



Amélioration de l'intelligibilité de signaux vocaux automobiles par traitements adaptés aux capacités auditives

V. Roussarie^a, M. Daniel^a, L. Vauchier^a, D. Jacquet^a et J.-F. Sciabica^b

^aPSA Peugeot Citroën, 1, route de Gizy, 78943 Velizy Villacoublay, France

^bAIRBUS Operations S.A.S., 316 route de Bayonne, 31060 Toulouse Cedex 09, France

vincent.roussarie@mpsa.com

Les capacités auditives évoluent avec l'âge ce qui a pour effet de dégrader la compréhension de la voix, en particulier dans les milieux bruités. Ces dégradations peuvent être liées à différents facteurs, principalement l'âge (presbycusie) et les expositions sonores trop fortes et répétées. Pour autant, les interactivités sonores sont en constante augmentation, avec notamment une forte inflation des échanges vocaux dans l'automobile. Ces échanges peuvent être de nature sécuritaire (systèmes d'aide à la conduite), fonctionnelle (GPS) ou de confort (téléphonie, radio). Un état de l'art intégrant l'âge des clients dans l'automobile montre qu'une forte proportion (on l'estime à 25% au moins pour les écoutes en milieu bruité comme l'automobile) ont probablement des difficultés de perception et d'intelligibilité des messages vocaux. Partant de ce constat, une première série de tests en condition automobile a été réalisée à partir de traitements classiques, à savoir une combinaison d'égalisation et de compression des signaux audio. Ces tests ont été réalisés sur 25 personnes dont la moyenne d'âge est de 50 ans. Chaque personne passe un audiogramme tonal puis vocal afin d'évaluer ses capacités auditives et d'intelligibilité. Le sujet a la possibilité d'améliorer lui-même l'intelligibilité des signaux sonores grâce à une interface graphique, précalibrée à partir de la norme ISO 2079-2000 en fonction de l'âge et du sexe du sujet, et qui adapte en temps réel les signaux aux réglages choisis. L'évaluation du gain d'intelligibilité obtenu se fait ensuite par un deuxième audiogramme vocal sur les signaux traités. Les résultats montrent une amélioration de l'intelligibilité avec une influence plus marquée pour certains critères de correction du son. Nous avons également observé les modifications les plus sollicitées en fonction des audiogrammes des sujets. Ces résultats encourageants nous ont motivé pour la proposition d'un projet ANR (AIDA, démarré en janvier 2014), avec l'ambition, d'une part, d'éprouver de nouvelles techniques de synthèse pour l'amélioration de l'intelligibilité, et d'autre part, de fournir des interfaces adaptatives pour la personnalisation des traitements.

1 Introduction

Avec le vieillissement de la population, les gênes sonores dues aux pertes auditives augmentent sensiblement. A l'heure actuelle, on estime que 16 à 17% de la population adulte européenne a une diminution auditive d'au moins 25 dB à la meilleure oreille [1], ce qui place la déficience auditive comme l'une des altérations sensorielles les plus répandues. Alors qu'il est tout naturel d'aller, dès le plus jeune âge, s'équiper d'une paire de lunettes correctrices, les problèmes d'audition ont encore une mauvaise image dans notre société (vieillesse, handicap). Tout comme la vision, l'audition baisse naturellement avec l'âge (presbycusie), mais d'autres facteurs tels que l'exposition au bruit, peuvent engendrer des troubles auditifs plus importants. Dans le même temps les interactions vocales augmentent dans notre environnement et dans l'automobile en particulier. Téléphonie, GPS, etc, sont autant de nouvelles fonctions interactives nécessitant un bon niveau d'intelligibilité.

Une solution pour améliorer ces interactions vocales est de réaliser en temps réel un prétraitement des voix diffusées afin de les rendre plus intelligibles par l'auditeur. Ces traitements des signaux numériques permettent d'en améliorer la perception dans un environnement bruité et de façon instantanée. Notre hypothèse est que les gains d'intelligibilité pourraient être sensiblement améliorés si les traitements des signaux sonores sont adaptés aux capacités auditives des auditeurs.

Afin d'initier cette méthodologie nous proposons dans la démarche présentée ici de tester l'efficacité de quelques traitements connus des signaux de voix en les calibrant au plus près de l'audition du conducteur. Cette étude ayant pour objectif de valider le potentiel de la méthodologie, nous développerons le concept plus largement dans le cadre du projet ANR AIDA [2].

2 Etat de l'art

La détection des troubles auditifs passe systématiquement par l'audiométrie tonale liminaire. Il s'agit de déterminer pour chaque oreille le seuil d'audibilité d'un son pur, aux fréquences 250, 500, 1000, 2000, 4000,

6000 et parfois 8000 Hz. La moyenne des seuils d'audibilité définit la perte moyenne et permet de déterminer le degré de surdité correspondant. Pour chaque degré de surdité, l'impact sur la perception de la parole est également décrit dans la Figure 1 [3]. Des méthodes permettant de mesurer l'intelligibilité existent et sont même déjà standardisées, comme par exemple l'indice d'intelligibilité dans la parole (SII : Speech Index Intelligibility) [4]. On constate cependant qu'elles ne sont pas encore utilisées de manière systématique et que leur déclinaison automobile est délicate.

Seuil moyen	Classification	Effets
0-20 dB HL	Audition normale ou subnormale	éventuellement atteinte tonale légère sans incidence sociale
21-40 dB HL	Surdité légère	La parole est perçue à voix normale, elle est difficilement perçue à voix basse ou lointaine. Cependant La plupart des bruits familiaux sont perçus.
41-55 dB HL 56-70 dB HL	Surdité moyenne (1er degré) (2ème degré)	La parole est perçue si on élève la voix. Le sujet comprend mieux en regardant parler. Quelques bruits familiers sont encore perçus.
71-80 dB HL 81-90 dB HL	Surdité sévère (1er degré) (2ème degré)	La parole est perçue à voix forte près de l'oreille. Seuls les bruits forts sont encore perçus.
> 90 dB HL	Surdité profonde	La parole n'est plus du tout perçue. Seuls les bruits très puissants sont perçus.

Figure 1: Classification des pertes auditives

Le premier symptôme d'une altération de l'audition est la difficulté à comprendre la parole dans un environnement bruyant, ces symptômes apparaissant dès les premières altérations du système auditif; au regard de ce constat, l'automobile s'avère être un contexte aggravant pour les gênes sonores.

En se basant sur des résultats d'études sociétales et des données audiométriques collectées [5,6], mais aussi sur des interviews d'audioprothésistes et de malentendants, une première estimation des problèmes potentiels rencontrés dans l'habitacle automobile liés à des capacités auditives réduites peut être dressée. Globalement sur la perception des sons dans l'habitacle, on constate une audibilité réduite des sons fonctionnels (comme les sons d'alarmes et de clignotants), une baisse d'intelligibilité (conversations entre

personnes à bord, messages vocaux type GPS, téléphonie et radio), une augmentation du stress et de la fatigue due à une attention accrue, mais aussi une baisse de performance de l'activité de conduite. Une des études [5] montre que la perte auditive chez les personnes entre 62 et 88 ans a un impact significatif sur l'activité de conduite en présence de distracteurs visuels ou auditifs. Une autre étude [6] met en évidence une augmentation significative du risque d'accidents routiers pour les personnes exposées quotidiennement à un bruit supérieur à 100 dBA et/ou ayant une perte auditive supérieure à 50 dB HL (moyenne calculée à partir des seuils d'audibilité aux fréquences 3, 4 et 6 kHz) induite par l'exposition au bruit.

La première cause d'altération de l'audition est la presbycusie, liée au vieillissement. Avec l'âge le tympan devient plus mince, la chaîne des osselets dans l'oreille moyenne (qui transmet le mouvement vibratoire du tympan à l'oreille interne) se rigidifie, et surtout les cellules ciliées de l'oreille interne (en particulier les cellules ciliées externes, notées CCE) disparaissent peu à peu. Une perte de CCE induit :

- une perte d'audibilité qui commence aux hautes fréquences, car les premières CCE touchées sont celles qui transmettent les fréquences aiguës
- une réduction de la plage dynamique (appelée recrutement), car les CCE amplifient surtout les sons de faible niveau sonore
- une diminution de la sélectivité fréquentielle, due à l'élargissement des filtres auditifs.

L'audiométrie tonale ne permet donc pas de pronostiquer l'altération auditive véritable car elle ne révèle pas les phénomènes de recrutement et d'élargissement des filtres auditifs.

Les cellules ciliées peuvent également être endommagées de manière irréversible quand la personne est exposée à des bruits d'intensité élevée et de façon répétée. L'augmentation des seuils d'audibilité dans la plage 3 à 6 kHz, appelée encoche, est caractéristique de cette pathologie.

D'après les données audiométriques d'un panel de patients d'un audioprothésiste (base de données de 7345 audiogrammes réalisés entre 1997 et 2011), on observe que les personnes qui se font appareillées sont majoritairement atteintes d'une surdité moyenne (80 % des patients), c'est-à-dire ayant des pertes auditives allant jusqu'à 70 dB. Une première conclusion est que les personnes, qui ont des capacités auditives altérées, attendent d'avoir atteint une surdité handicapante avant d'aller consulter un audioprothésiste. De plus, selon une étude européenne de 2009 [7], le taux de pénétration de l'appareillage chez les personnes qui pensent avoir une perte auditive est le plus faible pour les 45-64 ans (23-24 %), tranche d'âge qui concentre 44,4% des clients de PSA. On peut donc conclure que pour l'essentiel, les clients de PSA ayant déjà des capacités auditives réduites ne portent pas d'appareils auditifs. Notre cible sera donc les clients PSA souffrant de gênes auditives légères voir moyennes et qui ne sont pas encore appareillés. On considère les personnes atteintes de surdité plus sévères, déjà appareillées.

3 Expérience

3.1 Procédure

La première phase est un test d'audiométrie tonale. Dans une seconde phase, une interface est testée avec chacun des deux bruits de fond, précédée et suivie à chaque fois d'un audiogramme vocal. Ces audiogrammes vocaux permettent de quantifier l'efficacité des traitements.

L'idée est de proposer une interface simple permettant à l'utilisateur de régler par lui-même la configuration qui compense au mieux ses gênes. Afin de calibrer les corrections à apporter, la méthode CTM [8] a été appliquée. Une étape de précalibration de l'interface a été mise en place afin d'aider le sujet à trouver le meilleur réglage (cf. Figure 2). L'objectif de cette méthode est de relever le champ auditif du sujet au niveau de celui d'un sujet normo-entendant grâce à deux étapes. Une première phase d'égalisation sert à augmenter les niveaux sonores en fonction des bandes de fréquences afin de ramener le seuil de confort du sujet au même niveau que celui d'un normo-entendant. Elle est suivie d'une étape de compression qui rétablit la dynamique conversationnelle du sujet, c'est-à-dire qui va atténuer les sons proches du seuil de douleur tout en conservant ceux proches du niveau de confort.

À la fin du test, le sujet répond à un questionnaire concernant l'utilisation de l'interface.

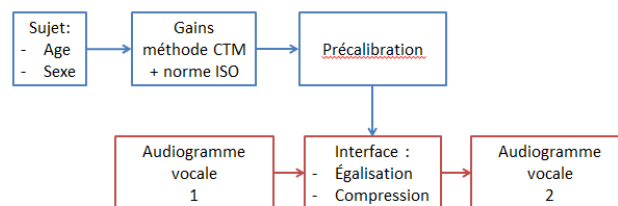


Figure 2: Interface individuelle proposée

3.2 Sujets

Afin de tester l'efficacité de notre interface, un panel de 25 hommes âgés de 45 à 60 ans a été sélectionné. L'âge moyen est de 50,2 ans et l'écart-type de 4,81.

3.3 Appareils

Les signaux audio ont été écoutés avec un casque Head Acoustics HA II. La chaîne de restitution est constituée d'un PEQ Head Acoustics 2482 PEQ IV, calibré à 74 dB pour l'audiogramme tonal et à 94 dB pour le reste du test, suivi d'un limiteur Head Acoustics 2485 PVA IV.2.

Le test se déroule dans une salle isolée phoniquement afin d'obtenir l'environnement le plus silencieux possible.

Une interface a été développée dans le but de permettre à l'utilisateur de procéder lui-même aux traitements. L'interface propose un pré-réglage adapté en fonction de l'âge et du sexe de l'utilisateur. Les gains d'égalisation et de compression de pré-réglage ont été calculés avec la méthode CTM sur chaque tranche d'âge et par sexe à partir des seuils liminaires fournis par la norme ISO 2079 – 2000 pour la presbycusie. Ces gains permettent un pré-calibrage de l'outil de traitement en fonction de l'audition donnée par la norme. L'égalisation a été réalisée à l'aide de deux filtres un passe haut et un passe bas. Dans les aigus, le sujet peut régler le gain de 1 à 100 (en amplitude), ce qui correspond à une augmentation de + 0dB à + 40 dB ; dans

les graves il peut régler le gain de 1 à 2, soit de +0 dB à +6 dB. Dans l'interface de réglage qui est utilisée dans le test, le réglage de l'égalisation se fait grâce à un panneau carré (cf. figure 3), qui permet de choisir indépendamment les deux gains (aigus et graves). Le gain du filtre agissant sur les aigus se fait sur la verticale, celui du filtre agissant sur les graves sur l'horizontale. Pour la compression on utilise une compresseur multibandes pour des bandes de fréquence situées à 125, 250, 500, 750, 1000, 1500, 2000, 4000, 6000, et 8000 Hz. Afin de conserver une voix naturelle, le temps d'attaque pour chaque bande est de 10 ms, et le temps de release de 200 ms [9]. Le sujet peut se déplacer entre les différents pré-réglages (en passant par les interpolations entre ces réglages) et choisir celui qui lui convient le mieux grâce à un curseur.

Une phase de correction qui n'appartient pas à la méthode a été ajoutée afin d'éliminer un souffle qui peut apparaître à cause des traitements précédents (égalisation, compression) et pour atténuer les sifflantes qui peuvent être trop agressives.

L'interface a été développée à l'aide du logiciel Max MSP [10] et est présentée dans la figure 3.

3.4 Stimuli

Les stimuli ont été réalisés avec l'enregistrement d'une voix de femme dans un fond sonore d'habitacle automobile roulant, avec et sans bruit de pluie successivement. La partie commune à ces deux fonds sonores est un bruit aérodynamique et de roulement, il permet de simuler un habitacle d'automobile silencieux (roulant à 120 km/h). Il a été choisi de ne pas ajouter de bruit de moteur car ce bruit n'est pas constant dans le temps et est trop variable d'un véhicule à un autre. Le bruit de pluie sera ensuite ajouté dans le second fond sonore car il s'agit d'une situation courante où le bruit est beaucoup plus masquant.

Les tests d'audiométrie vocale sont réalisés dans le même bruit que celui utilisé pour les réglages et les mots utilisés sont issus des listes dissyllabiques de Fournier calibrés à 60dB [10].

4 Résultats

4.1 Résultats sur l'audition

L'audiogramme moyen sur les sujets testés est représenté sur la figure 4 en noir et peut être comparé avec les courbes données par la norme ISO 7029-2000.

On remarque qu'en moyenne les sujets ont des pertes d'audition dans les graves qui semblent ne pas être dues à la presbycusie (probablement liées à un bruit de fond ou à

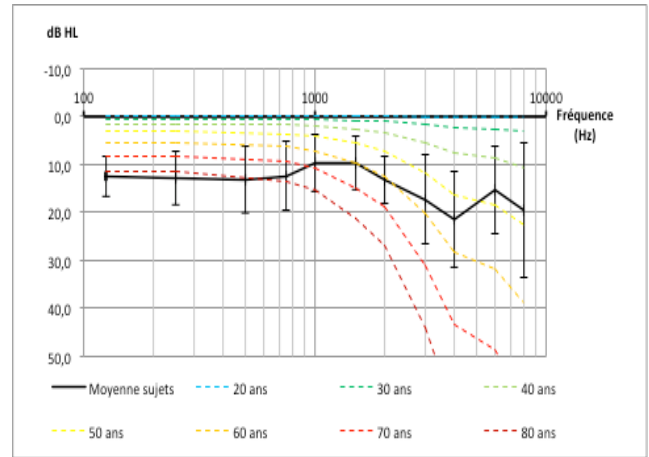


Figure 4: Audiogramme moyen des sujets testés en noir, comparé aux audiogrammes donnés par la norme ISO 7029-2000

la méthodologie lors la réalisation de l'audiogramme). Dans les aigus, les sujets, dont la moyenne d'âge est 50,2 ans, suivent approximativement la courbe d'un sujet de 50 ans avec de plus un pic autour de 4000 Hz. La présence d'un maxima à 4 kHz est bien connue des audioprothésistes, c'est une zone sensible de la membrane basilaire pour laquelle la perte des CCI/CCE est plus importante. C'est en général dû à des surexpositions à des bruits forts et répétées.

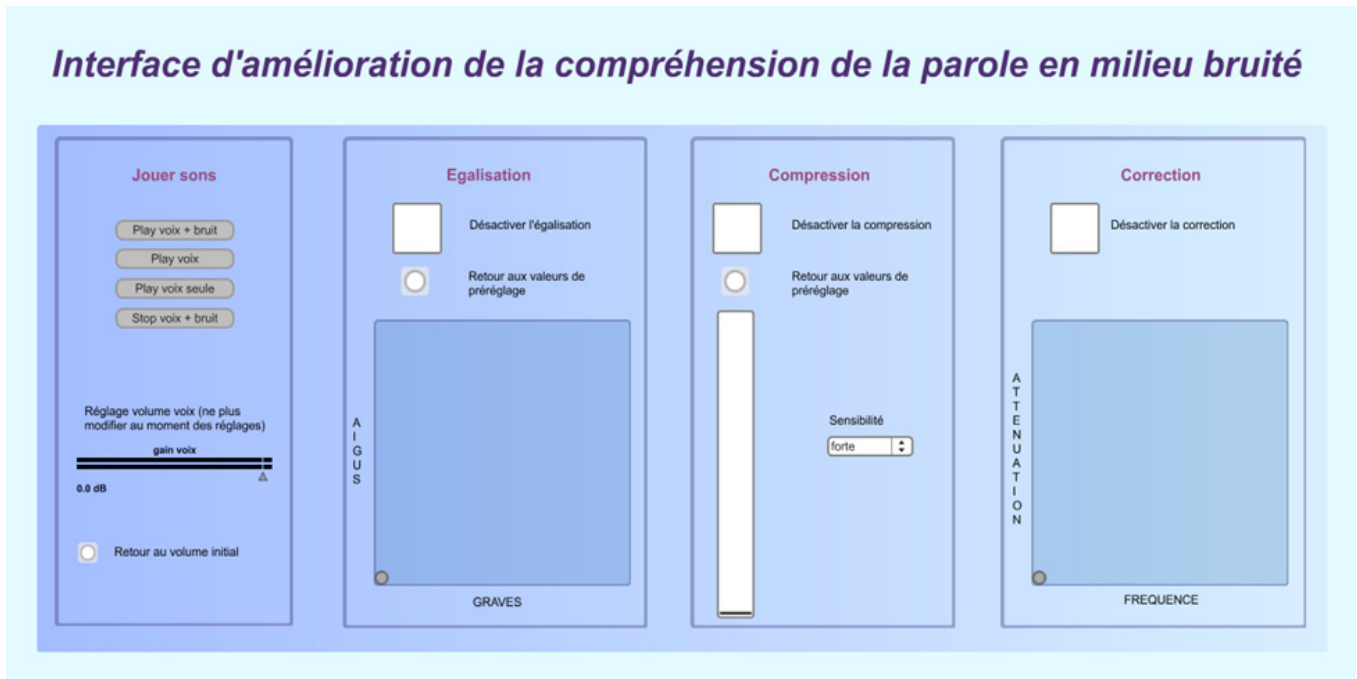


Figure 3: Interface individuelle proposée

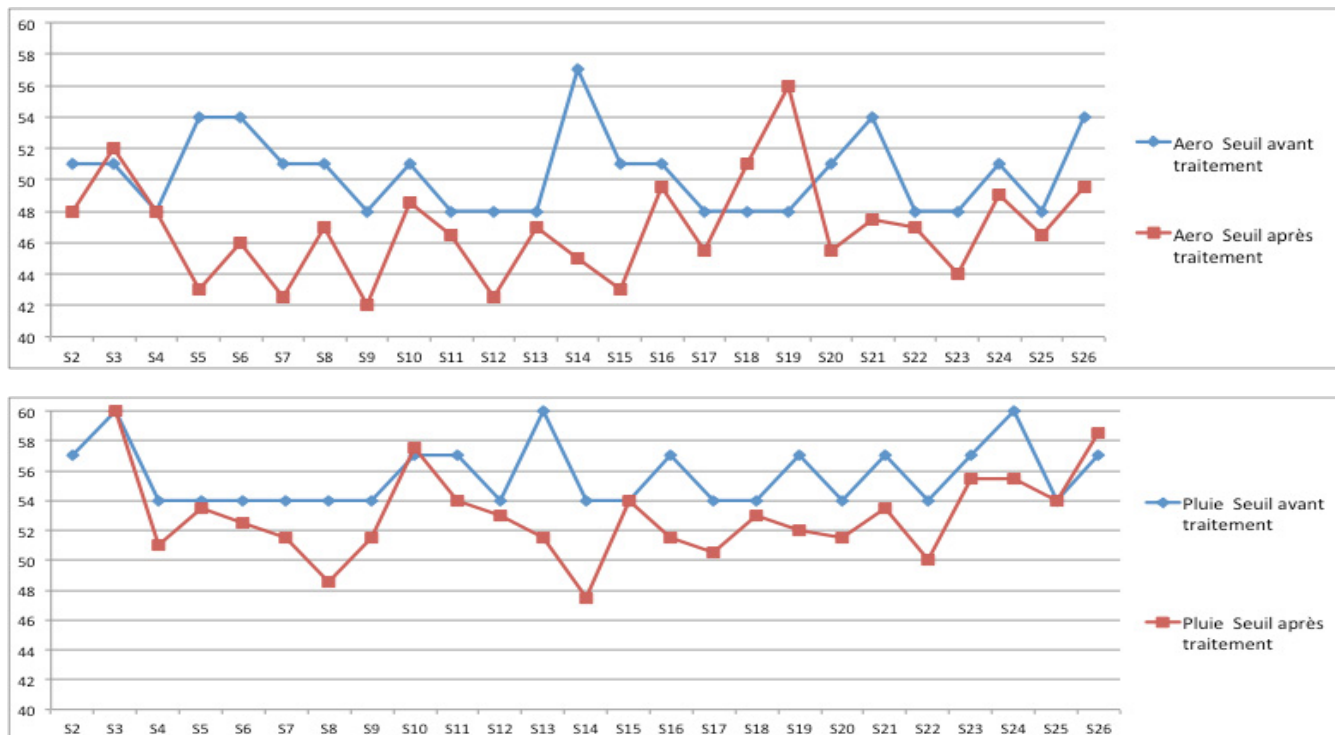


Figure 5: Résultats d'intelligibilité avant et après traitements, avec bruit aérodynamique (en haut) et avec bruit de pluie en plus (en bas). Augmentation des seuils après traitements dans la plupart des cas

4.2 Evaluation des traitements

On peut noter, sur la figure 5, une amélioration moyenne du seuil d'intelligibilité de 2,4 dB avec le bruit « pluie » et de 3,5 dB avec le bruit « aéro ». On peut supposer que le masquage supplémentaire dû aux impacts de la pluie rend l'amélioration plus difficile à réaliser. Cette amélioration varie cependant de façon assez importante selon les sujets : allant d'une dégradation de 8 dB à une amélioration de 12 dB pour le bruit « pluie » et une amélioration de 0 à 8 dB pour le bruit « aéro ».

4.3 Evaluation de l'interface individuelle

Un questionnaire a été rempli par chaque sujet pour évaluer les méthodes du test ainsi que ses interfaces. Les principales réponses ont été les suivantes :

- En général les réglages de la compression ont été trouvés beaucoup moins utiles à la compréhension que celui de l'égalisation. Il y a même eu des cas où leurs effets n'ont pas été entendus.
- La démarche la plus souvent mise en place a été d'identifier les effets en explorant les extrêmes pour ensuite revenir à des valeurs confortables.
- L'utilisation de l'interface a été simple, mais le choix d'un réglage adapté a été difficile
- Le test a paru intéressant à la plupart des sujets et plusieurs d'entre eux souhaiteraient disposer de dispositifs d'amélioration de la compréhension de la voix dans leur automobile.

5 Discussion

Notre échantillon de sujets, d'une moyenne assez jeune pour les cas de presbycousie (50 ans), présente pour la moitié une surdité légère, confirmant les statistiques et la part importante probable des clients de l'automobile gênés dans l'intelligibilité et potentiellement intéressée par nos traitements. Les résultats montrent également que dans la plupart des cas, une amélioration de l'intelligibilité de la voix est apportée grâce à l'interface proposée aux sujets. Bien que de nombreux progrès restent à faire et qui feront l'objet de nos prochaines recherches, le potentiel de ce type d'interface semble ainsi démontrée.

Cette amélioration reste légère puisqu'en moyenne proche de 4 dB dans le cas d'un bruit simulant l'habitacle d'un véhicule roulant, et de 3dB lorsqu'on ajoute le bruit de pluie. De nombreux facteurs permettent d'expliquer pourquoi le gain en intelligibilité est assez faible. D'une part, une grande partie de notre échantillon présente l'audition d'une personne quasi normo-entendante, le potentiel d'amélioration est forcément moindre pour ces personnes. D'autre part, l'interface présentée n'est pas optimale pour converger vers les meilleurs traitements. En effet, on peut noter que pour l'égalisation des aigus, les sujets ont en moyenne choisi un gain nettement plus fort que celui donné par la méthode CTM, ce qui n'est pas le cas pour l'égalisation des graves et la compression. De même, l'amplitude des réglages ne varie qu'assez peu entre les phases de test avec des bruits différents, ce qui ne semble pas forcément logique, le bruit de pluie étant plus masquant que le bruit aérodynamique seul. Enfin, de nombreux sujets ont déclaré ne pas avoir trouvé comment utiliser au mieux les réglages de compression.

De nombreuses voies d'amélioration se présentent donc à nous. La première étant bien sûr les traitements proposés, en effet cette méthode d'égalisation compression est celle la mieux connue et la plus utilisée. Il serait pertinent de tester d'autres types d'algorithmes. L'autre piste concerne l'interface et les leviers de contrôles mis à disposition afin

de mieux guider les sujets. Les méthodes de convergence des traitements ne sont pas optimisées ce qui entraîne les utilisateurs à pousser les réglages au maximum. Cela leur permet de comprendre l'utilité du traitement mais pas forcément d'avoir une meilleure intelligibilité. On pourra par exemple choisir des profils de compression plus pertinents pour les sujets. Une autre voie d'amélioration de l'interface est la mise en place de repères pour chaque panneau de réglage, qui ont manqué à certains sujets pour comparer les différents réglages qu'ils ont pu faire. On peut aussi noter que les sujets ont pu être influencés par l'ordre de présentation des panneaux dans l'interface lors de leurs réglages. Par exemple, presque tous les sujets ont commencé par régler l'égalisation en conservant un préréglage proposé pour la compression, puis ne sont pas revenus sur le panneau d'égalisation. On peut penser qu'un autre préréglage de compression aurait pu les mener à choisir d'autres paramètres d'égalisation. Une amélioration est donc à faire au niveau du protocole de réglage par l'utilisateur.

6 Conclusion

L'étude menée montre qu'il est possible d'obtenir une amélioration de la compréhension d'un signal vocal sans augmenter le niveau global de celui-ci, notamment grâce à des traitements d'égalisation et de compression adaptés à l'oreille du conducteur. Comme pour l'appareillage auditif des sujets, le préréglage ne suffit pas à avoir un traitement réellement adapté à l'audition du patient. Il s'agit d'un processus long entrepris par le couple audioprothésiste-patient afin d'affiner progressivement réglages et corrections. Il semble donc recommandé de suivre la même méthodologie avec les traitements que nous choisirons, en proposant un mode d'« apprentissage » des réglages, qui permettra de les faire varier dans le temps et de mieux guider les sujets dans leurs choix sur une période plus longue que celle testée. Cette première exploration a accompli son objectif en montrant le potentiel du traitement numérique pour l'intelligibilité et en mettant en évidence une population importante pouvant être ciblée. Elle suscite aussi de nombreuses questions que ce soit au niveau des traitements ou de l'interface et ses réglages. Nous évaluerons à l'avenir d'autres paramétrages des traitements existants et nous testerons de nouveaux algorithmes. Nous chercherons également à affiner nos méthodologies et interfaces utilisateurs afin d'optimiser les informations recueillies sans alourdir le processus pour le client. Enfin nous avons également constaté la nécessité de disposer d'outil de simulation des déficiences auditives afin d'alléger nos processus expérimentaux et de mieux calibrer nos prétraitements audio numériques.

Le potentiel étant démontré nous avons initié avec 3 laboratoires de renom un projet de recherche collaboratif proposant d'utiliser les méthodes et outils les plus en pointes dans la recherche, et de les adapter aux interfaces et aux besoins de l'automobile. Ce projet [2], AIDA (Automobile Intelligible pour Déficiants Auditifs), a démarré en janvier 2014 pour une durée de trois ans, il couvrira les domaines du traitement du signal, de la simulation des pertes auditives et enfin de la façon de présenter ces traitements de façon simple et intuitive aux utilisateurs via une interface (cf. Figure 6).

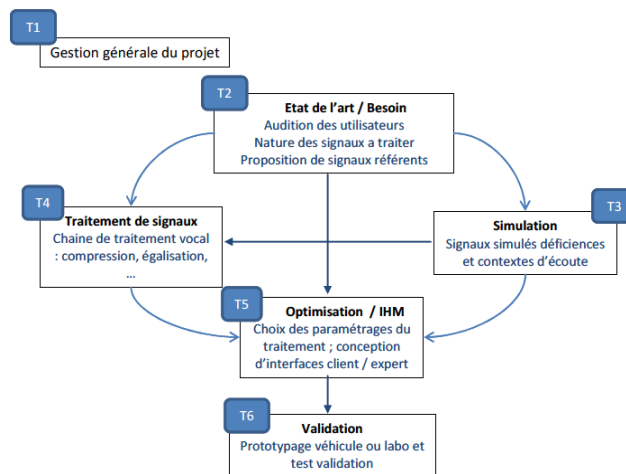


Figure 6 : Etapes du projet AIDA, Automobile intelligente pour Déficiants Auditifs

Références

- [1] B. Shield, *Evaluation of the social and economic costs of hearing impairment*, 2006.
- [2] CONTINT projet AIDA : http://www.agence-nationale-recherche.fr/projet-anr/?tx_lwmsuivibilan_pi2%5BCODE%5D=ANR-13-CORD-0001
- [3] Site <http://www.neurelec.com/fr/surdite/perce-d-audition>
- [4] ANSI: *ANSI S3.5-1997. American National Standard Methods for the Calculation of the Speech Intelligibility Index*. New York: ANSI, 1997.
- [5] L. Hickson, J. Wood, A. Chaparro, P. Lacherez, et R. Marszalek, *Hearing impairment affects older people's ability to drive in the presence of distracters*. J. Am. Geriatr. Soc., 58(6), 1097-1103, 2010.
- [6] M. Picard, S.A. Girard, M. Courteau, T. Leroux, R. Larouque, F. Turcotte, M. Lavoie, et M. Simard. *Could, driving safety be compromised by noise exposure at work and noise-induced hearing loss?* Traffic Injury Prevention, 9(5), 489-499, 2008.
- [7] Enquête *Focus sur la France des malentendants* Anovum 2009 European Hearing Instrument Manufacturers Association. EuroTrak France, 2009.
- [8] Interview de François Le Her par François Degove, *Question sur le CTM*, Les cahiers de l'audition – vol.21 – nov/dec 2008 – n°6 p. 39-53
- [9] Site http://www.ziggysono.com/htm_effets/index.php?art=c-ompression&titre=Compression
- [10] Site de Max : <http://cycling74.com/products/max/>
- [11] Listes dissyllabiques de Fournier : <http://www.college-nat-audio.fr/fichiers/img91a.pdf>
- [12] E. Parizet, *Evolution spatiale de l'intelligibilité de la parole dans un habitacle automobile*, JOURNAL DE

PHYSIQUE IV Colloque C5, supplément au Journal de
Physique III, Volume 4, mai 1994