

Simulation d'un Echogramme par Recomposition de Deux Réponses Impulsionnelles: Application à l'Evaluation d'un Modèle Prévisionnel d'Intelligibilité

Robert Ruiz, LA.R.A., Université de Toulouse II,
5 allées Antonio Machado, 31058 Toulouse cedex 1, France (rruiz@univ-tlse2.fr)

Laurent Faiget, S^{té} 01dB,
6 av. Louis Blériot, 31570 Ste Foy d'Aigrefeuille, France (lfg@toulouse.01db.com)

Résumé: De la mesure de deux réponses impulsionnelles, respectivement avec un microphone omnidirectionnel et un microphone bidirectionnel, est déduit l'échogramme qui aurait été obtenu avec un microphone cardioïde pour la même position source-captur. La méthode de reconstruction est utilisée pour évaluer l'aptitude d'un modèle prévisionnel de l'intelligibilité, à s'adapter au cas d'un message vocal diffusé dans un lieu d'écoute sourd, et préalablement enregistré dans un local très réverbérant à l'aide de microphones de directivité différentes.

INTRODUCTION

Les modèles de prévision de l'intelligibilité sont adaptés à l'estimation des scores dans le local où aura lieu la prononciation. Lorsque le message vocal est capté pour être diffusé dans un autre local et écouté en champ proche, la reconnaissance du vocabulaire peut dépendre du procédé de prise de son utilisé [1]. La question qui se pose est de savoir si un modèle prévisionnel [2] validé en situation normale d'écoute pour des locaux très réverbérants, modifié pour tenir compte de la prise de son, rend compte de cette modification des scores lorsque l'écoute a lieu dans un local sourd. L'adaptation du modèle à la situation de prise de son est présentée et les résultats sont confrontés aux scores réels. Préalablement, une procédure de reconstruction d'un échogramme est proposée: il s'agit de combiner les réponses impulsionnelles obtenues avec un microphone omnidirectionnel (cas classique) et un microphone bidirectionnel pour estimer l'échogramme (et surtout le D_{50}) qui aurait pu être obtenu avec un microphone cardioïde.

I. MÉTHODE ET HYPOTHÈSES

Le modèle prévisionnel d'intelligibilité basé sur le concept des énergies précoces utiles et des énergies tardives nuisibles de la réponse impulsionnelle du local sépare les contributions du diffuseur de celles de la salle [2].

Le score prévisionnel $I(\%)$ est calculé à l'aide du modèle comme suit:

$$I(\%) = 100 \left(10^{-\left[\frac{(S/N)_{eq} + 40}{10.8} \right]} \right)^{203} \quad (1)$$

avec

$$(S/N)_{eq} = 10 \cdot \log \left(\frac{D_{50}^s + R_{rf} + R_{dir}}{D_{50}^s} \right) 10^{-(S/N)/10}$$

(S/N) est le rapport signal-sur-bruit en dB(A) et R_{rf} est une grandeur qui mesure la régularité de la réponse en fréquence du haut-parleur. La grandeur R_{dir} désigne le rapport entre la définition D_{50} de

R.Thiele [3], calculée sur l'échogramme global (diffuseur et salle) et la définition modifiée D_{50}^s calculée sur l'échogramme issu de la réponse impulsionnelle globale déconvoluée de celle du haut-parleur dans son axe principal de rayonnement: $R_{dir} = D_{50} / D_{50}^s$. Les autres grandeurs intervenant dans le modèle ne présentent pas de lien avec le type de prise de son.

Les mesures de réponse impulsionnelles sont le plus souvent effectuées à l'aide de microphones de directivité omnidirectionnelle. Si le microphone est directionnel, l'hypothèse est ici que le rapport R_{dir} reste constant. En effet, l'influence de la directivité est équivalente sur la réponse impulsionnelle globale et sur celle déconvoluée puisque cette dernière résulte uniquement de la déconvolution de la réponse impulsionnelle du haut-parleur. La position du couple haut-parleur / microphone reste la même dans la salle quel que soit le microphone considéré. En disposant de la réponse impulsionnelle globale (R.I) obtenue avec un microphone directionnel et en admettant la constance du rapport R_{dir} il est donc possible d'en déduire D_{50}^s et de calculer le score prévisionnel à l'aide du modèle. La déconvolution n'est effectuée que lors de l'expérimentation avec le premier microphone. Les considérations qui suivent reposent également sur l'hypothèse que les directivités microphoniques sont conforme à l'équation polaire qui les définit sur tout ou une partie de la bande de fréquence audio ou au moins sur la bande de fréquence choisie pour la mesure des réponses impulsionnelles.

II. CALCUL DE L'ÉCHOGRAMME

A. Procédure

En l'absence de mesure de R.I avec un microphone de directivité cardioïde, il est possible de reconstituer l'échogramme correspondant à l'aide de deux R.I mesurées au même point avec deux

microphones respectivement de directivité omni- et bidirectionnelle.

Si $V^2(t_i)_{omni}$ et $V^2(t_i)_{bi}$ désignent respectivement l'amplitude de l'onde directe (t_i) sur l'échogramme mesuré avec un microphone omnidirectionnel et bidirectionnel, alors $V^2(t_i)_{omni}$ et $V^2(t_i)_{omni}$ désignent celles de la même réflexion i sur les deux échogrammes.

Le microphone bidirectionnel est positionné pour capter l'énergie sonore latérale: sa figure de directivité est perpendiculaire à l'axe principal de rayonnement du haut-parleur. L'égalité suivante exprime alors la relation entre les deux échogrammes pour chaque réflexion i considérée:

$$\frac{V^2(t_i)_{bi}}{V^2(t_i)_{omni}} = \frac{V^2(t_i)_{omni}}{V^2(t_i)_{omni}} \times \cos^2 \left(\theta_i - \frac{\pi}{2} \right) \quad (2)$$

L'origine des angles est choisie dans l'axe principal de la figure de directivité bidirectionnelle (la source sonore est donc à $\theta = \pi/2$ rad).

L'égalité (2) s'écrit encore comme suit:

$$\cos^2 \theta_i = 1 - \left[\frac{V^2(t_i)_{bi}}{V^2(t_i)_{omni}} \times \frac{V^2(t_i)_{omni}}{V^2(t_i)_{bi}} \right] \quad (3)$$

Pour $i=1$, on obtient $\theta = \pi/2$ rad ce qui correspond bien à l'onde directe.

Le calcul de l'échogramme qui aurait été obtenu avec l'utilisation d'un microphone cardioïde est basé sur l'équation polaire de la cardioïde en remplaçant l'amplitude dans l'axe principal par celle obtenue avec le microphone omnidirectionnel:

$$V(t_i)_{cardio} = V(t_i)_{omni} \times 0,5 \times (1 + \cos \theta_i) \quad (4)$$

où $V(t_i)_{omni}$ désigne l'amplitude de la réflexion présente sur la réponse impulsionnelle obtenue avec un microphone omnidirectionnel à l'instant t_i ; θ_i est l'angle correspondant à cette réflexion. L'équation (4) est modifiée pour tenir compte de l'origine des angles et élevée au carré pour obtenir l'échogramme:

$$V^2(t_i)_{cardio} = V^2(t_i)_{omni} \times 0,25 \times \left[1 + \cos \left(\theta_i - \frac{\pi}{2} \right) \right]^2$$

Après développement, l'équation devient:

$$V^2(t_i)_{cardio} = V^2(t_i)_{omni} \times 0,25 \times \left[2 - \cos^2 \theta_i + 2\sqrt{1 - \cos^2 \theta_i} \right] \quad (5)$$

La méthode de calcul de l'échogramme *cardioïde* consiste à obtenir $\cos^2 \theta_i$ grâce à l'équation (3) et aux deux échogrammes *omnidirectionnel* et *bidirectionnel*, puis à le remplacer dans (5) pour chaque t_i . Les t_i sont les échantillons des réponses impulsionnelles correspondantes.

B. Précisions sur le calcul

La programmation de la procédure de calcul implique un repérage précis de l'onde directe sur les deux échogrammes afin d'identifier $V^2(t_i)_{omni}$ et $V^2(t_i)_{bi}$. Cette opération est justement facilitée par la connaissance de deux réponses impulsionnelles car il est alors possible de comparer l'évolution des premiers échantillons et de repérer, par identité de formes, l'onde directe.

L'équation (3) peut conduire à des valeurs de $\cos^2 \theta$ négatives et très élevées en valeur absolue. En effet, l'amplitude $V^2(t_i)_{bi}$ peut être très faible par rapport à $V^2(t_i)_{omni}$ si pour l'échantillon considéré il n'existe pas de réflexion sonore ou si elle est "noyée" dans le bruit de calcul (lorsque les ordres de grandeurs sont contraires, les valeurs de $\cos^2 \theta$ sont inférieures à 1). Pour remédier à ce problème, la procédure consiste à affecter systématiquement le $\cos^2 \theta$ à la valeur 0 pour que l'amplitude de l'échantillon sur l'échogramme *cardioïde* soit égale à celle sur l'échogramme *omnidirectionnel* (équ.(5)). Ce choix annule l'action possible du microphone *cardioïde*.

III. RESULTATS

A. Mesures des réponses impulsionnelles

Les mesures ont eu lieu dans un local très réverbérant ($T_{60} = 6,2$ s en linéaire 20-20000Hz) le long de l'axe principal de rayonnement du haut-parleur. Quatre distances ont été choisies depuis la zone de champ direct jusqu'à celle de champ diffus: 2, 4, 8, 16 m. A ces distances les R.I ont été mesurées par une technique M.L.S à l'aide d'un microphone omnidirectionnel ½ pouce et d'un microphone bidirectionnel Senheiser MKH30P48 aux mêmes positions.

B. Echogrammes simulés

Les figures 3 et 6 montrent les échogrammes *cardioïdes* obtenus par reconstruction selon la procédure précédente à partir des échogrammes *omni-* (fig.1 et 4) et *bidirectionnel* (fig.2 et 5) mesurés à 4 et 16 m de la source (l'onde directe est confondue avec l'axe vertical). L'axe des abscisses des figures 1 à 6 est en nombre d'échantillons et la fréquence d'échantillonnage est de 8000 Hz.

Un grossissement des figures permettrait de constater que les réflexions prédominantes se retrouvent positionnées comme sur les échogrammes mesurés in situ.

Pour appliquer ces résultats à la prévision des scores d'intelligibilité, les D_{50} sont calculés sur chaque échogramme (tableau 1).

Distance	$D_{50\ omni}$	$D_{50\ bi}$	$D_{50\ cardioïde}$
2m	78,0 %	31,1%	82,4%
4m	54,4 %	32,9%	58,7%
8m	33,6%	18,8%	39,7%
16m	13,2%	3,3%	23,1%

Tableau1: Valeurs de D_{50} en fonction de la distance et de la directivité microphonique.

Les valeurs de D_{50} décroissent lorsque la distance augmente puisque l'énergie au-delà de 50ms croît avec la prédominance du champ réverbéré. La simulation de l'échogramme *cardioïde* satisfait à cette propriété. De plus, à distance égale, la logique est également respectée puisque le D_{50} est d'autant plus grand que le procédé microphonique employé sélectionne une part plus importante de l'énergie

directe: depuis la directivité bidirectionnelle à 90° de l'axe de la source, à l'omnidirectionnalité et à la directivité cardioïde. Enfin, l'effet du facteur de distance microphonique se manifeste également en observant les différences suivantes:

$(D_{50 \text{ cardio}} - D_{50 \text{ omni}}) \simeq 4\%$ à 2m et $(D_{50 \text{ cardio}} - D_{50 \text{ omni}}) \simeq 10\%$ à 16m. L'apport de la directivité est d'autant plus sensible que la distance grandit.

Les résultats obtenus pour le D_{50} montrent un comportement de la simulation tout à fait en accord avec les propriétés des RI dans les locaux [3].

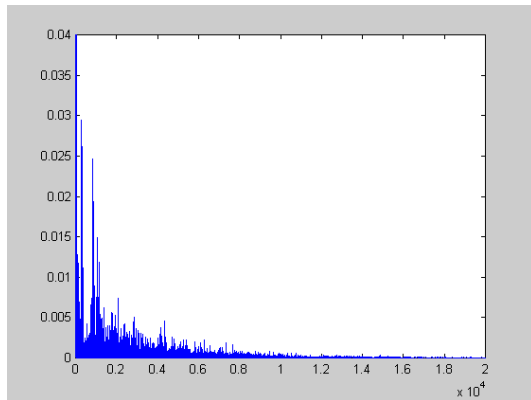


Fig.1: Echogramme mesuré avec le microphone omnidirectionnel à 4m.

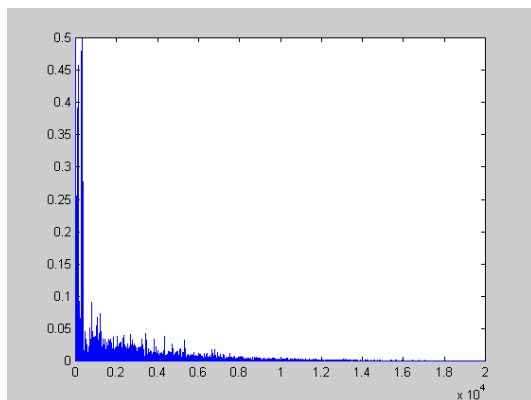


Fig.2: Echogramme mesuré avec le microphone bidirectionnel à 4m.

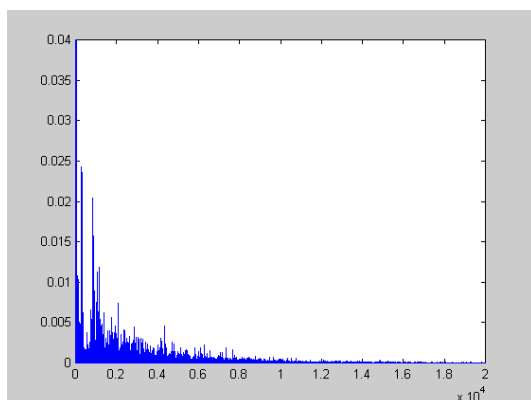


Fig.3: Echogramme simulé pour un microphone cardioïde à 4m.

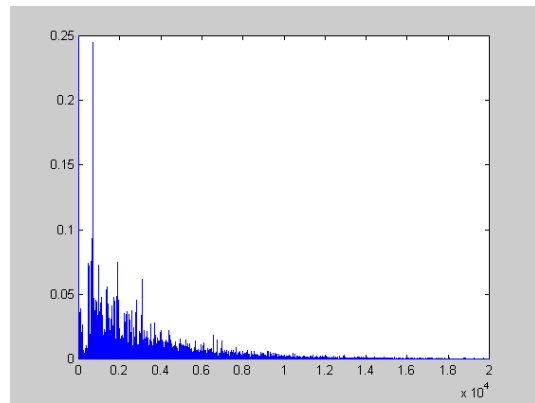


Fig.4: Echogramme mesuré avec le microphone omnidirectionnel à 16m.

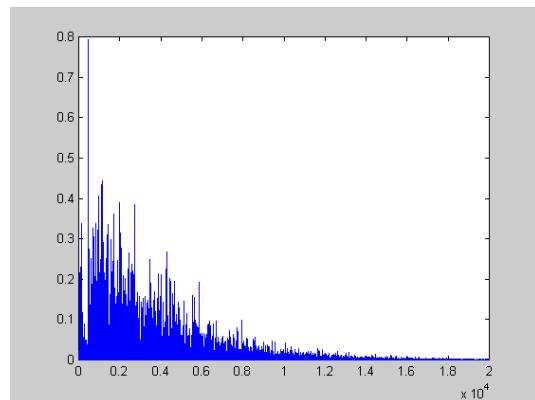


Fig.5: Echogramme mesuré avec le microphone bidirectionnel à 16m.

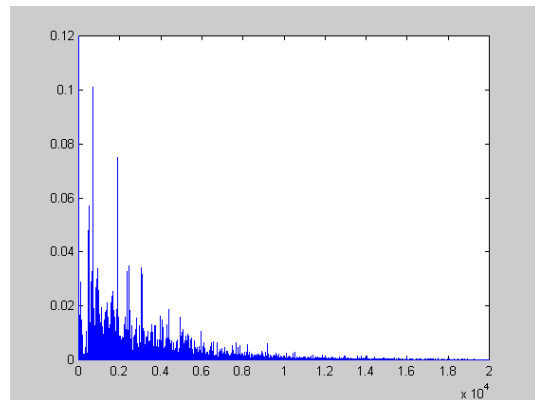


Fig.6: Echogramme simulé pour un microphone cardioïde à 16m.

IV. APPLICATION A L'EVALUATION DU MODELE D'INTELLIGIBILITE

A. Scores prévisionnels

Grâce aux valeurs de D_{50} du tableau 1, il est possible de calculer les scores prévisionnels à l'aide de l'équation (1) du modèle. Pour cela, il est admis que R_{dir} est une constante pour la position du couple source-microphone et le haut-parleur considérés. Les valeurs de R_{dir} sont donc celles calculées pour le microphone omnidirectionnel [2]: 1,8 à 4m; 1,5 à 8m et 1,2 à 16m. Le rapport signal-sur-bruit est toujours supérieur à 30 dB(A). Le tableau 2 et la figure 7 regroupent les résultats pour permettre la

comparaison entre score réels et scores prévisionnels mais aussi entre scores obtenus suivant le type de microphone utilisé.

Distance	Omni réel	Omni modèle	Cardio réel	Cardio modèle	Bi modèle
4m	95,0	89,2	98,7	89,6	87,1
8m	95,0	86,6	94,0	87,4	84,7
16m	78,0	82,5	80,3	84,4	80,6

Tableau 2: Scores d'intelligibilité (en %) mesurés et prédits par le modèle suivant la distance et le procédé de prise de son.

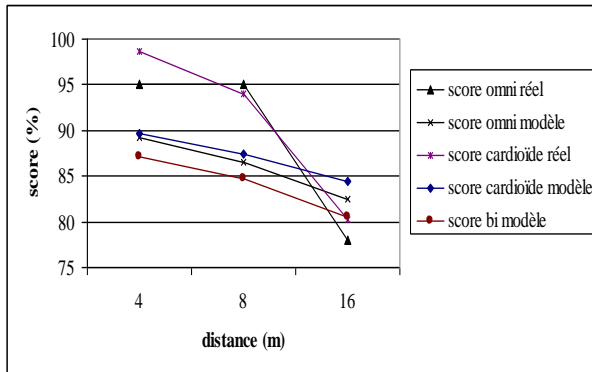


Fig.7: Scores d'intelligibilité du tableau 2.

Les scores réels sont mesurés à partir de listes de mots phonémiquement équilibrées et correspondent au nombre de phonèmes reconnus [2], ici pour un total de 306 phonèmes à reconnaître. Les enregistrements des listes avec les microphones de directivité différentes sont successivement diffusés dans le local sourd par les deux mêmes haut-parleurs que celui qui a servi dans le local réverbérant. Douze auditeurs se sont prêtés aux tests.

B. Evaluation du modèle

Dans une précédente étude [2], le modèle proposé a été validé dans le même local réverbérant et aux mêmes distances que pour cette expérimentation. La différence moyenne en valeur absolue entre les scores prédits et les scores mesurés *in situ* était de 6%, ce qui est également le cas ici pour l'écoute des enregistrements des listes. La décroissance des scores avec la distance est respectée mais le modèle surestime la réalité pour les enregistrements effectués en champ réverbéré et sous-estime ceux réalisés en champ direct. L'estimation de l'intelligibilité à partir des échogrammes *cardioïdes* simulés s'inscrit correctement dans l'évolution des scores, tant du point de vue de la distance que de la directivité, par rapport aux deux autres procédés de prise de son. Les résultats obtenus en utilisant ce modèle (adapté à l'évaluation des scores *in situ*) pour des messages enregistrés et diffusés dans un local beaucoup plus sourd que le lieu d'enregistrement présentent une tendance favorable et les écarts entre prévision et réalité correspondent à moins de 20 phonèmes non reconnus, ou omis, ou confondus sur 306 à

reconnaître dans un contexte exempt de prévisibilité et de redondance de l'information.

V. CONCLUSION

La recombinaison de l'échogramme *cardioïde* à partir des réponses impulsionnelles mesurées avec des microphones omni- et bidirectionnel peut être étendue à d'autres combinaisons de directivité. Un échogramme *bidirectionnel* peut être reconstruit par exemple, à partir d'une RI *omnidirectionnelle* et d'une RI *cardioïde*.

Bien qu'une mesure avec un microphone cardioïde reste à effectuer pour vérifier la procédure, tout comme des mesures pour s'assurer de la constance de la grandeur R_{dir} , les valeurs de D_{50} et le schéma des réflexions sonores obtenus témoignent du comportement de la technique mise en œuvre. Le nombre d'échantillons pour lesquels il est nécessaire d'affecter à l'échogramme *cardioïde*, l'amplitude de l'échogramme *omnidirectionnel* (cf II.B.) est de 901 pour la distance 16m, 4105 pour 8m, 9052 pour 4m et 7139 pour 2m. L'amélioration de la technique de calcul passe donc par la diminution de ces quantités, ce qui entraînera probablement une augmentation du D_{50} et par voie de conséquence celle des scores d'intelligibilité à 4m et 8m, distances auxquelles le modèle est le moins précis.

Les applications peuvent concerner le développement des algorithmes de prévision de l'acoustique interne en vue du calcul des critères, non seulement en monophonie mais surtout en stéréophonie, et en particulier pour le procédé de prise de son M-S.

L'application à l'évaluation du modèle prévisionnel de l'intelligibilité a confirmé le bon comportement des simulations d'échogrammes car très étroitement lié à celui des valeurs de D_{50} .

Le modèle doit être adapté pour être "transporté" au cas de la diffusion d'enregistrements effectués avec des directivités microphoniques variables, voire pour les techniques stéréophoniques. Mais les résultats obtenus permettent d'envisager que soit conservé le corps général du modèle et que l'adaptation au problème posé soit recherchée sur la base de ce modèle.

RÉFÉRENCES

- [1] R.Ruiz, I.Ballet, "Influence of Distance and Sound Recording System on Intelligibility in Highly Reverberant Conditions," 16th I.C.A and 135th Meeting of the A.S.A, Seattle, U.S.A, vol 103 (5), 3067(A) (1998).
- [2] L.Faiget, R.Ruiz, "Speech Intelligibility Model Including Room and Loudspeaker Influences", J. Acoust. Soc. Am., vol 105 (5), pp. 3345-3354 (1999).
- [3] L. Cremer and H. Muller, "Principles and Applications of Room Acoustics", Applied Science Publishers Ltd., London (1982).