 4: Irish Prime Minister (2 clips)

VIDEO LINK: Irish prime minister Leo Varadkar conference setting good sound https://www.youtube.com/watch?v=E1IJ0T402Wc VIDEO LINK: Irish prime minister Leo Varadhkar RSI videoconferencing setting, poor sound https://www.youtube.com/watch?v=3fSoNBeddGs&feature=emb_title

Example 5: French President, 2 clips

VIDEO LINK: French President Macron – Good quality sound feed broadcasted from the Elisée Palace https://www.youtube.com/watch?v=5ZuCesdY6XE VIDEO LINK: French President Macron – Poor quality sound feed (compression, bad "contrast", muffled sound) broadcasted from the Elisée Palace https://www.youtube.com/watch?v=l55Zy4fUZfs

Example 6: WHO general director, 3 clips

VIDEO LINK/ WHO General Director Tedros Adhanom Ghebreyesus good quality feed from an office: https://www.youtube.com/watch?v=VlJP17SRsMs&feature=emb_title VIDEO LINK: WHO General Director Tedros Adhanom Ghebreyesus poor quality feed, RSI setting https://www.youtube.com/watch?v=dbmiNId4OKk

Finally, the way your ears work has an impact on the quality of and the amount of effort needed to produce your voice. The two systems are closely interconnected (Tomatis 1987, Landzettel, Rohmert 2015) and hearing problems may result in voice issues. The opposite is also true. After all, the sound of your voice is what your ear is constantly exposed to during your waking hours, 12 months a year. Refining your voice is a long journey that requires work and dedication, but it is a good way to present your ears with a “good quality feed”, thereby reducing the burden your ears are subject to. But this is another story.

The author: Andrea Caniato: AIIC Conference Interpreter with EU accreditation, voice researcher and voice coach. Main qualifications: Trainer Certificate in Applied Physiology of the Voice, BA in Contemporary Saxophone, MA in Conference Interpreting.

Acknowledgements: this article would have been almost impossible to write without Cristian Guiducci's invaluable sound-engineering expertise and remarkable ability to hunt through tons of scientific publications. Thanks, Cristian. Additional help has come from Ben Mulvihill. Thanks, Ben.

Literature

K. C M Campbell, Impedance Audiometry. MedScape. 2018-09-12

Kawase, Tetsuaki; Takasaka, Tomonori; Hidaka, Hiroshi (June 1997). "Frequency summation observed in the human acoustic reflex". Hearing Research. 108 (1–2): 37–45.

Erwin H. Rock, OBJECTIVE TINNITUS AND THE TENSOR TYMPANI MUSCLE International Tinnitus Journal Vol. I, No.1, 1995

A. Tomatis, Alfred, l’Oreil et la Voix, 1987

R. E. Hooper, the Acoustic Shock Controversy, The Journal of Laryngology & Otology, Volume 128, Issue S2 July 2014, pp. S2-S9

M. Landzettel, G. Rohmert, Lichtenberger Dokumentationen, Band I, 2015, pp. 57-66,

Westcott, M. Acoustic Shock Injury. Acta Otol Supplement, 556, 2006: 54-58

Milhinch J. One hundred cases in Australian Call centres. Risking Acoustic Shock. A seminar on acoustic trauma from headsets in call centres. Fremantle, Australia, 18-19 September 2001. Morwell: Ear Associates Pty Ltd, 2001.

Westcott, M. Tonic Tensor Tympani Syndrome – an explanation for everyday sounds causing pain in tinnitus and hyperacusis clients. Poster IXth International Tinnitus Seminar, Gothenburg June 2008

Milhinch J. Acoustic Shock Injury – A report on injury following acoustic incidents in call centres. Melbourne, 2001 (personal communication).

Patuzzi R. Acute Aural Trauma in Users of Telephone Headsets and Handsets. In : Ching T editor. Abstracts of XXVI International Congress of Audiology, Melbourne 17th-21st March 2002. Aust N Z J Audiol (spec ed) 2002;23:2:132.

Patuzzi R, Milhinch J, Doyle J. Acute Aural Trauma in Users of Telephone Headsets and Handsets. Neuro-Otological Society of Australia Annual Conference. Melbourne 2000

K. B. Huttenbrink. Middle ear mechanics and their interface with respect to implantable electronic otologic devices. Oto-laryngol Clin North Am., 34(2):315–335, 2001.

J. A. N. Buytaert, W. H. M. Salih, M. Dierick, P. Jacobs,and J. J. J. Dirckx. Realistic 3d computer model of the gerbil middle ear, featuring accurate morphology of bone and soft tissue structures.Jaro, 12(6):681–696, 2011.

T. Cheng and R. Z. Gan. Mechanical properties of stapedial tendon in human middle ear.J. Biomech. Eng., 129(6):913–918, 2007.

T. Cheng and R. Z. Gan. Experimental measurement and modeling analysis on mechanical properties of tensor tympani tendon.J. Biomech. Eng., 30:358–366, 2008.

T. Cheng and R. Z. Gan. Mechanical properties of anterior malleolar ligament from experimental measurement andmaterial modeling analysis. Biomechanics and Modeling in Mechanobiology, 7(5):387–394, 2008.

S. A. Gelfand Essentials of Audiology. Thieme, 2009.

I. Kirikae. The structure and function of middle ear. Tokio University Press, page 60, 1960

The Hill Times, May the 6th, 2020, front page and page 5

Yu-Hsiang Wu,1 Elizabeth Stangl,1 Octav Chipara, Syed Shabih Hasan, Anne Welhaven, and Jacob Oleson, Characteristics of Real-World Signal-to-noise Ratios and Speech Listening Situations of Older Adults with Mild-to-Moderate Hearing Loss, Ear Hear. 2018 Mar-Apr; 39(2): 293–304.

8. THE PROPOSED PATHODYNAMICS OF THE JUNK SOUND SYNDROME: WHY RSI SOUND IS BAD FOR THE INTERPRETER'S EARS.

Caniato, A. published July 2020

Thinkstock

"Do no harm; so that the amplified signal is not unintentionally or undesirably altered by the hearing aid."

(Agnew, 1998, underscoring added by the author for emphasis)

ABSTRACT

RSI (Remote Simultaneus Interpreting) platform sound is poor quality sound. Its frequency range is limited to no more than one third of the audible spectrum and noise suppression, feedback cancelling and other algorithms originally developed to "improve" the listening experience end up manipulating and distorting the remaining frequency band. The huge chunks of the audible spectrum that are absent cannot be restored by any machine or sound engineer, and sound which has been manipulated by algorithms along the transmission chain cannot possibly be improved downstream, even in "hybrid" or "hub" mode as only minor adjustments are possible that do not make RSI sound less harmful.

Contrary to the dangerously widespread mantra "The higher parts of the audible spectrum are not really necessary when it comes to understanding speech", the upper portions of the human vocal spectrum are there for a number of reasons that turn out to be particularly

useful when people are required to listen, perfectly understand, process and simultaneously interpret speech, i.e. when the listening effort is made in the presence of interference and two streams must be monitored at the same time.

Receiving an artificial, narrow-spectrum "RSI type" feed (be it at home or at a fully-equipped, ISO-compliant professional interpreting hub) forces interpreters to misuse and overwork their hearing in a number of ways both at a peripheral (ears) and at a central nervous system (brainstem) level. Experience acquired during the Covid19 lockdown has shown that extremely undesirable short/medium/and potentially long term symptoms can arise as a consequence of even not-so-prolonged exposure to "junk", platform sound in the interpreting booth. These symptoms include tinnitus, hyperacusis, earache fullness of the ear, nausea, headaches and even vestibular disorders, that are usually mistaken for the result of "sudden peaks of noise". The present article provides a number of evidence-based, biomechanical, neurological, non-cognitive reasons why RSI-induced exposure to junk sound in the booth suffices to explain those symptoms without having to assume that a sudden loud peak of noise has somehow reached the interpreter's ears undetected.

HIGHLIGHTS

- Contrary to the requirements of applicable ISO standards, the sound delivered by videoconferencing and Remote Simultaneous Interpreting (RSI) platforms cuts off high and very high frequency information. Frequency range hardly ever exceeds 8k Hz, and spectral energy is typically concentrated in the 0-4k Hz segment at best. Frequency distribution is also arbitrarily modified by algorithms, all of which creates a very artificial acoustic environment and makes speech recognition difficult.
- High and very high frequencies are crucial to understanding speech in complex acoustic environments like the interpreting booth. They are key to sound localization, which allows the human brain to separate and simultaneously process two or more acoustic signals. Inability to localize and separate sound overworks the simultaneous interpreter's ears and nervous system.
- Lack of high and very high frequencies (7k-20k Hz) can be regarded as a form of sensory deprivation. With exposure, this might lead to hyperacusis and to the permanent loss of hair cells (the sensory receptors of the inner ear) in the long term, and result in permanent loss of hearing. A reduced spectrum stimulates a significantly smaller portion of the basilar membrane (which hosts the hair cells) in the cochlea. This means that the same amount of sound pressure (dB SPL) you would normally need with a full-spectrum signal is now applied to smaller area: pushing the bottom of a drinking glass against the palm of your hand with the equivalent of a 1 kg pressure for 10 minutes will cause you no harm, but pushing a needle against the same palm with the same amount of pressure for the same amount of time will make the needle pierce through your flesh.

- High and very high frequencies cause the same reflex reactions caused by dopaminergic stimuli. Dopamine is produced in the same areas of the brain that process spectral and sound localization cues, and protects hair cells from damage. It also regulates mood and attention. Low dopamine levels cause pain, stiffness of the muscles and a number of other problems.
- The characteristics of "platform" sound overwork middle ear muscles, including because lower frequencies require an additional amplification effort. They can also cause the interpreter's sensory system to process conflicting localization and direction cues and generate a sense of vection, which is believed to be the cause of vestibular issues like nausea, vertigo and postural imbalance.
- Junk sound is the product of the interaction between different factors, most of which end up affecting the brainstem. None of them is harmful for the interpreter's health if considered alone and if exposure is limited. But unfortunately, they tend to appear in combination, and cocktails of these factors are becoming the staple diet of interpreters around the world.

MAIN ARTICLE

WHAT RSI SOUNDS LIKE

The last few months have seen the use of Remote Simultaneous Interpreting (RSI) platforms skyrocket both in the private and in the institutional interpreting market. Though marketed as "new technology", all videoconferencing and RSI platforms rely on 25-year old VoiP technology (like Skype, WhatsApp) to deliver sound and video to distant locations over the internet and allow interaction between multiple meeting participants connecting from their homes or offices, with or without language interpretation, using whatever equipment they have available.

In order to save on bandwidth costs, reduce the effort needed to manage echo/feedback loops and background noise, and tackle the complex problems generated by a sound chain they do not fully control, all videoconferencing and RSI platforms currently limit the frequency bandwidth they broadcast to a maximum of 7-8k Hz (figures X, Y, Z) and use algorithms to process and “standardize” as much as they can what is left of the source sound.

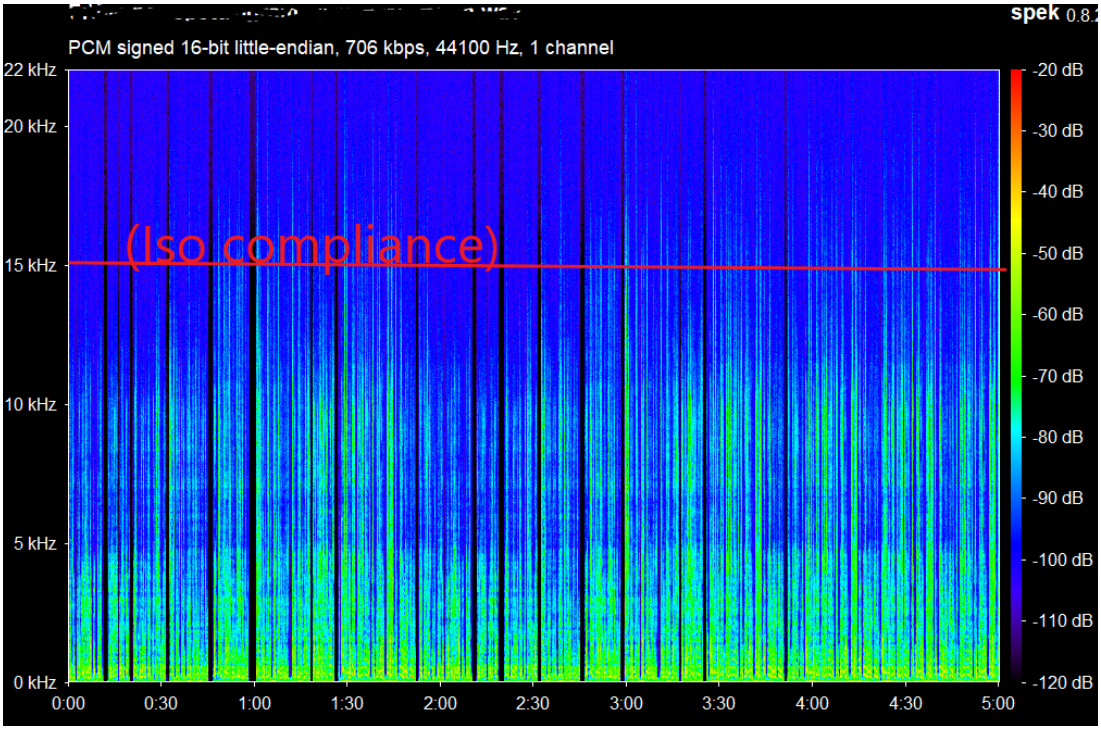


Figure 1 Spectrogram of a normal, adult female voice acquired using a RØDE NT1 microphone and a Steinberg UR22 interface. Normalized to -1 dB, 44.1 kHz, 16 bit. Areas in green and yellow show concentrations of spectral energy. Blue is not audible.

ISO standards applicable to the transmission of sound for the purposes of simultaneous interpreting require a minimum frequency range of 125-15k Hz (ISO 20109). The human ear can perceive sound in the 20-20k Hz range, and spectrograms of human speech clearly show activity up to 20k Hz (fig. 1, 2); most of its acoustic energy is found between 0 and 17-18k Hz and is organized in formants. Although it is true that speech remains intelligible even with a very reduced frequency range (telephony would be impossible without the human brain's ability to interpolate and virtually recreate the missing frequency information when listening to a limited bandwidth on the telephone), recent and not-so recent scientific studies (Moore et. a) have shown that the idea that all you need to hear to understand speech is limited to the 20-8k Hz, and that even 20-6k Hz "will do" ("the rest is for music", "the rest you cannot even hear", "the rest actually interferes with your ability to understand speech"), is fundamentally flawed, but extremely widespread both in the sound engineering and in the ENT doctor communities. Like all highly efficient systems, human speech (voice + hearing) is redundant by design, which means it produces and relies on much more information than the bare minimum in order to be able to function in all sorts of working environments and operating conditions.

Removing (and/or manipulating) critical chunks of information in the 20-8k Hz range will definitely hamper the intelligibility of speech, but in no way does that justify the idea that frequency information outside of that range is not useful or desirable to process and understand speech, especially when performing difficult auditory tasks that require high degrees of accuracy.

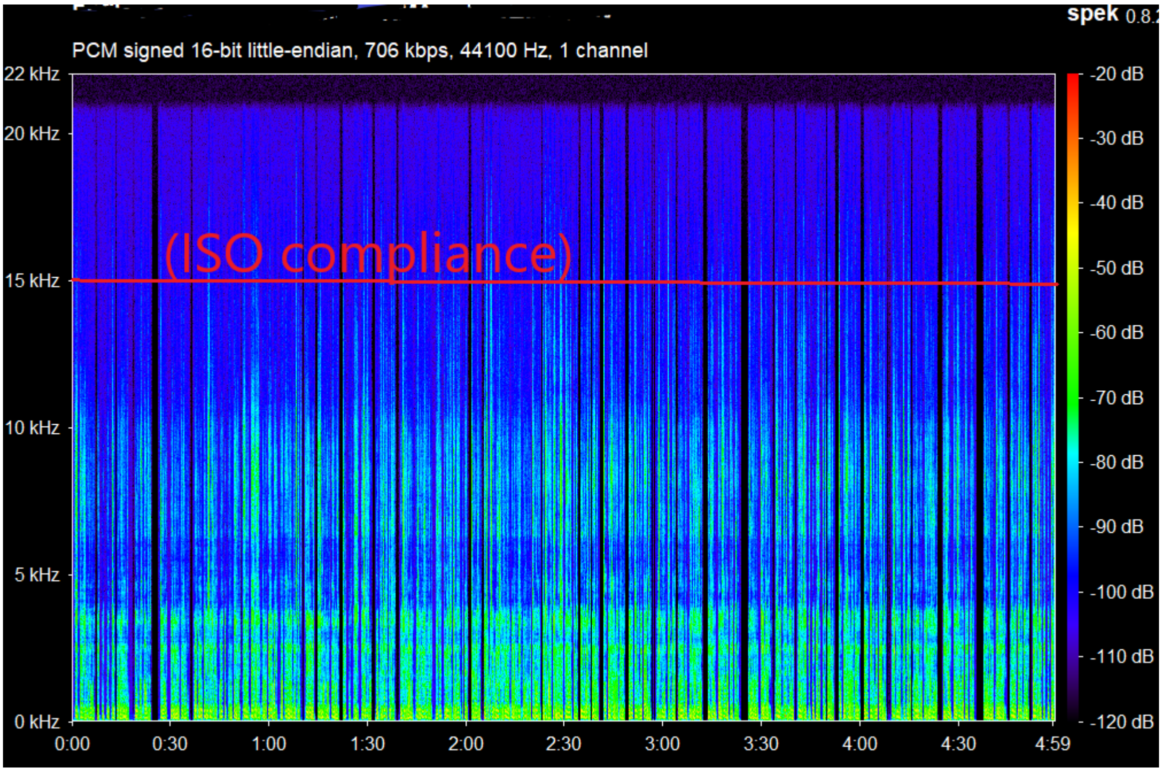


Figure 2 Spectrogram of a normal, adult male voice acquired using a RØDE NT1 microphone and a Steinberg UR22 interface. Normalized to -1 dB, 44.1 kHz, 16 bit. Areas in green and yellow show concentrations of spectral energy. Blue is not audible.

And yet when working with RSI feeds, both at home and at professional "RSI hubs" interpreters are expected to process speech with very limited frequency ranges (sometimes as low as 125-2-3k Hz, in any case hardly ever over 8k Hz), with overrepresented mid and low frequencies and huge spectral inconsistencies due to the activity of aggressive noise/feedback cancelling algorithms, that remove essential information along with noise and generate artefacts like spectral non-linearities, sound level fluctuations, unreal/inconsistent and misleading reverberation cues and harmonic and/or intermodulation distortion.

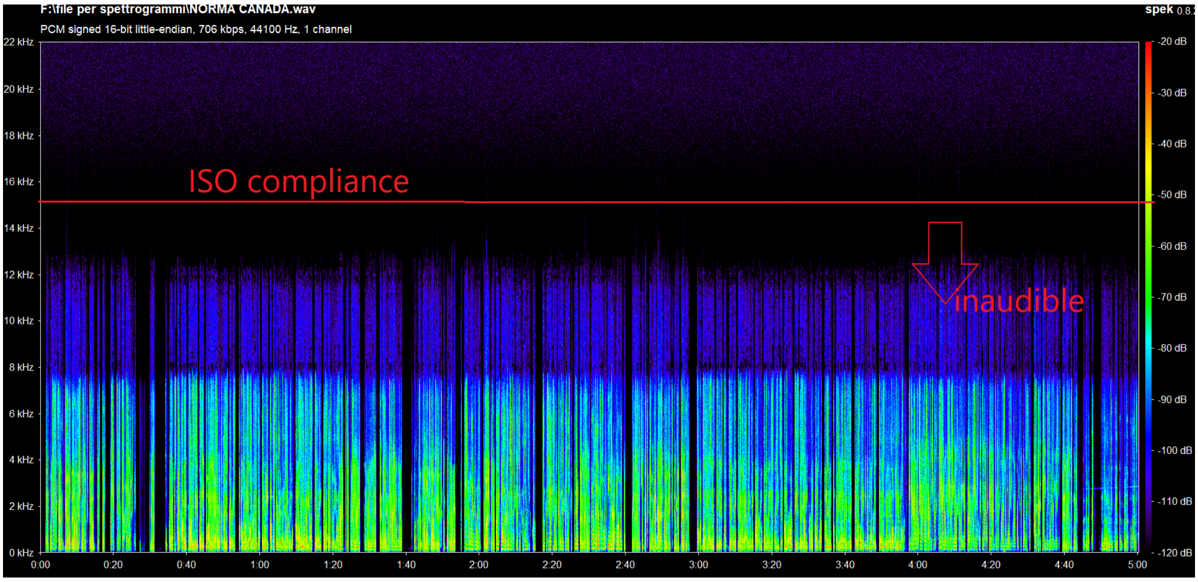


Figure 3 Spectrogram of male and female adult voices participating in a parliamentary debate via a very well known videoconferencing platform, normalized to -1 dB, 44.1 kHz, 16 bit. All speakers use an ISO-compliant headset. Areas in green and yellow show concentrations of spectral energy. Blue is not audible. Black is digital silence. The signal is limited to 8k Hz and spectral energy is distributed across the available range in an abnormally even way. The debate was interpreted in 2 languages.

The best-case scenario (fig 3) is a digitally limited bandwidth (125-8k Hz) with spectral energy either arbitrarily or evenly (and therefore unrealistically) spread across the range). The manipulation of critical frequency bands has been shown to have major negative impacts on speech intelligibility (Shannon et al.). Some platforms are even beginning to resort to digital tricks to to surrogate the missing frequencies in the upper portion of the spectrum. Needless to say, this might help them pass compliance tests in the laboratory setting, but what looks like "filler material" on a spectrogram does not replace the real missing frequencies. It does not solve the problem and it generates additional spectral non-linearities (Fig. 4).

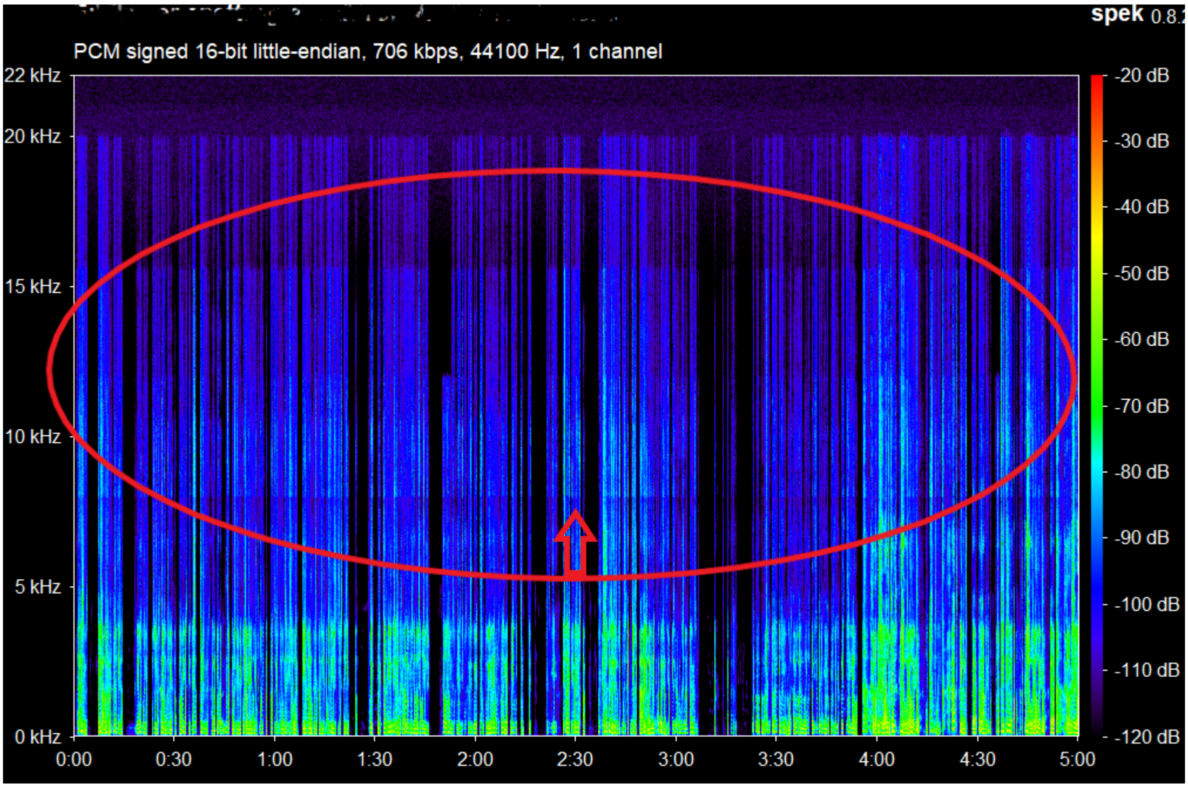


Figure 4: Spectrogram of the same male adult voice shown in fig. 2 acquired using a RØDE NT1 microphone and a Steinberg UR22 interface and broadcasted through a well known RSI platform. Normalized to -1 dB, 44.1 kHz, 16 bit. Areas in green and yellow show concentrations of spectral energy. Blue is not audible. This spectrogram shows a modular structure. The "module" between 8k and 16k Hz almost appears to have been copy-pasted from the "module" below. Non-linearities are evident when contrasted to fig. 2

A FAULTY APPROACH

When assessing sound quality and looking for the roots of hearing problems, one critical mistake is usually made: the real life situation is left completely out of the picture, and individual variables are assessed separately without considering their interaction in the interpreting booth. Thus, sound checks are often run by just listening to a feed (that is probably not even the feed that will have to be interpreted) and checking if it can be "heard". And even when speech intelligibility is assessed, the fact that an interpreter will actively be producing interference (listening and speaking at the same time, so the feed will have to extracted from "background noise"), is given no consideration whatsoever. And when the causes of hearing problems are looked for, even in the clinical setting a quantitative approach ("how loud was it?) is traditionally chosen instead of a qualitative approach ("What did it consist of? For how long? What did it interact with?"). There is no doubt that exposure (including prolonged exposure) to very loud sounds (firearms, explosions etc.) can cause hearing problems, but so can the huge sensory effort and the sensory conflicts caused by prolonged exposition to artificial, difficult to decipher, poor quality sound that needs to be separated from an additional, interfering source: the interpreter's voice (and perhaps from additional background noise if the feed is "dirty").

ON GOES THE MICROPHONE: LET THE COCKTAIL PARTY BEGIN

In order to understand what simultaneous interpreters' ears are subject to when interpreting "junk", platform sound, the problem must be characterized properly. The ability to separate and follow two or more acoustic signals has been investigated for decades and is known in the scientific literature as the "Cocktail Party Problem". At a cocktail party, in order to follow one conversation, we have to be able separate it from all the other conversations going on a the same time and from the background noise. Studies in this field have shown that sound localization/direction and spectral cues play a key role in performing this separation. When working, simultaneous interpreters also need to execute a so called "dichotic listening task" (Cherry, 1953), that involves separating and processing two different signals: a) the speaker; b) their own output; and perhaps even c) background noise if the source feed is dirty.

The best possible tool available to human beings to successfully separate the target signal from background noise at a cocktail party is binaural hearing (Hawley et a., 2004). In order to localize an acoustic source, the human brain contrasts interaural time differences (ITDs, the difference in arrival time of a sound between two ears) on the horizontal plane and interaural level differences (ILD, difference in loudness and frequency distribution between the two ears) to establish the position of the source on the vertical plane (Roffler et a., 1968). On the horizontal plane, middle and low frequencies are evaluated to locate the source, but high and extremely high frequencies also contribute to accurate positioning (Ibid.). On the vertical plane, our brain depends entirely on high and extremely high frequencies (anything over 7k Hz, Ibid.).

While sound localization based on high and very high frequencies has been shown to also work monaurally (i.e. using one ear only) both on the vertical and, to some extent, on the horizontal level (Butler et a., 1992), binaural localization is virtually impossible in the booth, as most interpreters harness auditory system laterality and rely on their "best ear" (Singer et a., 2012) to listen to the speaker and on their other ear to monitor their own output. This can be regarded as an external "mechanical" way of beginning to separate the two feeds. Even if interpreters do keep both ears covered binaural localization is still not possible, because interpreting consoles typically deliver a "double mono" sound instead of real stereo. Supplying a double mono feed to both ears might increase perceived loudness but also reduces the interpreter's ability to separate signals and also generates an artificial acoustic scene. Therefore the only localization (i.e. separation) mechanism available to interpreters while working is monaural localization based on high and very high frequencies.

High and very high frequencies are also used by our brains to estimate distance, as they are best heard if the speaker is nearby and facing the listener because of their shorter wavelenght (shorter sound waves travel less far than longer waves of equal amplitude). Absence of high frequencies therefore signals distance, while the presence of high frequencies means proximity.

Since videoconferencing and Remote Simultaneous Interpreting platforms remove all information above the 7-8k Hz threshold, monaural localization becomes impossible, and the source language signal is perceived as coming from the middle of a distant nowhere. Reverberation cues are also normally used to assess the sources' distance and position in space, but these are either absent in the RSI setting or extremely inconsistent. If all reverberation is absent, the brain receives contradictory information as in any enclosed space, a distant voice would normally produce a fair amount of reverberation. Inconsistent phase cues can also cause sound to appear to be moving in an erratic and difficult to determine pattern. Other spectral cues are also not so helpful when it comes to solving the "Cocktail Problem" in the RSI setting, because in addition to limiting bandwidth, platform algorithms also cause major spectral manipulation and what is fed into the interpreter's best ear is both limited and incoherent/unrealistic. The auditory centres performing sound localization in the midbrain (Szymanski et al., 2020) are at a loss, and the brain cannot solve the "Cocktail Party Problem". The speaker's voice remains bidimensional at best, and separating a foreground from a background becomes a daunting task.

On top of all this, a full-spectrum voice signal (the interpreter's voice) is masking a reduced-spectrum voice signal (the speaker's voice delivered by the platform). All of the above makes the reduced-spectrum signal even more difficult to process.

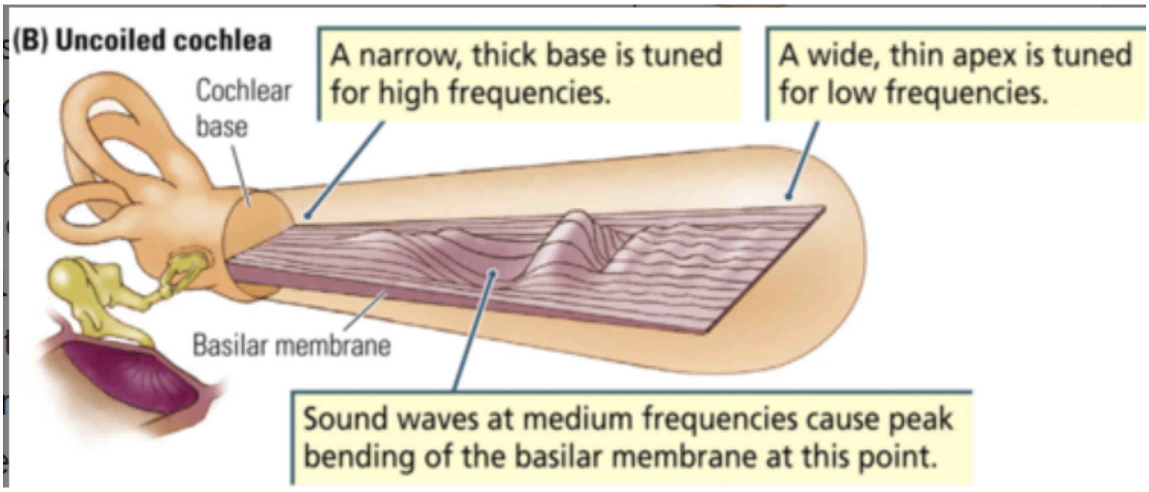
So what? How can all this cause hearing and vestibular problems?

Severe medium and long term hearing and vestibular problems can arise because of the sheer deprivation of high and very high frequencies, constant exposition to inconsistent spectrums and because of the interpreter's brain continuing failure to solve the "Cocktail Problem".

Use it or lose it

Frequencies above 7-8k Hz are perceived and processed by healthy individuals (including adults) to increase auditory performance, especially in challenging auditory situation (Lieberman et a., 2016). Their function also appears to be associated with arousal, attention and mood regulating mechanisms in the brainstem (see below). Loss of the higher area of the spectrum decreases auditory performance in challenging situations and can be regarded as a predictor of hearing loss in the lower range (20 to 7-8k Hz) (Hyun Joon Shim et al, 2009;

Moore et a., 2017). It is also regarded as one of the potential causes for hyperacusis (Wei Sun, 2009).



source: https://quizlet.com/198700897/lecture-15-temporal-lobes-1-flash-cards/

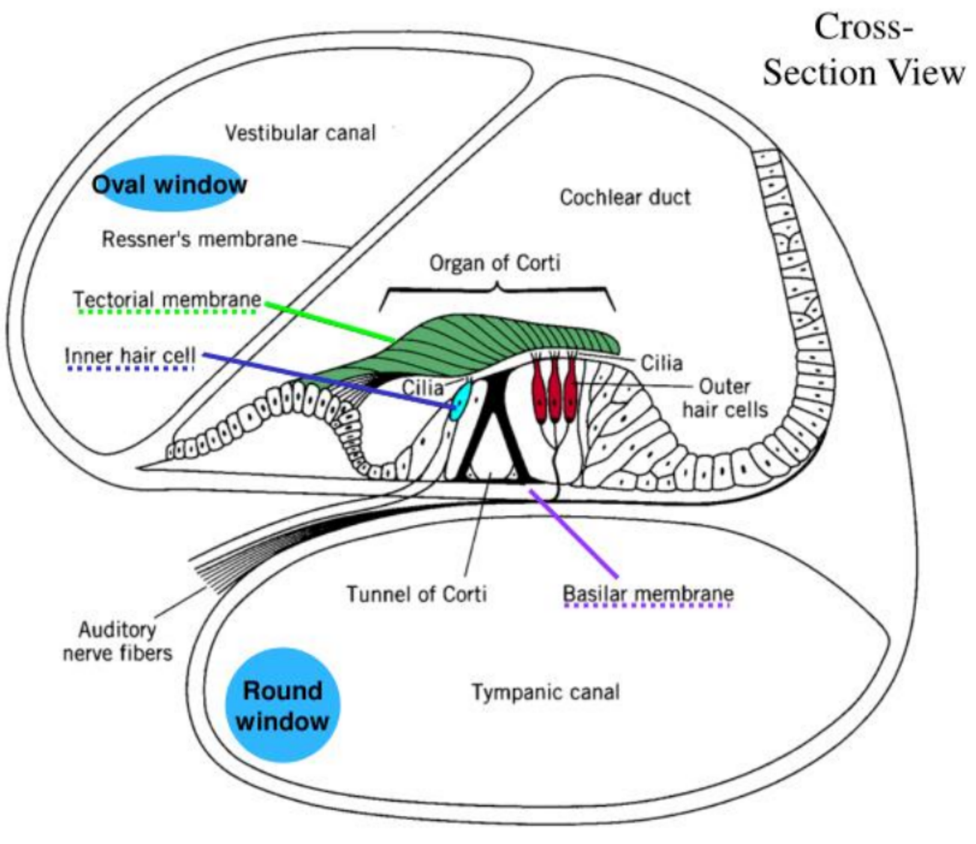
Experiments on conductive hearing loss indicate that blocking out specific frequency ranges leads to losing the ability to hear those frequency ranges in the medium and long term, as frequency deprivation leads to the death of the neural sensors (hair cells) of the mammalian cochlea (the frequency analyser in our inner ear) in charge of processing those frequency ranges (Lieberman et a., 2015). Since high and very high frequencies are not present in RSI sound and headphones (including open design headphones) prevent high and very high frequencies from reaching the ear from the outside as they flatten the outer ear and block its entrance, frequent exposition to reduced-spectrum VoiP/platform sound over headphones can easily be regarded as a form of high-frequency deprivation.

Compensation for loss

One typical reaction to the lack of high frequencies is turning up the volume knob: higher sound intensities stimulate the basilar membrane (the cochlear structure hosting the hair cells in the mammalian inner ear) in its entirety thus partially (but unspecifically) activating unused areas at the membrane's base (Ren, 2002) where higher frequencies are detected. Louder sound also causes the peaks of basilar membrane stimulation to shift backwards towards the higher-frequency area of the basilar membrane (Ibid.). Needless to say, turning up the volume only provides partial compensation for high-frequency deprivation and exposes the interpreters' ear to higher-than-needed average sound levels over prolonged periods of time.

When high-frequency information is absent, all information must be extracted from middle and low frequencies, that because of their physical properties must be amplified by the middle ear and require a discrimination effort by the tensor tympani muscle (especially in the 20-1k Hz range). The stapedius muscle also needs to allow larger stapes displacement (Greene 2017) to amplify lower frequencies, and it does so by exerting traction on the annular ligament. In so doing, the middle ear can amplify sound up to a factor of ten. Interestingly, otosclerosis is the consequence of calcification of the very same ligament (one of the proposed causes of tendon calcification is fiber degradation and wear and tear: Megumi Matsuda et a., 2018) and causes people to lose the ability to amplify those areas of the spectrum. As mentioned above, untreated otosclerosis can lead to permanent damage of the inner ear. Having to adopt this particular middle ear configuration for 7 hours a day, for a number of days a week might lead to highly undesirable outcomes.

Signal amplification also occurs inside the cochlea, where outer hair cells can increase gain levels (Motallebzadeh et al, 2018) up to a surprising 50dB (Byung In Han et al. 2009), especially when stimuli are faint, in the presence of background noise and when harmonic partials have been removed that in a redundant system would have helped unmask and discriminate the target frequencies. This mechanism is known as the “cochlear amplifier”. But hyper-functioning outer hair cells are believed to be among the causes of cochlear tinnitus (Hesse et al, 2008) and hyperacusis.



The cochlear amplifier. Source: https://www.slideserve.com/daniel_millan/presentation_118411

Cochlear lesions due to excessive sound intensity are a typical cause of tinnitus and loss of critical hearing bands (Ibid) and could also well be the result of functional imbalance and excessive sound amplification by the human ear itself to compensate for limited spectrum and to overcome the difficulties inherent in having to solve very difficult “Cocktail Party” problems.

Cocktail Party Fatigue

In order to accurately perform complex auditory tasks involving sound localization and target signal separation, our brain and ears need to mobilize all available resources:

The pinna (or auricle: the outer ear) is generally viewed as a high frequency amplifier (Roffler, 1968) that can be directed towards the source and/or moved so that the source sound will impact on its complex structures in ways that will select and amplify specific frequency ranges (Musicant et a., 1984). This can be done by moving the head and neck (including very slightly) and by performing micro-adjustments of its structural tension by means of facial and pericranial muscles and outer/inner pinna muscles. The latter are generally believed to be vestigial organs, but their neural connections to the brainstem, where sharp-tuning, orientation and sound localization are performed, are perfectly functional, and they appear to respond to sound stumuli (Iurkianets, 1973). The inner pinna muscles have been shown to act as sphincter of the auditory meatus (Matsuo et a., 1987) and at least one of the outer pinna muscles, the postauricular muscle, has been evidenced to react to high-frequency sounds (more on this muscle further down). The tone of pinna muscles contributes to giving the outer ear its sophisticated shape (Zerin et a., 1982), which plays a key role in the way the outer ear functions. A flattened pinna for instance reduces the human ear’s ability to process high-frequency sounds (Butler, 1992). The pinna muscles are supplied by the same nerve roots that also supply the facial muscles and the stapedius muscle (VII cranial nerve).

Postural changes also result in intracranial pressure adjustments (the pressure exerted by cerebrospinal fluid (CSF) that flows between the arachnoid and the pia mater) that have been shown to help optimize the listening and sound localization effort and to improve sound / frequency discrimination (Büki et al., 2000). These adjustments also seem to have a hydrostatic influence on the stapedius ligament via intralabyrinthine transmission (Ibid.).

Although receiving one of the target signals over headphones makes all of the above adjustments superfluous or much less effective, at least as far as the source language feed is concerned, the continuing and unresolved need to separate the two target signals keeps the entire system busier than normal and overworks the sensory apparatus.

Difficult listening tasks (when sounds to be detected are masked by interference and/or approach the hearing threshold, below which they cannot be detected correctly) also activate a number of pericranial and facial muscle reactions aimed at facilitating the relaxation of middle ear muscles (and thus contribute to mid and low frequency discrimination) and decreasing inner body noise (Stekelenburg et al., 2001) . Upper facial muscles like ***frontalis*** and ***corrugator supercilii*** will activate to support the listening effort, while lower facial, masticatory and pericranial muscles in general will need to relax (Ibid.) in order to achieve the same goal. But some of these muscles (like ***temporalis***, ***orbicularis oris*** and ***mylohyoideus***) are also involved in speech articulation. Difficult perceptual tasks would require them to perform accurate fine-tuning to help detect faint or otherwise difficult stimuli, but the constant vocalization and high-speed articulation typical of simultaneous interpreting drive the system in the opposite direction. Two conflicting tasks are being run at the same time and the auditory system must carry on solving a difficult cocktail problem while relying on an impaired function. Perichranial muscles need to relax to increase perceptual acuity, but again high-speed vocalization and articulation usually require significant activity from exactly the same muscles. Although the physical reactions described here have been shown not to be dependent on cognitive loads, pericranial muscles are also known to contract in response to increasing cognitive loads, which adds to the conflict. On top of all this, heart and respiratory rates are slowed down when the system needs to improve sensory acuity (Ibid.), but this is also difficult to imagine while a simultaneous interpreter is working. All in all, it's a little bit like requiring a professional athlete to compete in a hurdle race with one of his/her knees mechanically locked into a fixed position allowing only limited movement. It would make little sense and it would cause the athlete physical damage.

Needless to say, the only option left in a similar situation is to turn up the volume knob on the interpreting console, though it must be borne in mind that amplification is already happening inside the inner ear (as shown above). But a hyperactive cochlear amplifier (a biological function of the inner ear relying on mechanical stimulation of inner hair cells by outer hair cells, directly activated by the auditory brainstem) will in time cause inner hair cell damage owing to a) overstimulation of limited portions of the basilar membrane; even if intensity (measured in dB SPL) is assumed to remain stable, the same amount of sound pressure is concentrated on a smaller surface: a decrease in the area over which force is applied results in an increase in pressure on the area corresponding to the limited spectrum of RSI sound: fig.5); b) increased wear and tear; and c) the additional amount of inner amplification needed to unmask the target frequencies and resolve direct and indirect beating within the same or between related critical bands when the spectrum is inconsistent owing to arbitrary reconstruction/equalization by algorithms.

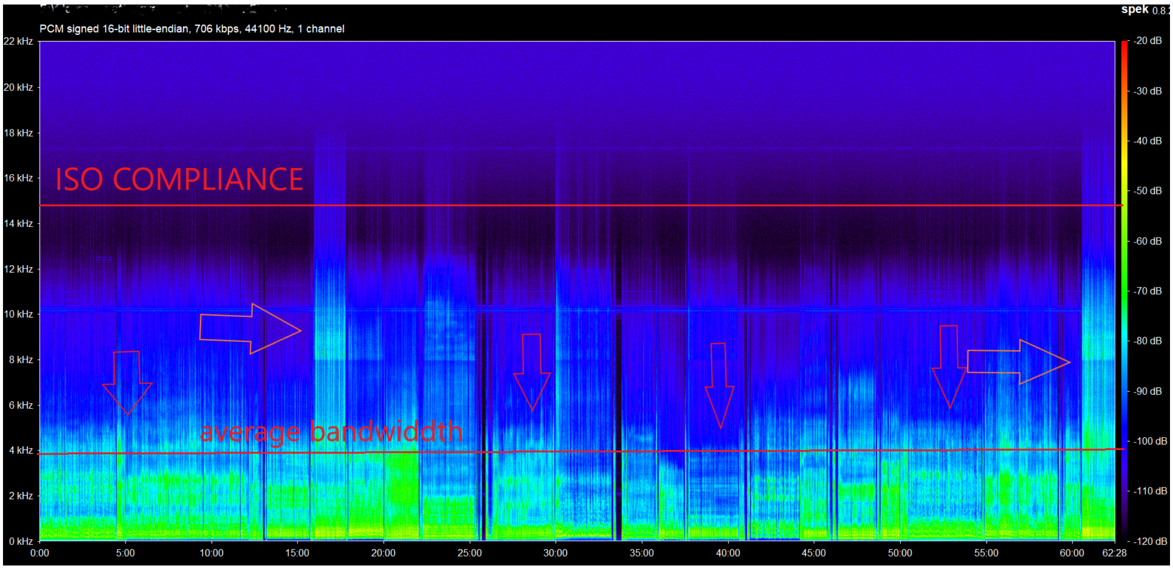


Figure 5: Spectrogram of around 60 minutes of "hybrid" parliamentary debate involving the use of a well known RSI platform. Normalized to -1 dB, 44.1 kHz, 16 bit. Areas in green and yellow show concentrations of spectral energy. Blue is not audible. Most spectral energy concentrates in the 0-4k Hz range. Orange arrows show "filler material" similar to the one shown in fig. 4.

In a similar context, <u>symptoms like hyperacusis, tinnitus and loss of hearing should come as no surprise</u>.

The Little Telltale Muscle

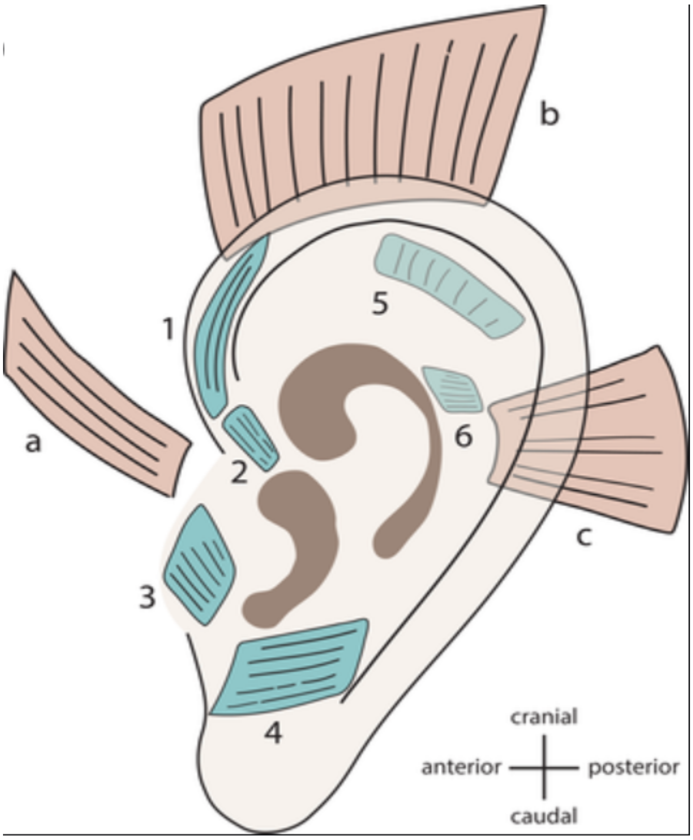
High-frequency deprivation also has a number of impacts on the interpreter's nervous system. Its consequences are better understood by studying the Postauricular Muscle Reflex (PAMR), a reflex causing the postauricular muscle to display contractile activity (evoked potential), and originating in the brainstem as a response to diverse stimuli. This particular muscle belongs to the same nervous circuit as the cochlea and the facial nerve, and its response is activated:

a) by erotic images and images of palatable food (Benning 2011, 2018);

b) by high and very high frequencies (Agung et al., 2005);

c) by "pleasant sounds" (Benning, 2011);

d) behind an ear performing an auditory detection task (here, reflex activity decreases with the difficulty of the auditory task) (Hackley et al., 1987)



Interestingly, besides activating the PAMR, erotic or appetizing images (a) also stimulate the production of dopamine in the brainstem's substantia nigra (Sonne et al., 2020). Dopamine in turn is known to regulate arousal and intracranial pressure levels. Experiments (Oohashi et a., 2000) indicate that full-spectrum sound, even beyond the 20k Hz threshold, is perceived as "more pleasant" (b). Pleasant stimuli are also associated with the production of dopamine.

Meinke et al., 2018, the PAM is "C"

In animal models, dopaminergic fibres have been found to prevalently originate in the high-frequency responsive lateral olivar complex of the brainstem, and to mainly project to the high-frequency, basal half of the cochlea (Maison et al., 2000), which makes the hypothesis that high and very high frequencies (b) also correlate with dopamine production very plausible.

Dopamine levels are low and the PAMR is absent in depressed subjects. Some signs and symptoms of conditions related to a dopamine deficiency include: muscle cramps, spasms, or tremors. aches and pains, stiffness in the muscles (which surely cannot help the middle ear and pericranial muscles perform complex auditory tasks). Low dopamine also generates a sense of frustration.

Dopamine has also been shown to modulate the function of neural cells reacting to low-frequency sounds and to protect hair cells from the lethal consequences of overstimulation, including by loud sound.

The Postauricular Muscle Reflex (PAMR), which originates in the brainstem thus seems to correlate with dopamine production (also a function of the brainstem) and one interesting theory (Hackley et al., 2015) postulates that the PAMR comprises one portion of a multifaceted orienting (again) system.

The brainstem's dopaminergic system is also an arousal system, and arousal appears to be key in governing attention (Dalton et al., 2009, Narayan et al., 2007, Conway et al., 2001). The Yerkes-Dodson law predicts that arousal will be optimal at moderate levels and that performance (including auditory performance: Narayan et al., 2007) will be poor when one is over- or under-aroused.

Interestingly, dopamine levels also appear to affect intracranial pressure (that in turn affects the way we use our head to optimize sound detection), as dopamine plays a role in keeping cerebrospinal fluid (CSF) pressure up (Altaf et al., 2017). Sudden loss of CSF pressure caused by CSF leaks notably generates symptoms including nausea, headaches, muffled sound, ringing in the ear and sense of imbalance, that are also reported by simultaneous interpreters required to work with RSI sound and with other, similar forms of "junk sound", and by other professionals working with VoiP sound feeds (Ayse Coskun Beyan et al. 2016).

A balanced, natural spectrum including high and extremely high frequencies and can legitimately be regarded as the lubricant that ensures the good functioning of a system that governs the interpreter's auditory perception, attention, orienting ability and mood. Lack thereof seems to have negative impacts on the vestibular system as well.

The labyrinth goes virtual

Be it from a fully-equipped ISO-compliant booth or from a laptop computer in the interpreter's kitchen, interpreting online videoconference streams means interacting with a low-quality virtual environment (Guiducci, 2020). While the image can be easily and stably positioned in space, sound is all but stable and, as shown above, virtually impossible to locate in space.

Minor sound level fluctuations are typical of VoiP feeds, and unreal/inconsistent (and therefore misleading) reverberation cues are the norm, which gives the listener's sound localization centres in the brainstem the impression that the source might be moving. Studies (Ystad et a., 2010) have found that moving acoustic stimuli can give listeners the illusion of self-motion (vection). A mixture of stable video, impossible-to-localize audio containing "moving" stimuli and lack of lip-sync (which is also typical of RSI) could contribute to explaining vestibular symptoms like nausea and loss of balance. Motion-sickness is normally viewed as the outcome of contradictory sensory stimuli being processed by the brain at the same time (Sensory Conflict Theory). Loud mid/low frequency sounds have also been found to stimulate the vestibular system (in particular, the vestibular postural reflex in the lower limbs) when delivered monaurally (Alessandrini et al., 2006). Given the characteristics listed above, the typical RSI sound might be capable of overstimulating or interfering with the normal function of the vestibular system.

Conclusions: Cocktail effects at the Covid19 Cocktail Party

Sound interacts with the human ear and brain in multiple and extremely complex ways. Junk VoiP sound is the staple diet of call centre workers, who are known to develop symptoms like tinnitus, otalgia, hyperacusis, fullness of the ear, nausea and vestibular problems (Ayse Coskun Beyan, 2016) that are often assumed to be the result of sudden peaks of loud noise. In only a few weeks, junk, VoiP sound delivered by videoconferencing and RSI platforms has caused a steep rise in the prevalence of the very same symptoms in conference interpreter populations required to adapt to the "new normal" and embrace a "new delivery mode" to ensure "hybrid" continuity of service in large multilingual institutions during the Covid19 pandemic (The Hill Times, 2020). Junk sound is an elusive beast and a multifaceted phenomenon, a toxic cocktail consisting of many different factors that would probably turn out to be harmless if assessed separately or require unrealistically high dosage or exposition to become the sole cause of the Junk Sound Syndrome. A long, non-exaustive list of these factors has been provided in this article to raise awareness in the interpreting community of the health problems that begin to appear under their combined impact of these factors.

ACKNOWLEDGEMENTS

Many thanks to Cristian Guiducci for helping me brainstorm and organize ideas and hunt through the scientific literature. Credit also goes to Ben Mulvihill for his feedback and proof-reading help.

REFERENCES

Agnew, 1998 The Causes and Effects of Distortion and Internal Noise in Hearing Aids, Trends Amplif. 1998 Sep; 3(3): 82–118

Agung K., Suzanne C. Purdy,Robert B. Patuzzi, Greg A. O'Beirne, Philip Newall, Rising-frequency chirps and earphones with an extended high-frequency response enhance the post-auricular muscle response, 2005, International Journal of Audiology;

Alessandrini M, Lanciani R, Bruno E, Napolitano B, Di Girolamo S (2006) Posturography frequency analysis of sound-evoked body sway in normal subjects. Eur Arch Otorhinolaryngol 263: 248–252.

Altaf F., DE Griesdale, L Belanger, L Ritchie, J Markez, T Ailon, MC Boyd, S Paquette, CG Fisher,J Street, MF Dvorak,and BK Kwon, The differential effects of norepinephrine and dopamineon cerebrospinalfluid pressure and spinal cord perfusion pressure after acute human spinal cord injury, Spinal Cord (2017) 55,33–38

Ayse Coskun Beyan, Yucel Demiral,1 Arif Hikmet Cimrin, and Alparslan Ergor, Call centers and noise-induced hearing loss, Noise Health. 2016 Mar-Apr; 18(81): 113–116.

Benning S. D: The Postauricular Reflex as a Measure of Attention and Positive Emotion, 2018, Oxford Handbooks Online

Benning, S. D. Postauricular and Superior Auricular Reflex Modulation during Emotional Pictures and Sounds, 2011, HHS Author Manuscripts

Büki, B., A Chomicki, M Dordain, J J Lemaire, H P Wit, J Chazal, P Avan, Middle-ear Influence on Otoacoustic Emissions. II: Contributions of Posture and Intracranial Pressure, 2000 Feb;140(1-2):202-11, Hear Res

BUTLER and HUMANSKI, Localization of sound in the vertical plane with and without high-frequency spectral cues, Perception & Psychophysics /1992. 5/ (2), /82-/86

Byung In Han, MD,corresponding author a Ho Won Lee, MD,b Tae You Kim, MD,c Jun Seong Lim, MD,d and Kyoung Sik Shin, Mde, Tinnitus: Characteristics, Causes, Mechanisms, and Treatments, J Clin Neurol. 2009 Mar; 5(1): 11–19.

Cherry, E. Colin (1953). "Some Experiments on the Recognition of Speech, with One and with Two Ears". The Journal of the Acoustical Society of America. 25 (5): 975–79.

Conway AR, Cowan N, Bunting MF (June 2001). "The cocktail party phenomenon revisited: the importance of working memory capacity". Psychon Bull Rev. 8 (2): 331–5.

Dalton, Polly; Santangelo, Valerio; Spence, Charles (2009). "The role of working memory in auditory selective attention". The Quarterly Journal of Experimental Psychology. 62 (11): 2126–2132

Guiducci C. Why Interpreter Hubs Cannot Fix Toxic Sound, 2020, https://www.linkedin.com/pulse/why-interpreter-hubs-cant-fix-toxic-sound-cristian-guiducci/?trackingId=AAR6plYGTwK0LnD5xBa7%2BA%3D%3D

Greene N. T., Herman A Jenkins, Daniel J Tollin, James R Easter Stapes Displacement and Intracochlear Pressure in Response to Very High Level, Low Frequency Sounds . 2017 May;348:16-30, Hear Res

Hackley, S. A. Evidence for a vestigial pinna-orienting system in humans. Psychophysiology, 2015 52(10), 1263–1270.

Hackley, S. A., Woldorff, M., & Hillyard, S. A. (1987). Combined use of microreflexes and event-related brain potentials as measures of auditory selective attention. Psychophysiology, 24(6), 632–647.

Hawley ML, Litovsky RY, Culling JF (February 2004). "The benefit of binaural hearing in a cocktail party: effect of location and type of interferer". J. Acoust. Soc. Am. 115 (2): 833–43.

Hesse, G., Andres, R., Schaaf, H. ***et al.*** DPOAE und laterale Inhibition bei chronischem Tinnitus. (DPOAE and Lateral Inhibition in Chronic Tinnitus) ***HNO*** **56,** 694–700 (2008).

Hyun Joon Shim, MD, Sun Ki Kim, MD, Chul Ho Park, MD, Sung Hee Lee, Sang Won Yoon, MD, A Ram Ki,1 Dae Han Chung, MD,1 and Seung Geun Yeo, MD, Hearing Abilities at Ultra-High Frequency in Patients with Tinnitus, Clin Exp Otorhinolaryngol. 2009 Dec; 2(4): 169–174.

ISO 20109:2016 Simultaneous interpreting — Equipment — Requirement

Iurkianets EA, Matiushkin DP. [Electrical activation of human external auricular muscles (at rest and during perception of acoustic signals)]. Bulletin of Experimental Biology and Medicine. 1973 Mar;75(3):16-9

Liberman C., Leslie D. Liberman, Stéphane F. Maison. Chronic Conductive Hearing Loss Leads to Cochlear Degeneration. PLOS ONE, 2015; 10

Liberman C., Michael J Epstein, Sandra S Cleveland, Haobing Wang, Stéphane F Maison, Toward a Differential Diagnosis of Hidden Hearing Loss in Humans PLoS One 2016 Sep 12;11(9)

Maison S., Xiao-Ping Liu, Ruth Anne Eatock, David R. Sibley, David K. Grandy, and M. Charles Liberman, Dopaminergic Signaling in the Cochlea: Receptor Expression Patterns and Deletion Phenotypes J Neurosci. 2012 Jan 4; 32(1): 344–355.

Matsuo K, Hirose T. Tragicus and antitragicus muscles as constrictors of the external auditory meatus. Eur J Plast Surg (1987) 10(2)

Meincke J, Hewitt M, Reischl M, Rupp R, Schmidt-Samoa C, Liebetanz D (2018) Cortical representation of auricular muscles in humans: A robot-controlled TMS mapping and fMRI study. PLoS ONE 13(7): e0201277. https://doi.org/10.1371/journal.pone.0201277

Megumi Matsuda, Asako Yamamoto, Jun Sasahara, Hiroshi Oba, and Shigeru Furui, quoting Hans Uthoff in Symptomatic calcification of the lateral collateral ligament: a case report, Acta Radiol

Moore, David PhD; Hunter, Lisa PhD; Munro, Kevin PhD, Benefits of Extended High-Frequency Audiometry for Everyone The Hearing Journal: March 2017 - Volume 70 - Issue 3 - p 50,52,55

Motallebzadeh, H. , Joris A. M. Soons,c and Sunil Puriaa, Cochlear amplification and tuning depend on the cellular arrangement within the organ of Corti, Proc Natl Acad Sci U S A. 2018 May 29; 115(22): 5762–5767.

Musicant A D, Butler R A. The influence of pinnae-based spectral cues on sound localization[J]. The Journal of the Acoustical Society of America, 1984, 75(4): 1195-1200.

Narayan, Rajiv; Best, Virginia; Ozmeral, Erol; McClaine, Elizabeth; Dent, Micheal; Shinn-Cunningham, Barbara; Sen, Kamal (2007). "Cortical interference effects in the cocktail party problem". Nature Neuroscience. 10 (12): 1601–1607.

Oohashi T, E Nishina, M Honda, Y Yonekura, Y Fuwamoto, N Kawai, T Maekawa, S Nakamura, H Fukuyama, H Shibasaki, Inaudible high-frequency sounds affect brain activity: hypersonic effect, 2000, J. Neurophysiol. Jun;83(6):3548-58.

Ren, Tianying, Longitudinal pattern of basilar membrane vibration in the sensitive cochlea, 2002 Proceedings of the National Academy of Science, vol. 99, Issue 26, p.17101-17106

Roffler, S. and Robert A. Butler,, Factors That Influence the Localization of Sound in the Vertical Plane The Journal of the Acoustical Society of America 43, 1255 (1968);

Shannon, F G Zeng, J Wygonski, Speech Recognition With Altered Spectral Distribution of Envelope Cues, J Acoust Soc Am 1998 Oct;104(4):2467-76.

Sininger and Bhatara, Laterality of Basic Auditory Perception, Author manuscript; available in PMC 2013 Mar 1. Published in final edited form as: Laterality. 2012 Mar; 17(2): 129–149.

Sonne J.; Vamsi Reddy; Morris R. Beato. Neuroanatomy, Substantia Nigra, 2020, StatPearls Publishing

STEKELENBURG J.J. and A. VAN BOXTEL, Inhibition of pericranial muscle activity, respiration, and heart rate enhances auditory sensitivity, Psychophysiology, 38 2001, 629–641. Cambridge University Press.

Szymanski; Geiger. Anatomy, Head and Neck, Ear 2020, StatPearls Publishing

The Hill Times, May the 6th, 2020, front page and page 5

Wei Sun The Biological Mechanisms of Hyperacusis, 2009, the ASHA Leader

Ystad S., Mitsuko Aramaki, Richard Kronland-Martinet (editors) Auditory Display: 6th International Symposium, Cmmr/Icad 2009, Copenhagen, Denmark, May 18-22, 2009 Revised Papers, 2010

Zerin M, Van Allen MI, Smith DW. Intrinsic auricular muscles and auricular form. Pediatrics (1982) 69(1):91

9. RSI Sound Myth Buster: Ten Misconceptions that result in RSI sounding terrible

Caniato, A. 26/06/2021

"Better" is the enemy of good (Voltaire)

True. Except when what you call "good" is harmful and "better" is well within reach. (Yours truly)

Author's note: Links containing video and sound clips are provided in this article: please use good, wired headphones (no earbuds, no in-ear headsets) to listen and fully appreciate their content.

Poor sound has proven to be one of the biggest nightmares in the videoconferencing and Remote Simultaneous Interpreting (RSI) setting. It makes listening unpleasant, causes meeting participants to tune out (bad sound causes listening fatigue) and makes simultaneous interpreting an arduous and hazardous business.

Poor sound undeniably hampers the interpreter's performance. Evidence gathered by various studies place poor sound on top of the list of suspects when it comes to the recent, major surge in hearing problems among conference interpreters (Reported Impacts of RSI on Auditory Health at International Organisations, Auditory Health Survey Canada), including debilitating and career-ending hearing conditions. Published scientific papers show that similar issues are not uncommon in other professions exposed to poor-quality sound over headphones (eg. call-centre workers) even when the use of peak limiters and compressors make sudden peaks of loud noise mathematically impossible. Conversely, an uncommonly high incidence rate of similar issues is not known to have been found among categories of professionals who are exposed to reasonable levels of high-quality sound over headphones (radio anchors, voice actors etc).

Videoconferencing and RSI do not need to sound artificial, robotic, tinny and heavy on the ear. Current technology and average internet connections already allow the transmission of decent image and, above all, pristine, radio-quality sound. So when your remote speakers sound like this instead of like this, it simply means that your remote event is not being organised and run properly with trained staff and the right platform / equipment.

As will be shown below, if technology is not the real hurdle, the problem is of a much more human and organisational nature. Following is a list of widespread misconceptions that stand in the way of interpreters getting the sound they need and deserve to stay healthy and deliver satisfactory quality to their listeners:

1) “SOUND IS GOOD WHEN I CAN UNDERSTAND WORDS WELL OR AT LEAST WELL ENOUGH, AND I DON’T MISS ANY CHUNKS OF INFORMATION”.

False. Sound is good when it is natural, and when listening is pleasant and completely effortless. When your feed sounds artificial but still remains intelligible and can be interpreted by making some sort of “extra effort”, a warning alarm should go off in your head.

Even when performed on perfect sound, simultaneous interpreting is the auditory equivalent of walking a tightrope, as - unlike other people - **interpreters have to understand their source while they are generating interference with their own voices**. Speakers typically sound “artificial” in the RSI setting because platforms save on bandwidth and server costs and create an environment where sound engineers are no longer necessary. Substandard microphones are allowed into the circuit and speakers are permitted, if not outright encouraged, to take the floor from noisy environments using whatever device they have available.

In order to make all of this possible, **RSI sound usually conveys a heavily reduced and processed portion of the original input, that is, the frequency content naturally present in the timbre of a human voice**. However, to manage multiple audio streams at the same time, interpreters need to rely on the richness and redundancy of a natural sounding voice. The way this works is explained here (scroll down to “On goes the microphone: let the Cocktail Party begin” to jump to the relevant section).

This is the main reason why simultaneous interpreting needs to happen within a strictly controlled environment. When exposed to typical videoconferencing / RSI sound, **interpreters put both their sensory and their cognitive systems under a great deal of pressure.** Conference interpreting turns into telephone interpreting and becomes the equivalent of walking a tightrope in high heels while juggling burning torches. As such, the type of damage that far too many colleagues are experiencing these days should come as no surprise.

The question therefore needs to change from “can I understand it?” to “**does this sound natural**?”. Sound is good (and harmless) when you can close your eyes and can say “Yes, this is what a real human voice would sound like”.

2) "WHEN SOUND ISN'T GOOD, THAT'S BECAUSE THE SPEAKER HAS A SLOW CONNECTION"

This is a narrative of convenience: "people have bad/slow connections, bandwidth is limited so quality sound is impossible". But when an Ethernet cable is used, **the average home connection in developed countries is powerful enough** to receive and broadcast both high quality video and sound, because **sound does not use up much bandwidth.** Video does. Therefore, allocating more/additional bandwidth to sound if perfectly feasible.

Full scale experiments conducted at major international organisations have clearly confirmed this, so it is rather difficult to believe that connectivity is the real issue here. Why then are RSI platforms almost invariably telling users that their connections are too weak? And why does this happen even in places where professional, corporate subscriptions guarantee huge download and upload speeds?

An educated guess would be that offering better sound would require platforms to process additional data to achieve a result that IT developers and marketing people don't really regard as necessary (sound is already "good enough", it's "speech optimised" or it's "more than enough to understand speech"). Providers (not users!) would need to allocate additional bandwidth/server/computing power at additional cost to them which is probably not a particularly palatable option for SIDPs. Why make a bigger organisational and financial effort, if you can sell it the way it is to clients who listen through their phones or computer speakers and aren't particularly "fussy" anyway?

And you can simply tell work-starved interpreters that people have weak connections, and that narrow-band headset sound is the gold standard they should aspire to.

3) "IN ORDER TO IMPROVE SOUND, SPEAKERS SHOULD USE A USB HEADSET WITH A BOOM MICROPHONE"

Absolutely false. 99% of **USB headsets come with low quality microphones** and onboard sound cards that **heavily process their input**. They are *designed* to be used for low-quality telephony applications on videoconferencing platforms that will typically not broadcast full-band quality sound. Why manufacture a Mercedes if users are going to drive it down a bumpy, unpaved country road?

Professional headsets with boom microphones cost hundreds and usually come with connectors that would be too complex for the average home setup. They need professional interfaces, pop filters and very accurate placing, as **a microphone positioned close to the mouth will pick up all sorts of annoying plosive and breathing sounds**, and will even make scratching noises against bearded cheeks. A boom microphone in the hands of an unassisted, inexperienced user is almost invariably a badly positioned microphone. So, if you expect people to simply put it on and use it, what **you have to do is process its input** heavily and remove a lot of the signal. **The result sounds artificial, robotic**, often sharp and heavy on the ear, especially when fed into a platform that reprocesses its input. This video contrasts the performance of a tabletop microphone and a USB headset that was provided to all members of staff working for a major international organization.

Why are platforms recommending USB headsets then? RSI platforms are run by software developers and marketing people and do not necessarily have sound engineers on their payroll. If I were to adopt the perspective of a developer working for a small/mid cap company I would probably think that if the microphone input is already narrow-band and what I consider to be useless information is filtered out, then less data is fed into the platform at the source. Data equals bandwidth, and bandwidth is a cost. Yet I have failed to consider whether the lost information is universally useless.

A much better option to obtain satisfactory quality sound is a **USB tabletop microphone** (excellent solutions are available starting from 50$ / 60€). These microphones are designed to produce radio/podcasting quality, and their onboard sound cards do not over process sound.

4) "A headset with a boom microphone is always better than no headset at all"

It really depends. What better USB headsets provide compared to bad laptop microphones is higher intelligibility. But as shown above, **intelligibility does not necessarily mean safety or quality**. Intelligible can still be harmful. When a headset mic is used, you might be able to soldier on through a presentation with less cognitive effort than when your speaker is using a really bad laptop microphone, but this will just help you tolerate a higher amount of toxic sound for longer, thus increasing your exposure. Moreover, many telephones/tablets and high-end computer mics (especially Apple) perform better (and process less) than a lot of USB headsets. Which means that **neither integrated mics nor USB headsets are a viable solution for RSI.** Tabletop microphones and selected clip-on microphones are statistically a much better option.

5) "Tabletop and lapel microphones can cause trouble if misused. USB headset mics are more reliable"

Tabletops and lapel mics can be misused like all other microphones, but when used properly they are the only proven way to deliver a rich, natural signal. Exactly like in the conference room, speakers talking too close or too far away from the microphone need to understand that they are doing it wrong. Exactly like in the conference room, no paper documents or other objects should come between their mouth and their microphone. Exactly like in the conference room, any background noise will be picked up by the speaker's microphone. **Speakers who can follow simple instructions will be able to manage tabletop and lapel microphones correctly. Speakers who cannot follow simple instructions are likely to mismanage all sorts of microphones**, including conference room goose-necks.

No matter how it is used, a USB headset mic will almost invariably deliver a heavily processed, artificial signal, because as shown above, **a USB headset mic is almost by definition a poorly placed microphone**. That is the reason why it comes with sound-processing onboard electronics in the first place, while tabletops / lapels do not.

6) "But hi-fi platforms and quality microphones are not *idiot-proof*"

No they are not. But neither are telephone-quality platforms and bad microphones. Is the expectation of having a fully plug-and-play, completely "idiot-proof" solution that guarantees quality, trouble-free videoconferencing (and simultaneous interpreting) even from the middle of the road legitimate?

Quality is achieved if sound checks are run in advance for every speaker and if microphones are properly configured with the help of remote sound engineers/moderators who actually know what they are doing. Simultaneous interpreters are dependent on sound quality like trout are dependent on cold, clean and clear water. Quality also requires a sufficient, and above all. effective use of *human* resources. Videoconferencing is, in general, a much more difficult environment than in-person meetings, so **expecting solutions that will work hassle-free and out-of-the-box in any situation with only minor human intervention is unrealistic.**

7) “Convincing speakers to use a headset is already difficult enough, so asking them to use a proper microphone is virtually impossible”

A logical fallacy. **Nobody really wants to look and sound like a call-centre worker** (no offense intended) **on camera** while addressing a conference that will likely remain on YouTube for the next 10 years, so the notion that speakers will drag their feet when proposed an unobtrusive device that will make them both look and sound professional (like they would on a TV or radio show) and would rather opt for a USB headset if they really have to is beyond all logic and understanding, but it’s a mantra I get a lot from USB headset prophets. A headset with a boom microphone is also not necessarily something most speakers will have available at their home/office and the less horrible-sounding models will cost as much (even twice as much) as a tabletop microphone with impressive performance. Practical experience shows that **when offered a real choice between a headset with a boom mic and a tabletop/lapel microphone, very few speakers opt for the headset solution**.

But **the biggest problem behind this misconception is the unquestioned assumption that speakers should be the one in charge of sourcing their own peripherals**. If we want RSI to sound good enough to allow safe and good quality simultaneous interpreting going forward, speakers must be provided with the right equipment by organisers and platforms. Decent lapel microphone solutions start from as little as 30 €, they fit in an envelope and **getting a parcel containing a small, 50$ plug-and-play USB tabletop microphone to the speaker’s location anywhere around the world and devoting 10 minutes to remote configuration support cannot be considered an insurmountable problem in 2021**.

Much more complex and expensive logistical efforts are usually made when organising multilingual in-person meetings. There is a huge difference between “technically or logistically impossible” and “we simply cannot be bothered with getting the right equipment, you just get used to it and do your best” or “not 100% compatible with a low-cost business model based on the delusion that quality equipment is no longer needed and nobody will ever notice the difference anyway”. **These attitudes have never been compatible with acceptable working conditions before**, let alone by AIIC.

8) “Noise cancelling is crucial if you want decent quality sound on the internet”

Nothing could be farther from the truth. No algorithm on earth will remove annoying background noise from a live audio feed without significantly affecting the quality of the signal. “Clean” sound does not mean you no longer get to hear any background noise: it means you hear a pristine, natural-sounding representation of whatever is picked up by a decent microphone. Any experts will tell you that active noise-cancelling is utterly incompatible with professional, hi-fi sound.

When they get an opportunity to try a good sounding videoconference on a clean platform, many colleagues wake up to the realisation that **if sound is rich and natural, background noise is usually not as annoying as they would expect.** But **in the typical RSI model**, where noise cancelling is aggressively performed by both headsets (or many integrated computer mics) and the platform, the resulting **signal is particularly poor** (lots of missing frequency information) **and muffled. At that point, any noise still making it through ends up being a much bigger nuisance than it should** because:

a) softer components get artificially pumped up by automatic gain control algorithms and become particularly disruptive and;

b) when you are struggling to keep a natural and therefore full-spectrum signal (your voice) from overpowering a muffled and heavily processed signal (your RSI feed), any additional noise becomes unbearable no matter how small it might be.

An audible example of the extent to which noise cancelling can degrade a good input from a good microphone is provided here (in this clip, noise canceling is applied using Krisp).

Yet we are bombarded by claims that you desperately need noise-cancelling to be able to interpret people who join a meeting while a vacuum cleaner is being used in the same room, dogs are barking, ambulances are passing by and loud construction work is being carried out outside the speaker's open window. While none of this is impossible, any similar situation would be an extreme nuisance for both the speaker and the other participants, and the idea of having to "interpret it anyway" while everybody else struggles to hear is hardly compatible with the notion of interpreters being highly-skilled and, above all, self-respecting professionals. As a matter of fact, these situations are not particularly frequent, and **the**

price for being able to solider on through a few isolated incidents by means of aggressive noise cancelling algorithms **is having to struggle with muffled sound the rest of the time**.

In reality, **background noise is probably more of a nuisance for platforms than it is for interpreters**. Harmless noise is "superfluous" information that codecs need to encode and broadcast. Noise removal at headset mic level proactively removes sizable chunks of the audible spectrum where noise (but also precious voice signal) can show up; it also reduces the frequency content of the chunks that get broadcast, resulting in… you guessed it: **lower bandwidth and server costs**.

Moreover, recently published research would seem to indicate that background noise is an obstacle for the human-machine interface. Given that a number of RSI platforms are known to be using your output to train interpreting machines, computer algorithms might be hampered by background noise much more than human beings. Unlike interpreters, **speech-to-text algorithms appear to like processed, muffled, telephone quality.**

9) "HI-FI QUALITY IS FOR MUSIC LOVERS, NOT FOR INTERPRETERS. WE ARE PROCESSING WORDS, NOT MUSIC"

Hi-fi means a high-fidelity reproduction of the original sound. The human voice produces much more information than the bare minimum needed to understand speech in an otherwise silent environment. Voice is a multilayered, redundant signal where the same information is repeated over and over again on different levels (harmonics) and **our ears harness this redundancy whenever we are required to perform a difficult auditory task** involving background noise or multiple signals building up a complex soundscape.

Simultaneous interpreting clearly qualifies as a difficult auditory task. People who lose their ability to hear high and very high frequencies (the part of the auditory spectrum that provides redundancy) struggle to process speech when concomitant sounds are present (read this to find out more). Spectral complexity is not just for the pleasure of demanding music lovers. It is a non-negotiable requirement for the performance of simultaneous interpreting. People who are forced to overspecialise their listening behaviour in order to compensate for loss of high and very high frequency tend to develop hyperacusis. (published paper here)

10) "OK, YOU WIN, MAYBE IT CAN BE DONE ON ZOOM HI-FI BUT ZOOM IS NOT AN RSI PLATFORM. RSI PLATFORMS ARE MUCH MORE COMPLEX AND CANNOT GIVE YOU HI-FI QUALITY"

Zoom currently accounts for a huge portion of online events and is currently used by many international organisations. Its Hi-Fi function, which speakers can quickly activate, works well but appears to have gone unnoticed in the language services industry. Interestingly, even WebEx has recently introduced a "music mode" which sounds better than regular WebEx (when a decent microphone is used) but still does not compete with Zoom Hi-Fi. **Big players are sensing the trend. Nobody wants to be listening to robotic sound for hours.** The gaming platform Twitch has also been offering and promoting high-bitrate audio for a while, as a way of keeping streams entertaining and preventing viewers from tuning out. Skype has already made an option available to deactivate background noise removal, although it hasn't begun offering a "music mode" yet.

RSI platforms with "high quality sound" releases specifically developed for individual institutional clients have been tested and used by some international organisations for a couple of months. When decent microphones are used, these releases deliver good quality audio, both on the floor and on over 20 different interpretation channels. With almost zero connection crashes and some packet loss due to poor WIFI. But not all that glitters is gold. Cases of platforms claiming that they can offer Hi-Fi quality (but obviously only if *your* connection is good enough, if *the speaker* has a good connection and is using a headset with a boom mic etc), or who even claim they can "improve" incoming feeds if they are not good enough, are also known. Quality can only be preserved, not improved.

Can a Hi-Fi platform giving interpreters radio-quality sound, guarantee a better-than-in-the-room interpreting experience at all times? It probably cannot. But it is a safer, less frustrating and more conducive tool to provide to a decent, professional output when interpreting remote speakers.

Credit: Thanks to Cristian Guiducci for his audiovisual wizardry, technical and sound-engineering advice!

10.Headsets Won’t Work Miracles: Here is How Digital Sound Gets Degraded in the 21st Century

Giaducci, C. 26/05/2020.

Highlights

- Sound quality in the booth is influenced by two main macro-components: a) the audio transmission chain; b) headphones and microphones. The audio chain is what can degrade sound the most, and yet for some reason the discussion seems to be focusing on headphones and microphones alone.
- Interpreters have no control of the complex transmission chain where most audio mismanagement can happen and often happens. This is the primary cause of poor sound in our headsets and remote simultaneous interpreting (RSI) platforms complicate things even further: Since they do without the necessary dedicated infrastructure, platforms have to rely on artificial intelligence and algorithms whose efficacy is by far insufficient to provide good quality audio.
- But if we still want to address peripherals (headsets and mics), two things have to be kept in mind: a) the main problems concern the sound feed coming from meeting participants, not from interpreters; b) one of the main purposes of manufacturer specifications is marketing, so manufacturer specifications and even ISO compliance have to be interpreted and understood within the proper technical context.
- Our ears are not ISO compliant machines. Auditory perception is an extremely complex “analogue” system and deserves to be treated as our supreme judge. No doubt, as any judge, it needs adequate training to distinguish poor sound from quality sound, otherwise it will never be able to tell good from evil.

Discussion

As a blind interpreter with an audio engineering background, I have to rely on good sound not only in the booth, but also in my daily life. Too often have I heard colleagues dismissing sound quality as a minor problem, and then making superhuman efforts to make sense of unintelligible speech owing precisely to poor sound. Even in the EU conference setting sound is often poor, and that in spite of the fact that headsets and microphones comply with ISO standards. A growing number of colleagues suffer from acoustic disorders like tinnitus, partial hearing loss, Menière syndrome etc. Canadian parliamentary interpreters also have experienced the consequences of poor sound, both before and after the transition to remote simultaneous interpreting.

Why can we get poor sound despite headphone compliance with ISO standards?

The transmission chain, i.e. everything that comes between the participants' microphones and the interpreters' headphones, usually introduces most of the sound alterations leading to poor or degraded audio. Merely concentrating on end user peripherals to improve sound quality is therefore pointless. Without an adequate and well functioning audio chain, no ISO compliant headset or microphone, cheap or expensive, can restore good quality sound.

LINK TO VIDEO EXAMPLE (2"30): poor vs good management of the sound chain

https://www.youtube.com/watch?v=GFJG7Jz7QmA&t=10s

Why is concentrating only on headsets or microphones not enough?

Until the late 90's, an analogue conference audio system - though full of cables and hidden gear - was relatively simple in terms of its electronic components. The typical chain from speaker to interpreter would consist of a room microphone, a low-noise preamplifier, a professional mixer, the interpreter's console (functioning as a headphone amplifier) and, to close the chain, the interpreter's headphones. In this "old school" setting, headsets undeniably were, along with preamplifiers, the weakest link in the chain.

Good quality microphones with good frequency response were rare and expensive. And robust, durable, lightweight open-back headphones with a wide frequency response were neither cheap, nor easy to manufacture.

But times change fast and today even a pair of well chosen and above all properly managed 2$ condenser microphones can produce studio-like stereo recordings of a symphonic orchestra.

How come then excellent sound gets butchered before it reaches our headphones?

Did you ever wonder why a friend's voice message sounds much clearer, crispier or pleasant to your ears than that long online meeting that left your ears and brain exhausted? That voice message is the result of a self-contained and above all, well designed and properly managed sound chain.

Now in the conference setting, and more so in the RSI setting, no matter how good and/or expensive the equipment you and other participants are using might be, the audio chain is often poorly managed: poorly tuned equalisers, compressors, limiters or feedback prevention mechanisms will still result in huge sound degradation even if on paper, the whole installation is ISO compliant. Furthermore, AI algorithms have no ears, so they have no idea what the final result of their activation will sound like.

Theoretically, modern technology enables the digital transmission of high quality sound either on site or across the globe with reduced cabling, cheaper setups and greater language regime flexibility and/or scalability.

However, audio chains between meeting participants and interpreters are very complex systems and are full of pitfalls. In large, institutional conference settings involving IP-based equipment, recent trends have seen their management outsourced to a remote location and there is not much sound technicians operating on site can do to improve things no matter how hard they try: access to advanced functions is restricted.

What is then so complex about these chains?

First, a preamplified speaker microphone output is transformed into binary code at a given sampling rate and with a specific bit depth by an analogue-to-digital converter. This digitised signal then undergoes digital compression via specific audio codecs relying on a compression algorithm with a given bitrate. Binary information travels on site through various network facilities. A mixing board is then used to process this data flow, and sound manipulation capabilities are endless: voice compression, echo cancellation, automatic gain control, noise reduction, automatic or feedback prevention, parametric equalisers with fine frequency band adjustment just to name a few. An exhaustive list clearly goes beyond the scope of this article.

Hundreds of different variables can exert a negative influence on the transmission chain. If this is not managed with great care, the sound quality delivered to our headphones can become much lower than that of the original sound.

What happens when we add remote participants or RSI to the mix?

On site, processed data is fed into our interpreting consoles, converted back to analogue sound by a digital-to-analogue converter (DAC), amplified and sent into our headphones.

But when a distant site is added, data will often travel thousands of miles through the internet and, here too, transmission protocol resilience, bandwidth capacity, overall network latency, average packet loss, all play a crucial role in determining what reaches the interpreters' headphones (and what does not), and in what condition.

Sound is usually muffled due to poor equalisation or exaggerated use of feedback control mechanisms. The overall frequency response on the interpreter's end of the chain is severely reduced by low sampling rates or low quality real time compression codecs. Destabilising speech artefacts originate from the use of noise reduction filters or echo suppression algorithms that will inevitably end up cutting out some useful voice information.

Microphone sound cuts due to poor network jitter can also puzzle interpreters and negatively affect speech intelligibility. The list could go on for pages… In such a complex situation, our modern microphones and headphones are very seldom the weakest link in the whole chain. Extreme sound quality degradation can occur at various levels throughout the digital path and our poor headphones cannot work miracles, no matter how expensive or funky they are.

Manufacturer specifications are marketing tools: they are not necessarily reliable

When sound is poor, lack of headphone compliance with ISO standards is very unlikely to be the primary suspect, especially when transmission chains are so polluted by overprocessing. But if we really insist on talking standards and headsets, the discussion should at least be technically sound and precise. First, real technical specifications must not be confused with marketing devices: manufacturers often take advantage of their customer's lack of technical expertise and make aggressive, dubious and misleading marketing statements that tend to overstate the real performance of their equipment. In their guidelines, interpreter organisations rely almost solely on raw frequency response, sensitivity and weight.

AIIC lately published recommendations on headsets almost exclusively focus on frequency response without an adequate discussion of other physical characteristics that make a real difference in the booth.

Headphones are for instance either open-back (or half-open) or closed-back (or half-closed). Interpreters should clearly prefer open-back headphones independently of their circumaural (on ear) or sovraural (over the ear) characteristics. Open back headphones notoriously provide a relaxed sound for long listening sessions, produce much less auditory fatigue, and interfere less with phonation. Using closed-back headsets, we can not hear our voice naturally through our eardrums an have to rely on bone transmission only, which typically leads to increased vocal effort to compensate for reduced proprioception: when one or both ears are covered by a closed-back headset, we tend to force our voice through the “barrier”, and this strains our vocal folds. Hearing ourselves well is crucial to control our output and prosody, and ensure customer satisfaction. Ironically and sadly enough, most headphones included in AIIC recommendations are closed-back or half-opened-back headphones at best.

As far as frequency response and marketing statements are concerned, 10$ mics or headphones have, on paper, the same specifications of microphones and headphones that cost hundreds of dollars, but of course sound completely different. It is basically like comparing lead to gold. You can hear the difference in this videoclip:

LINK TO VIDEO EXAMPLE ("2.23): Mics contrasted

https://youtu.be/OuGoZwrMYEY

Peripherals claiming a frequency response range between 100 and 18000 Hz might be ISO compliant and still sound horrible. Those figures alone will give no clue on how that mic or headphone sounds in the real world. Why? Because it is totally useless to say that a headset transducer can reproduce frequencies from 20 to 20000 Hz if no efficiency curve (plus or minus 3db) is provided, and the same is true for microphones.

This additional specification, which I could not find in AIIC's latest recommendations is absolutely necessary to understand how linear frequency response is and it indicates how flat the curve is at the lower and higher end of the frequency spectrum: on paper, I can easily claim that my headset can reproduce 20Hz frequencies from a church organ, but it might reproduce them 60 dB softer than the rest of the spectrum, so a human ear will most probably not be able to hear them. Conversely, if frequencies around 2000 Hz are reproduced 70 dB louder than frequencies ranging from 10000 Hz and above, it is very unlikely that the user will be able to use the upper part of the frequency spectrum. Quoting just a raw marketed frequency response doesn't mean good sound, though compliance with ISO PAS 24019 / ISO 20109 might be ensured.

Noise cancelling is not a desirable feature and often pure "marketing jargon".

Noise cancelling microphone capsules simply do not exist and noise cancelling headphones are not suitable for interpreters.

Marketing claims on noise cancellation need debunking. AIIC's recommendations include noise cancelling in the specs, but these specs are pure marketing devices, and they mean almost nothing.

So-called noise-cancelling microphones can be fitted both to USB headsets and to 3.5mm jack headphones. However, analogue headphones have no electronics and by definition no noise-cancelling function is possible, neither for headphones nor for microphones. Most of them use neodimium passive dynamic transducers in headphones and condenser microphones relying on plug-in power supplied by the connected sound interface. Thank God, no electronic circuit capable of noise cancellation is present: its function would hamper our ability to work well.

Indeed, some headsets can be fitted with a cardioid mic. Cardioid microphones have a directional polar pattern, with greatest sensitivity at the front and rejection at the back, but although this is marketed as "noise cancelling", it has nothing to do with noise-cancelling at all. If we elaborate further, USB headsets with built-in electronics, may theoretically have some form of active noise cancelling function both for mics and headphones, though this would require additional mics to create an "inverted phase" signal to cancel external noise. But in the booth, this is not desirable, neither for headphones, nor for microphones. Good noise cancelling requires very expensive, patented algorithms and its use should be limited to very noisy environments and most importantly, these algorithms usually create sound artefacts and generate additional pneumatic pressure on our ears.

Wrong technical claims are more dangerous than marketing jargon

The Recent AIIC Checklist on RSI reads: "❑ Does the RSI platform provide adequate protection against acoustic shock (at least 102 dBSLP peak loads as per G616 guideline or ISO 20109-compliant: 94 dBA SPL for any duration longer than 100ms)".

Though full of legal references and technical jargon, this is totally nonsensical and electronically meaningless.

An RSI platform is just a "middle man" and output limiting can never occur in the middle of the chain where these platforms operate. If you listen to your favourite radio station, you can, on your end of the audio chain, decide if you want to listen at comfortable headphone sound level or if you want to amplify the sound using your 300 watts loudspeaker system and blast your windows. If this happens, for sure you will not be able to sue the radio station for not limiting sound "pressure", which, by the way, doesn't even exist as such in the digital chain.

A pair of well trained human ears is our only supreme court

Training our ears to distinguish good from bad quality sound is essential and no algorithm or ISO standard will ever replace that. Stringent food labelling is very useful, but if you keep eating junk food manufactured in compliance with all the applicable standards and rules day in and day out, no “labelling” standard will protect your health.

To conclude on a positive note, As a blind interpreter I am grateful and open to all technological advances that can offer me good sound and help me protect my auditory system. ISO standards are a milestone and their further development will improve our profession even further. But extreme caution is needed: misinterpreting these standards or neglecting to monitor compliance therewith, relying on headphone limiters when literature has shown they do not solve problems, or limiting our analysis to marketing specs while neglecting muddy, distorted, muffled and artificially manipulated sound can be really dangerous and give interpreters a false sense of safety. Along the same lines, underestimating what happens down the digital audio chain and not knowing about its countless, sound-degrading algorithms, both in the conference and in the RSI setting, or ignoring frequent packet loss, which swallows fragments of useful speech, is extremely risky.

If we miss this fast moving target and fail to rise up to all these challenges, our ears will suffer a lot more in the not so "remote future", much more than we dare imagine today.

A great thank you to Andrea Caniato for the excellent peer review along with both graphical and textual editing.

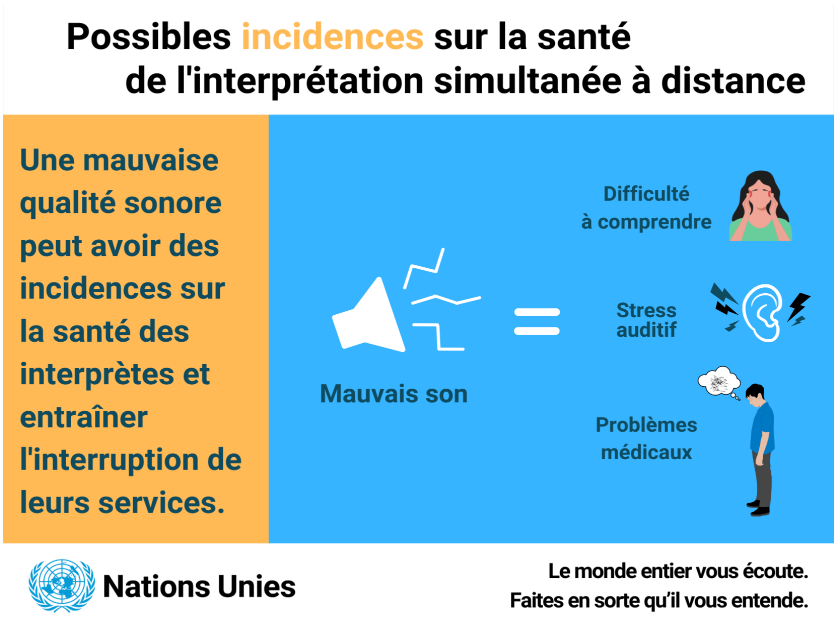
11.Echo des Cabines, AIIC n° 46, 02/2022 (extrait)

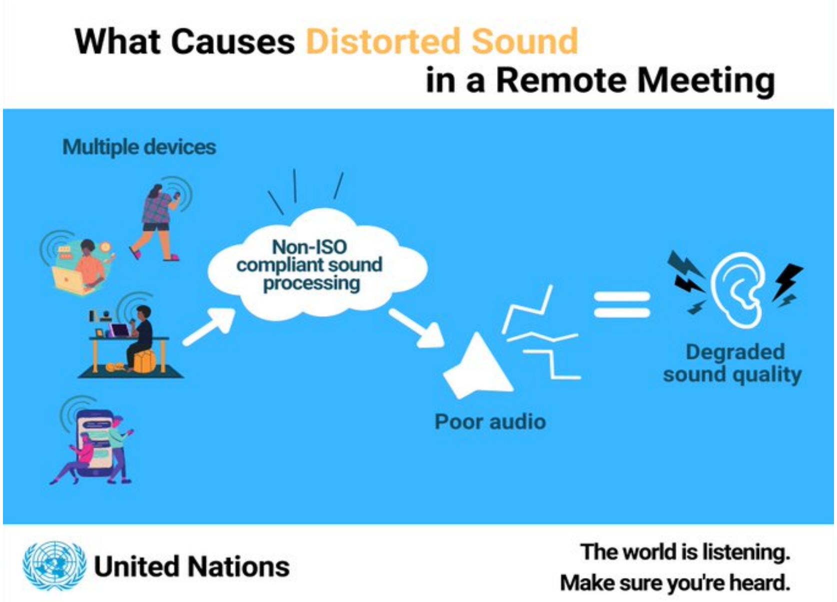
Point RSI et ses consequences

La santé auditive

…les conséquences de la RSI sur la santé auditive commencent à préoccuper non seulement les interprètes qui en sont les premières victimes, mais également leurs employeurs.

Voici une campagne de sensibilisation du Département de l'Assemblée générale et de la gestion des conférences DGACM des Nations Unies sur le son toxique et ses conséquences :





12. Media release CAPE (Canadian Association of Professional Interpreters)

L'ACEP présente une plainte contre le Bureau de la traduction pour manquement vis-à-vis de la santé et la sécurité des interprètes

Pour diffusion immédiate

OTTAWA, le 2 février 2022 – L'Association canadienne des employés professionnels (ACEP) a présenté aujourd'hui une plainte contre le Bureau de la traduction auprès du Programme du travail d'Emploi et Développement social Canada, car l'employeur n'a pas assuré la sûreté des membres de l'ACEP en milieu de travail.

Dans cette plainte, présentée au nom de ses membres interprètes issus de la fonction publique fédérale, l'ACEP affirme que le Bureau de la traduction a manqué à son obligation prévue à l'article 124 du Code canadien du travail, L.R.C. (1985), ch. L-2, qui énonce que l'employeur « veille à la protection de ses employés en matière de santé et de sécurité au travail ».

Plus précisément, l'ACEP fait valoir que le Bureau de la traduction n'a pas pris les mesures adéquates pour protéger les interprètes contre les blessures et les préjudices causés par la mauvaise qualité du son durant l'interprétation à distance, une situation qu'ils ont subie durant les deux dernières années, soit depuis que les activités du Parlement se déroulent en ligne.

Comme l'infrastructure technique est inadéquate et que les consignes ne sont pas respectées, les interprètes sont exposés à des risques considérables de subir des blessures propres à leur profession. Plusieurs interprètes ont signalé des blessures comme des chocs acoustiques, des maux de tête, des nausées et des acouphènes, qui, à terme, peuvent entraîner une perte auditive permanente. Durant les deux dernières années, un nombre anormalement élevé de rapports d'incident ont été recueillis.

« Bien que des échanges aient eu lieu pendant deux ans avec le Bureau de la traduction et la Chambre des communes, nos membres continuent d'être exposés à ces risques et pourraient devoir composer avec des effets irréversibles sur leur santé, a déclaré le président de l'ACEP, Greg Phillips. Assez c'est assez. C'est pourquoi nous avons pris la décision de présenter cette plainte; le Bureau de la traduction doit être tenu responsable et il doit régler ce problème. »

Cliquez ici pour en connaître davantage sur les risques pour la santé et la sécurité des interprètes depuis le début de la pandémie.

Chronologie des évènements

En mai 2020, **l'ACEP a tiré la sonnette d'alarme** à la Chambre des communes sur les blessures subies par les interprètes. Les **recommandations** de l'ACEP ont été incluses par la suite dans le **rapport de mai 2021** du Comité permanent des langues officielles.

Le 26 mai 2021, l'ACEP a publié les **constatations préliminaires d'un sondage** pour évaluer les risques sur la santé et la sécurité des interprètes dans le contexte du Parlement virtuel.

Plus de 60 % des interprètes représentés par l'ACEP ont répondu au sondage. Les constatations préliminaires publiées font état d'une situation déplorable :

- 92 % des répondants sont préoccupés par la perte auditive liée au travail qu'ils pourraient subir à l'avenir;

- 79 % d'entre eux se sont retrouvés dans une situation qu'ils considéraient comme dangereuse, selon le *Code canadien du travail*, en offrant des services d'interprétation simultanée à distance;

- 79 % affirment avoir présenté au moins un rapport d'incident pour des problèmes de son depuis mars 2020.

À propos de l'Association canadienne des employés professionnels (ACEP) L'ACEP représente plus de 21 000 employés de la fonction publique fédérale au Canada, ce qui en fait le troisième syndicat en importance de la fonction publique fédérale au pays. L'ACEP représente des économistes, des analystes de politiques, des chercheurs de la Bibliothèque du Parlement, des analystes du Bureau du directeur parlementaire du budget, des statisticiens, des traducteurs, des interprètes et des terminologues. **www.acep-cape.ca**

Suivez-nous sur Twitter, Facebook, LinkedIn et Instagram.

Pour en savoir plus à ce sujet, communiquer avec :

Katia Thériault, Directrice des Communications et affaires publiques,ktheriault@acep-cape.ca Téléphone cellulaire : 819-431-1015

13. Text sent out by the Canadian Bureau of Translation

In addition, as of February 7, 2022, all active (speaking) participants in virtual events with interpretation MUST use either a headset with integrated boom microphone or a unidirectional table-top microphone. Omnidirectional shared microphones, built-in computer microphones, wireless (Bluetooth) microphones and earbuds (Earpods, Airpods, iPhone headphones, etc.) are not permitted. If participants are not properly equipped, the following instructions must be followed:

If a speaker is not compliant with the new requirements, please interrupt service and make the following announcement on both channels (EN and FR):

French channel : « *L'interprétation ne sera pas* assurée *pour cet(te) intervenant(e), car il/elle n'utilise pas un microphone approprié. Les services reprendront dès que possible*. »
English channel : « *Interpretation will not be* provided *for the current speaker because a suitable microphone is not being used. Services will resume as soon as possible*. »

14. AIIC General Assembly Declaration on Auditory Health (adopted January 2022)

The 38th AIIC Assembly, meeting in Geneva from 13 to 16 January 2022, has adopted a resolution on distance interpreting (DI) stating the following:

“The Assembly is aware of alarming reports of negative health impacts attributed to distance interpreting, including a variety of auditory complaints such as tinnitus and hyperacusis, and other medical complaints including dizziness, nausea, confusion, mental fog, insomnia, headaches, concentration difficulties, optic nerve sensitivity, etc. The Assembly therefore advocates the precautionary principle in relation to DI and DI Agreements, and calls on the Executive Committee to ensure the commissioning of urgent research to be undertaken in this area.”

The Assembly, therefore, draws the attention of all professional freelance and staff conference interpreters, international organisations that employ interpreters, private-market entities that hire interpreter services and companies that provide equipment for conference interpreting to the increasing evidence that sustained exposure to substandard, compromised sound quality, especially in the remote simultaneous interpretation context, may cause serious damage to the hearing of conference interpreters, who - unlike other videoconference participants - must be able to decipher the audio signal above the sound of their own voice in order to perform their work, and produce other adverse health effects.

This damage may take the form of severe tinnitus (persistent ringing in the ear) and hyperacusis (acute and painful sensitivity to sound) as well as partial hearing loss, vertigo, acute migraines, eye problems and persistent disruption to sleep patterns. In many instances, such hearing damage is irreversible and, in some cases, it has already led to permanent disability, prematurely ending the careers of the interpreters so affected.

These problems increasingly appear to be the result of consistent exposure to digitally-altered, narrow-band, frequency-deficient and/or dynamically compressed defective sound (widely known under the umbrella term toxic sound because of its noxious effect on human health), also characterised by a typically low signal-to-noise ratio, which compels interpreters to increase the volume in their earphones to dangerous levels in order to decipher what is being said. Digitally-altered sound may also significantly aggravate unrelated auditory damage caused by such hazardous phenomena as acoustic shocks, which become much more likely to occur when the middle ear has already been made more vulnerable by consistent exposure to the compromised sound routinely experienced by interpreters working in the videoconferencing setting.

The Assembly calls for urgent medical studies and technical research into the precise nature of these auditory health problems and the precise technological and anatomical mechanisms involved in generating them. Pending the conclusive results of such research, the Assembly also calls for the application of the precautionary principle by employers of interpreters in ensuring shorter periods of exposure and much longer breaks between such exposures in order to allow for the ear to recover before being re-exposed, and it above all calls for the use, throughout the entire sound chain and especially by remote speakers, of microphones and other technical equipment required for audio transmission capable of reproducing the full ISO frequency response (125-15,000 Hz) and which do not manipulate the audio signal in any way. Such technologies and equipment already exist and are readily available.

The Assembly expresses its deep solidarity with, and strong support for, those conference interpreters who have already incurred damage to their hearing as a consequence of toxic sound and calls for the Association's relevant bodies to be granted the resources required to take every feasible step to protect the health and well-being of its members and to raise global awareness of these hazards among interpreters, international organisations, private-market employers and interpretation equipment providers.

15.WHO Make Listening Safe Initiative :

Over one billion people are at risk of hearing damage due to unsafe recreational listening practices. To combat these risks WHO created the **Make Listening Safe** initiative in 2015.

“Make listening safe” **aims** to realize a world where people of all ages can enjoy recreational listening without risk to their hearing.

The **approach** of this initiative is to change listening practices and behaviours. WHO aims to achieve this through:

- raised awareness about the need for and means of safe listening, and
- implementation of evidence-based standards that can facilitate behaviour change in target population groups.

The Make Listening Safe mission is performed via three main pillars, developed and carried out through in collaboration with all stakeholders in the field.

Creation of evidence-based standards

WHO creates standards that outline safe listening features for a variety of situations where unsafe practices are common. These include:

- the WHO-ITU Global standard for Safe listening devices and systems
- the Global standard for safe listening venues and events (link to be added)
- WHO offers support to its Member States, private sector entities, and civil society in adoption and implementation of these standards.

Increasing awareness

WHO develops and disseminates evidence-based awareness materials for safe listening. These include:

  - Be healthy, be mobile – A handbook on how to implement mSafeListening (link to be added)
  - mSafeListening message libraries (link to be added)
  - Media brief on safe listening (link to be added)
  - Communication materials such as flyers, posters, brochures, infographic

Investing in research

Research into safe listening is performed in collaboration with our global partners to better understand the current state of affairs, to ensure WHO leverages current best practices around the globe, and to uncover future need of safe listening interventions.

Useful links:

Link to the WHO-ITU global standard for safe listening devices (in support of the Make Listening Safe Initiative)

https://apps.who.int/iris/bitstream/handle/10665/330020/WHO-NMH-NVI-19.4-eng.pdf?sequence=1&isAllowed=y

WHO Make Listening Safe leaflet ("Once you lose your hearing, it won't come back!")

https://www.who.int/pbd/deafness/activities/1706_PBD_leaftlet_A4_English_lowres_for_web170215.pdf

Making Listening Safe:

https://www.who.int/activities/making-listening-safe/making-listening-safe