

# Optimisation d'un critère d'intelligibilité de la parole dans un contexte bruité automobile

## *Speech intelligibility enhancement by equalization for in-car applications*

**Enguerrand Gentet**

LTCl, Télécom ParisTech  
Université Paris-Saclay  
75013 Paris  
et  
Groupe PSA  
Centre Technique de Vélizy  
Route de Gisy  
78943 Vélizy Villacoublay CEDEX  
E-mail : enguerrand.gentet@telecom-paristech.fr

**Bertrand David, Gaël Richard**

LTCl, Télécom ParisTech  
Université Paris-Saclay  
75013 Paris

**Sébastien Denjean, Vincent Roussarie**

Groupe PSA  
Centre Technique de Vélizy  
Route de Gisy  
78943 Vélizy Villacoublay CEDEX

**Résumé**

Ce travail s'inscrit dans le cadre de l'amélioration de l'intelligibilité des signaux de parole en contexte bruité automobile. Contrairement aux approches de rehaussement, il s'agit ici de réallouer l'énergie de la parole pour améliorer sa compréhension, mais sans modifier le Rapport Signal sur Bruit (RSB) global. L'approche étudiée consiste à effectuer cette réallocation en maximisant une fonction objectif basée sur le Speech Intelligibility Index (SII). Ce critère est calculé à partir d'une somme pondérée de RSB sur différents canaux fréquentiels et il est alors possible d'appliquer un égaliseur dynamique aux signaux de parole sous une contrainte énergétique donnée. À la différence d'autres méthodes, nous proposons une résolution exacte du problème d'optimisation du SII avec adaptation dynamique aux signaux (parole et bruit). Elle prend en compte la sensibilité acoustique de l'utilisateur en utilisant une échelle de puissance adaptée au contexte de manière à ne pas augmenter l'intensité perçue. Les résultats montrent une nette amélioration du critère d'intelligibilité et des tests d'écoute subjectifs viennent valider la méthode.

**Abstract**

In this paper, we propose a new speech intelligibility enhancement method for typical in-car applications in noisy environments. While traditional speech enhancement algorithms aim at increasing the Signal to Noise Ratio (SNR), the goal here is to increase intelligibility by applying dedicated voice transformation techniques without changing the original SNR. The proposed method consists in an adaptive equalizer which reallocates the energy of frequency bands to maximize the Speech Intelligibility Index (SII). The validation of the algorithm is carried out by means of a perceptual test derived from the Hearing in Noise Test (HINT) that demonstrates the merit of the algorithm for a high speed car noise.



La diffusion de signaux de parole dans les habitacles automobiles est sans cesse grandissante à travers notamment les télécommunications, les émissions de radio, les podcasts, les informations transmises par l'ordinateur de bord ou le système de navigation... Cependant, malgré les efforts et les avancées mécaniques, beaucoup de bruit persiste au sein de l'habitacle. Son origine peut être de diverses natures (roulement, moteur, frottements aérodynamiques...) et sa présence parasite très souvent la bonne écoute des signaux audio. L'objectif de ce travail est donc de traiter les signaux de parole avant leur diffusion dans l'habitacle afin d'améliorer leur intelligibilité en fonction du bruit ambiant.

En débruitage, et plus particulièrement dans le domaine de la réduction de bruit, la parole et le bruit sont mélangés dans le signal à traiter. Au contraire, lorsque le signal de parole et le bruit sont séparés, ce sont des méthodes de contrôle de bruit qui sont majoritairement employées.

Notre étude se place à la croisée de ces deux domaines en tentant de contrecarrer l'effet néfaste du bruit en agissant, en amont, sur le signal avant qu'il ne s'intègre au bruit. Contrairement aux approches de rehaussement, il s'agit donc de modifier les signaux de parole afin d'améliorer leur compréhension sans modifier le Rapport Signal sur Bruit (RSB) global. Ce domaine communément appelé *near-end listening* enhancement suscite un intérêt croissant [1, 2], d'ailleurs un challenge *Hurricane* [3] a été mis en place sur cette problématique lors de l'édition 2013 de la conférence internationale *Interspeech*.

Nous avons choisi le *Speech Intelligibility Index* (SII) [4] comme outil d'évaluation de l'intelligibilité des signaux de parole. Ce critère étant calculé à partir d'une somme pondérée de RSB sur différents canaux fréquentiels, notre approche consiste à appliquer un égaliseur dynamique aux signaux de parole afin d'optimiser le SII et, par suite, à améliorer l'intelligibilité des signaux au sens du SII.

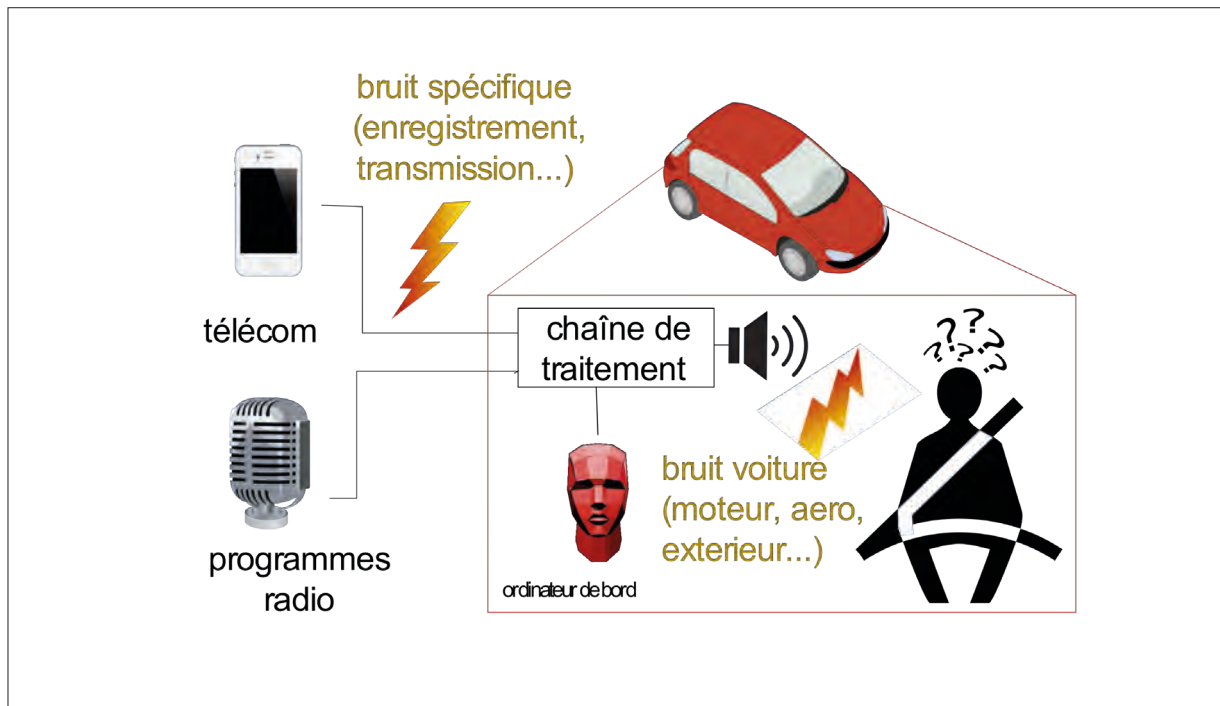


Fig. 1 : Contexte de l'étude qui cherche à contrecarrer l'effet néfaste du « bruit voiture » sur l'intelligibilité des signaux de parole en les traitant avant leur diffusion dans l'habitacle

Cette stratégie a déjà été suivie dans la littérature, que ce soit par une approximation linéaire du SII [5], ou une approximation non linéaire [6]. Dans tous les cas, la contrainte énergétique n'est pas pondérée et, au vu des spectres traités obtenus, il est probable que le niveau perçu ait été augmenté. Cette augmentation du niveau perçu pourrait expliquer l'augmentation significative de l'intelligibilité lors de tests subjectifs. Dans notre contexte automobile, si l'utilisateur souhaite une intensité perçue supérieure, il peut l'augmenter manuellement. Au contraire, si nous l'augmentons, l'utilisateur la ramènera à un niveau confortable. Ainsi, en plus de proposer une méthode de résolution exacte du problème de maximisation du SII, nous proposons une contrainte énergétique supplémentaire qui consiste à maintenir le niveau perçu constant. Les résultats montrent une nette amélioration du SII dans un contexte automobile à grande vitesse et des tests d'écoute subjectifs viennent valider la méthode.

L'article est organisé de la façon suivante. La justification du choix du SII comme mesure d'intelligibilité et sa méthode de calcul sont détaillées d'abord. La procédure d'optimisation et son exploitation pour traiter les signaux de parole sont présentés ensuite. Enfin, nous décrivons la validation de l'approche par des tests subjectifs et analysons les résultats obtenus. Une conclusion sur les performances et les limites de notre algorithme est finalement proposée avec des propositions d'améliorations et d'extensions pour le futur.

Par souci de clarté, de nombreuses simplifications ont été faites mais un lecteur intéressé pour avoir plus de détails est invité à lire la version étendue de l'article publiée dans les actes du congrès CFA 2018 [7].

## Mesure de l'intelligibilité de la parole

Plusieurs mesures objectives d'intelligibilité de la parole existent, chacune adaptée à des cadres bien précis. Les mesures d'intelligibilité s'apparentent souvent de deux mesures principales qui sont le *Speech Transmission Index* (STI) [8] et le *Speech Intelligibility Index* (SII) [4]. D'un côté, le STI et ses extensions s'appuient sur la Fonction de Transfert de Modulation (FTM) du canal de transmission et sont donc adaptés à l'étude de l'influence des caractéristiques du canal de transmission sur l'intelligibilité de la parole. D'un autre côté, le SII et ses extensions se basent sur le Rapport Signal sur Bruit (RSB) et sont donc plus adaptés à l'étude de l'intelligibilité de la parole dans un environnement bruité. L'environnement bruyant étant la principale cause de la baisse d'intelligibilité des signaux transmis dans un habitacle automobile, cela justifie le choix du SII comme mesure de référence et cet article s'appuiera sur sa définition originale introduite dans la norme ANSI/ASA S3.5 [4] en 1997.

## Méthode de calcul du SII

Les mesures objectives sont basées sur de solides connaissances empiriques et de nombreuses hypothèses. L'hypothèse principale du SII décompose la parole suivant  $i_{\max}$  canaux fréquentiels qui sont porteurs d'informations indépendantes. Ainsi, le SII est calculé à partir des niveaux du spectre équivalent de parole  $E_i$  et des niveaux du spectre équivalent de bruit  $N_i$  dans chaque bande, exprimés en décibels. Ces niveaux s'obtiennent en intégrant le périodogramme de chaque signal sur leurs canaux respectifs et en normalisant par la largeur de bande associée. À partir de ces spectres équivalents, deux coefficients par bande sont calculés: les coefficients d'audibilité et de distorsion.

Ces coefficients suivent des principes simples qui sont expliqués ci-après mais leur calcul fait intervenir des opérations non linéaires (prise en compte de la propagation du masquage, phénomènes de seuil...) qui ne seront pas détaillés ici. Les coefficients d'audibilité  $A_i$  peuvent s'apparenter à un RSB par bande, ils représentent la proportion du spectre audible au-dessus des diverses perturbations qui impactent l'intelligibilité. Les coefficients de distorsion  $L_i$  prennent en compte la distorsion introduite lorsque les niveaux par bande s'éloignent trop des niveaux d'un spectre équivalent de parole de référence.

Toutes les bandes ne contiennent pas la même quantité d'information relative à la parole, elles n'ont donc pas la même importance vis-à-vis de l'intelligibilité. Ainsi, une Fonction d'Importance de Bande (FIB), dont les coefficients sont notés  $I_i$  est appliquée pour pondérer chaque bande. Plusieurs FIB sont mises à disposition dans la norme ANSI/ASA S3.5 [4] en fonction du matériel vocal utilisé e.g. des syllabes sans sens particulier, des mots monosyllabiques ou des courts passages de discours.

La formule du SII correspond donc à une somme pondérée de ces différents facteurs dans chaque bande, comme indiqué par l'équation suivante :

$$SII(\{E_i\}, \{N_i\}) = \sum_{i=1}^{i_{max}} I_i \cdot A_i(E_i, N_i) \cdot L_i(E_i).$$

## Optimisation et traitement

### Définition et résolution du problème

Afin de prendre en compte la sensibilité acoustique de l'utilisateur, nous utilisons une échelle de puissance adaptée, le dB(A), de manière à ne pas augmenter l'intensité perçue. L'objectif de l'optimisation est donc de trouver les niveaux optimaux  $E_i^{opt}$  qui maximisent le SII pour un bruit donné (et donc des niveaux  $N_i$  fixes), sans augmenter le niveau perçu en dB(A) noté. On obtient alors un problème d'allocation de ressource classique, aussi appelé problème du sac à dos.

Comme indiqué précédemment, le calcul du SII fait intervenir de nombreuses opérations non linéaires qui rendent l'optimisation du critère complexe. C'est pourquoi, jusqu'ici, toutes les approches basées sur l'optimisation du SII passaient par des approximations linéaires [5] ou non linéaires [6]. Une originalité de notre méthode consiste à optimiser de manière exacte et permet une réallocation de l'énergie des signaux plus optimale du point de vue de la maximisation du SII.

En décomposant le problème en sous problèmes, il est possible d'utiliser un algorithme de séparation et évaluation, ou *branch and bound*, [9] pour choisir méthodiquement les branches à résoudre et converger vers la solution globale du problème. Chaque sous problème répond aux critères de résolution du problème continu convexe d'allocation non-linéaire sous contraintes et se résout donc directement en utilisant la méthode des multiplicateurs de Lagrange [10].

Avec cette méthode, le nombre d'itérations nécessaire pour la résolution du problème est trop important pour être effectué en temps réel. Par contre, l'optimisation ne dépendant que des niveaux du spectre équivalent de masquage, du seuil d'audition de l'auditeur et du niveau de référence de la parole en dB(A)  $S^{dBA}$ , il est possible de calculer préalablement les spectres équivalents optimaux si les profils sont connus et c'est dans ce cadre que nous nous plaçons dans la suite.

### Interprétation des spectres équivalents optimaux obtenus

Afin d'obtenir une visualisation complète du processus d'optimisation pour un bruit et un auditeur donnés, il est possible de calculer les spectres équivalents optimaux pour plusieurs niveaux du signal de référence  $S^{dBA}$ . Nous prenons un intervalle allant du niveau minimal pour obtenir un SII non nul, au niveau nécessaire pour obtenir un SII maximum.

Les résultats de l'optimisation obtenus pour un normo-entendant ainsi que l'allure du spectre équivalent du bruit automobile grande vitesse utilisé sont visibles Figure 2.

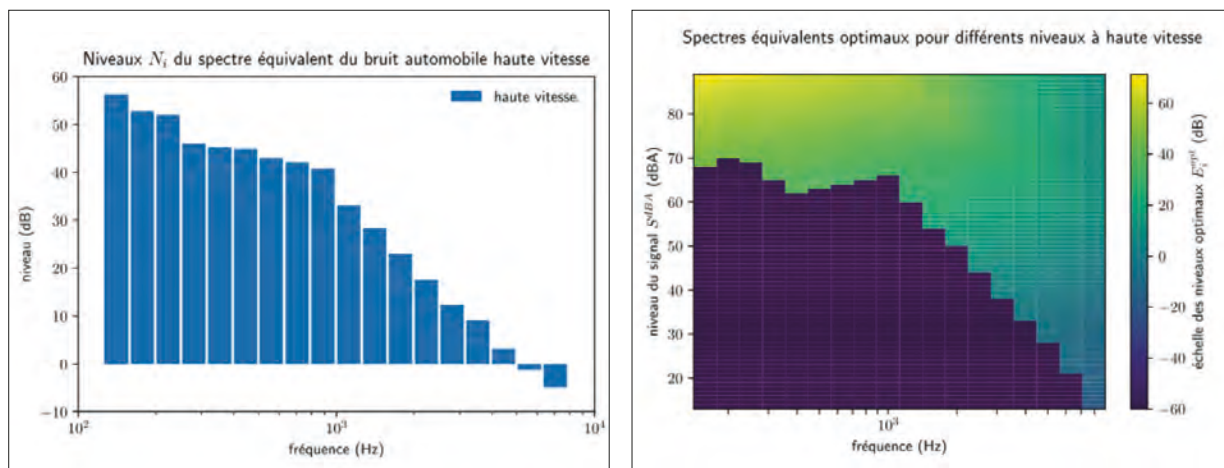


Fig. 2 : Comparaison entre les spectres équivalents optimaux maximisant le SII pour différents niveaux du signal à traiter avec le spectre équivalent du bruit donné

Le processus d'optimisation semble cohérent : l'activation des bandes de fréquences croît avec l'augmentation du niveau du signal  $S^{dB(A)}$ . Les premières bandes à recevoir de l'énergie sont celles dont l'énergie d'activation est faible i.e. où le bruit est peu présent. En effet, la répartition énergétique suit de près l'inverse du profil du spectre équivalent de bruit. On alloue alors de l'énergie aux bandes en fonction de leur rendement non-linéaire. Les différents facteurs de pondération introduits par la FIB vont aussi conditionner l'allocation énergétique en donnant plus ou moins de rentabilité aux bandes.

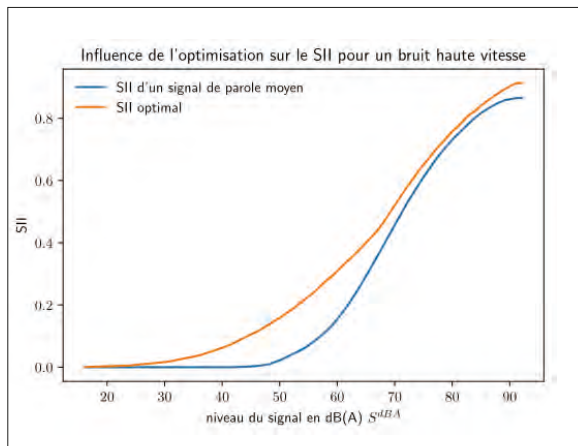


Fig. 3 : Évolution du SII en fonction du niveau du signal en dB(A) calculé à partir d'un spectre de parole moyen et d'un spectre de parole optimal

La figure 3 représente l'évolution du SII pour un signal de parole moyen à différents niveaux comparée à celle pour des signaux optimaux de même niveau en dB(A). On reconnaît les allures de sigmoïdes récurrentes aux courbes d'intelligibilité. L'amélioration du SII est notable, surtout dans la zone [30 dB(A), 70 dB(A)]. On obtient une augmentation maximum de 16% d'intelligibilité théorique pour un niveau de signal à 55 dB(A).

### Traitement de la parole

L'obtention des spectres équivalents optimaux est une étape importante dans notre approche et nous détaillons ci-dessous comment ils sont exploités pour traiter efficacement les signaux de parole afin d'améliorer leur intelligibilité. Nous avons exploré deux stratégies : une approche locale qui consiste à appliquer un égaliseur dynamique afin d'optimiser le SII sur de courtes fenêtres du signal, et une approche globale qui consiste à appliquer un égaliseur fixe calculé à partir des spectres équivalents long terme.

L'optimisation locale contraint, à chaque instant, le spectre équivalent court terme du signal à correspondre au spectre équivalent optimal. Cependant, si le signal n'a pas d'énergie, i.e. d'information, dans les «bandes optimales», ce sera le bruit qui sera amplifié pour atteindre les niveaux d'énergie optimaux. Même en utilisant des fenêtres de plus en plus larges pour lisser le traitement, la qualité du signal est largement impactée et l'intelligibilité n'en est que réduite.

L'optimisation globale consiste à calculer un filtre qui permet de passer du spectre équivalent long terme du signal vers le spectre équivalent optimal. Le filtre est calculé à partir d'un spectre équivalent de parole de référence normalisé au bon niveau perçu. Le calcul des niveaux du filtre équivalent en dB est alors une simple soustraction entre le spectre optimal visé et le spectre de référence. Ce filtre constant est ensuite appliqué sur l'intégralité du signal et c'est cette approche, plus douce, qui a été retenue pour traiter nos signaux.

### Tests subjectifs

Dans cette partie, nous détaillons les tests subjectifs mis en place pour valider notre approche et nous discutons des résultats obtenus.

#### Protocole

##### Principe

Le protocole mis en place suit majoritairement celui du *Hearing In the Noise Test* (HINT) [11] qui est un test d'intelligibilité dans le bruit destiné aux malentendants. De façon très simplifiée, il consiste à faire répéter des phrases noyées dans du bruit à différents RSB. Le niveau de présentation de chaque phrase dépend de la réussite du sujet à répéter la phrase précédente ce qui permet de converger vers un niveau qui caractérise l'intelligibilité des signaux. Ce niveau est appelé Seuil de Réception de la Parole (SRP), ou *Speech Reception Threshold* (SRT) et son estimation est détaillée par la suite. Compte tenu de notre contexte, plusieurs adaptations ont été faites au test, ces dernières sont présentées et justifiées ci-après.

##### Stimuli et présentation

Le matériel vocal choisi suit les recommandations du HINT [11] et est construit avec des phrases au vocabulaire standard représentatives d'un discours classique. Ce sont des phrases tirées du HINT franco-canadien [12] et des phrases de Fournier enregistrées par le Collège National d'Audioprothèse [13]. Le corpus est prononcé par un unique locuteur (homme) de sorte à pouvoir comparer les résultats.

Le type de bruit utilisé classiquement dans les tests d'écoute est synthétisé à partir d'un spectre de parole moyen appelé *Long Term Average Speech Spectrum* (LTASS) généré à partir de l'ensemble des signaux enregistrés par le locuteur. Cela a l'intérêt de proposer un RSB à peu près constant dans toutes les bandes de fréquence. Dans notre cas, le bruit d'étude est imposé par notre application automobile. Le bruit automobile grande vitesse utilisé est le même que précédemment, il a été enregistré en conditions réelles avec une tête acoustique (HMS IV, HEAD acoustics GmbH) et est retransmis en haute-fidélité à l'auditeur au moyen de l'interface dédiée et d'un casque audio (Sennheiser™ HD 650) calibré.

##### Équilibrage des phrases

Une étape cruciale dans la mise en place d'un test d'écoute est l'équilibrage du matériel vocal. Cela consiste à s'assurer que tous les éléments vocaux ont la même difficulté à être répétés dans le bruit considéré. C'est une étape qui demande généralement plusieurs itéra-



tions de corrections sur des groupes de personnes différentes [11, 12]. Effectuer cette étape sur l'ensemble des bruits automobiles sur lesquels nous travaillerons durant notre étude n'est pas concevable, c'est pourquoi nous avons décidé d'effectuer un équilibrage moyen sur un bruit synthétique dont le spectre correspond à la moyenne, en décibels, de bruits automobiles qui décrivent des situations à différentes vitesses par exemple. La méthode d'équilibrage employée est très proche de celle proposée par Nielsen et al. [14] pour la mise en place du HINT danois [15]. Elle a été réalisée en deux itérations avec deux groupes de six sujets chacun. À l'inverse des méthodes classiques, cette méthode fait intervenir le jugement subjectif des sujets dans l'équilibrage ce qui permet, théoriquement, d'obtenir de meilleurs résultats avec moins de sujets.

À l'issue de cette phase d'équilibrage nous avons créé dix listes de vingt phrases. Deux listes sont composées uniquement de phrases nécessitant un équilibrage trop important (correction supérieure à 3 dB), elles serviront de listes d'entraînement pour éviter que l'effet d'apprentissage n'impacte nos résultats. Les huit autres listes sont composées de manière à minimiser la variance des corrections, tout en essayant de conserver les phrases des listes originales ensemble car ces dernières sont équilibrées phonétiquement.

#### Procédure adaptative d'estimation du Seuil de Réception de la Parole (SRP)

Le SRP correspond au niveau de présentation le plus bas pour que 50% du matériel vocal soit répété. La procédure adaptative d'estimation du SRP qui a été choisie est celle détaillée dans les travaux de Brand et al. [16] qui est une généralisation de la méthode adaptative de Hagerman et Kinnefors [17]. Elle consiste à présenter une première phrase à un niveau très bas relativement au SRP supposé et d'augmenter progressivement le niveau jusqu'à ce que le sujet soit capable de répéter au moins un mot. Les dix-neuf phrases suivantes ne sont présentées qu'une seule fois et leur niveau dépend de la réponse donnée par le sujet pour la phrase précédente i.e. si plus de la moitié des mots a été correctement répétée le niveau de présentation est baissé, sinon il est augmenté.

Lorsque toutes les phrases ont été présentées, le SRP est alors estimé en utilisant un estimateur du maximum de vraisemblance. Cette procédure permet d'obtenir une estimation du SRP dont l'erreur est inférieure à 1 dB [16].

### Résultats

Les tests ont été réalisés sur une population de treize normo-entendants dont l'acuité auditive a été vérifiée par un examen d'audiométrie tonale [18]. Les tests se déroulent alors de la façon suivante. La procédure d'estimation du SRP est appliquée sur les deux listes d'entraînement afin d'habituer le sujet à la tâche. On applique ensuite la procédure à deux listes équilibrées en conservant la voix originale pour l'une et en la traitant pour l'autre. Il est important que les listes, les phrases et le traitement soient présentés de façon pseudo-aléatoire afin d'éviter l'effet d'ordre.

	SRP moyen (dB)	SRP écart type (dB)	Valeur p	Intervalle de confiance à 95%
Voix originale	57,88	1,07	1,42E-08	]-∞ ; - 5,90]
Voix traitée	51,01	1,84		

Tabl. 1 : Résultats des tests

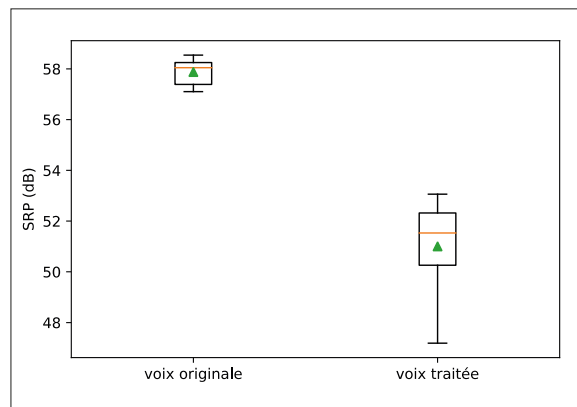


Fig. 4 : Profil statistiques des SRP estimés pour les signaux non traités et traités

Les résultats des tests d'intelligibilité ont été synthétisés dans le tableau 1 et par des diagrammes en boîte de Tukey sur la figure 4. On remarque une nette amélioration du SRP avec une diminution moyenne d'environ 6,9 dB. Un test de Student nous permet de justifier que les SRP estimés à partir du matériel vocal traité sont significativement et strictement inférieurs à ceux estimés à partir du matériel original. La borne supérieure de l'intervalle de confiance dit qu'au moins 95% des différences entre SRP devraient être d'une amplitude d'au moins 5,9 dB.

On remarque aussi une augmentation relative de 72% de l'écart type des SRP pour les voix traitées par rapport à celui pour les voix originales. Cela indique que des paramètres liés au matériel vocal ou aux sujets semblent influencer sur les performances du traitement.

Une question ouverte à chaque sujet concernant leur ressenti sur le naturel et la qualité de la voix n'a pas mis en évidence de différences majeures entre celle traitée et non traitée.

### Interprétations

En ré-allouant l'énergie vers les hautes fréquences où le bruit est beaucoup moins présent, nous favorisons l'émergence du signal au détriment des composantes basses fréquences qui, de toute façon, seraient masquées par le bruit trop important dans cette zone. Cette procédure permet d'améliorer grandement l'intelligibilité tout en conservant un niveau perçu constant.

En écoutant la voix traitée sans bruit, on est vite dérangé par l'absence de basses fréquences. Par contre, une fois la voix noyée dans le bruit, le retour des sujets semblerait indiquer que le traitement altère peu le naturel et la qualité. Des tests de qualité restent tout de même nécessaires afin de quantifier et valider ces impressions qui, si elles s'avèrent fausses, motiveraient la recherche d'un bon compromis entre traitement et qualité.

## Conclusion

En appliquant un égaliseur de fréquences à des signaux de parole de façon à maximiser le SII à intensité perçue constante, nous avons mis en évidence une approche permettant d'améliorer l'intelligibilité théorique de ces signaux diffusés dans un environnement automobile bruité spécifique. Le SII étant un critère mathématique objectif, des tests subjectifs rigoureux ont été mis en place afin de valider l'approche auprès d'une population de treize sujets normo-entendants. Les résultats obtenus montrent une amélioration de l'intelligibilité avec une diminution significative moyenne du SRP d'environ 6,9 dB pour le matériel traité par notre algorithme.

Par contre, la généralisation de ces résultats sur d'autres profils de bruit nécessite des tests supplémentaires. Il serait notamment intéressant de travailler sur des spectres plus équilibrés en fréquence et non stationnaires, comme en présence de pluie, ou en phase d'accélération.

Aussi, une spécificité du SII qui n'a pas été développée ici est sa prise en compte de l'audiogramme de l'auditeur : la méthode s'adapte naturellement à d'éventuelles pertes auditives. Des tests sur déficients auditifs sont alors prévus afin de valider la personnalisation de la méthode.

L'utilisation d'outils déjà existants pour compléter cette approche peut aussi être intéressante avec, par exemple, l'utilisation de la compression dynamique, une méthode très utilisée dans les aides auditives.

[11] M. Nilsson, S. D. Soli and J. A. Sullivan, "Development of the Hearing in Noise Test for the measurement of speech reception thresholds in quiet and in noise", *The Journal of the Acoustical Society of America*, vol. 95, pp. 1085-1099, 1994

[12] V. Vaillancourt, C. Laroche, C. Mayer, C. Basque, M. Nali, A. Eriks-Brophy, S. D. Soli and C. Giguère, "Adaptation of the hint (hearing in noise test) for adult canadian francophone populations: Adaptación del hint (prueba de audición en ruido) para poblaciones de adultos canadienses francófonos," *International Journal of Audiology*, vol. 44, pp. 358-361, 2005

[13] C. N. «CD d'audiométrie vocale.» 2006. [Online]. Available: <http://www.college-nat-audio.fr/listes-cd-audiometrie-vocale.html>

[14] J. B. Nielsen and T. Dau, "Development of a Danish speech intelligibility test," *International journal of audiology*, vol. 48, pp. 729-741, 2009

[15] J. B. Nielsen and T. Dau, "The Danish hearing in noise test," *International journal of audiology*, vol. 50, pp. 202-208, 2011

[16] T. Brand and B. Kollmeier, "Efficient adaptive procedures for threshold and concurrent slope estimates for psychophysics and speech intelligibility tests," *The Journal of the Acoustical Society of America*, vol. 111, pp. 2801-2810, 2002

[17] B. Hagerman and C. Kinnefors, "Efficient adaptive methods for measuring speech reception threshold in quiet and in noise," *Scandinavian audiology*, vol. 24, pp. 71-77, 1995

[18] B. S. Audiology, "Recommended Procedure for Pure-tone air-conduction and bone-conduction threshold audiometry with and without masking," 2012. [Online]. Available: <http://www.thebsa.org.uk>

## Références bibliographiques

[1] K. Nathwani, G. Richard, B. David, P. Prablanc and V. Roussarie, "Speech intelligibility improvement in car noise environment by voice transformation," *Speech Communication*, vol. 91, pp. 17-27, 2017

[2] G. Kim and P. C. Loizou, "Improving speech intelligibility in noise using environment-optimized algorithms," *IEEE transactions on audio, speech, and language processing*, vol. 18, pp. 2080-2090, 2010

[3] M. Cooke, C. Mayo and C. Valentini-Botinhao, "Intelligibility-enhancing speech modifications: the hurricane challenge.," in *Interspeech*, 2013

[4] A. N. S. I. ANSI, "S3. 5-1997, Methods for the calculation of the speech intelligibility index," *New York: American National Standards Institute*, vol. 19, pp. 90-119, 1997

[5] B. Sauert and P. Vary, "Near end listening enhancement optimized with respect to speech intelligibility index and audio power limitations," in *Signal Processing Conference, 2010 18th European*, 2010

[6] C. H. Taal, J. Jensen and A. Leijon, "On optimal linear filtering of speech for near-end listening enhancement," *IEEE Signal Processing Letters*, vol. 20, pp. 225-228, 2013

[7] E. Gentet, B. David, S. Denjean, G. Richard and V. Roussarie, "Optimisation d'un critère d'intelligibilité de la Parole dans un Contexte Bruité Automobile," in *Congrès Français d'Acoustique (CFA)*, 2018

[8] H. J. M. Steeneken and T. Houtgast, "A physical method for measuring speech-transmission quality," *The Journal of the Acoustical Society of America*, vol. 67, pp. 318-326, 1980

[9] S. A rali and J. Geunes, "Solving knapsack problems with S-curve return functions," *European Journal of Operational Research*, vol. 193, pp. 605-615, 2009

[10] K. M. Bretthauer and B. Shetty, "The nonlinear knapsack problem-algorithms and applications," *European Journal of Operational Research*, vol. 138, pp. 459-472, 2002