

# Evaluation d'estimateurs de sonie

**Jeremy Marozeau**

**Isabelle Boulet**

**Sabine Meunier**

LMA-CNRS

31, chemin Joseph Aiguier

13402 Marseille CEDEX 20

E-mail : marozeau@lma.cnrs-mrs.fr

E-mail : boulet@lma.cnrs-mrs.fr

E-mail : meunier@lma.cnrs-mrs.fr

**Fabrice Gagneux**

**Patrick Boussard**

Genesis S.A.

Bâtiment Beltram

Domaine du Petit Arbois

BP 69

13545 Aix-en-Provence CEDEX 4

E-mail : fabrice.gagneux@genesis.fr

E-mail : patrick.boussard@genesis.fr

*La sensation de force sonore d'un son, la sonie, ne peut être mesurée que par des tests psychoacoustiques impliquant plusieurs auditeurs et nécessitant des moyens importants. Afin d'éviter ces expériences systématiques et longues, des modèles perceptifs ont été proposés. Leur but est de permettre la conception d'algorithmes pour estimer la sonie à partir du signal physique. Cependant, ces différents modèles ne donnent des estimations fiables qu'avec des signaux stationnaires. Le projet de recherche que nous avons entrepris dans le cadre de l'Appel à Proposition de Recherche « Bruit et Nuisances Sonores » a pour but l'élaboration d'un estimateur « global » de sonie pouvant s'appliquer aussi bien aux signaux stationnaires que non stationnaires. Le présent article décrit deux aspects du projet : 1) l'élaboration d'une méthode objective permettant de juger la qualité d'un estimateur de sonie, 2) la comparaison de certains outils d'estimation déjà existants. Nous illustrerons à cette occasion les limitations de la mesure du niveau perceptif à l'aide d'échelles de décibels pondérés (A, B ou C).*

*Loudness, the auditory sensation relative to sound intensity, can be measured only by psychoacoustic tests, implying several listeners and requiring significant investment. To avoid long and systematic experiments, perceptive models have been proposed, the goal being to design algorithms that estimate loudness from the physical signal. However, these various models tend to give reliable estimates only when applied to stationary signals. The aim of the research project presented here is the development of a «global» estimator of loudness, which can be applied to both stationary and non-stationary signals. This article describes two aspects of the project: 1) the development of an objective method to judge the quality of a loudness estimator, 2) the comparison of some existing estimation tools. The limitations of the perceptive level measurement using scales of weighted decibels (A, B or C) will be illustrated in the comparison.*

## Introduction

### La sonie, une caractéristique perceptive

La sonie est l'équivalent sensoriel de l'intensité physique d'un son. Elle dépend non seulement du niveau acoustique, mais aussi du contenu fréquentiel et de la durée du signal. La sonie est une grandeur subjective et s'exprime en sonies. Par convention, la sonie d'un son pur de 1 kHz à 40 dB SPL correspond à une valeur de 1 sonie. Un son de 2 sonies est perçu deux fois plus fort qu'un son de 1 sonie. La sonie d'un son test peut être aussi mesurée par comparaison avec un autre son (un son pur à 1 kHz ou une bande de bruit centrée sur 1 kHz de largeur inférieure à 160 Hz, correspondant à la largeur d'une bande critique). Le niveau d'isonie est alors obtenu en phones et correspond au niveau en dB SPL que doit avoir

le son de comparaison pour que les deux sons (son test et son de comparaison) aient la même sonie.

Une relation déterminée à partir de la fonction de sonie d'un son pur de 1 kHz, permet de passer de l'échelle des sonies à l'échelle des phones [1]. Une augmentation de 10 phones entraîne ainsi un doublement de la sonie. Par exemple, si la sonie d'un son pur de 1 kHz à 40 dB (ou 40 phones) est de 1 sonie, elle est de 2 sonies à 50 dB (ou 50 phones), de 4 sonies à 60 dB (ou 60 phones) etc...

En tant que grandeur subjective, le seul moyen de mesurer la sonie est de réaliser des tests psychoacoustiques sur un grand nombre d'auditeurs. Toutefois, ces tests sont contraignants car ils nécessitent des ressources humaines et matérielles importantes. Des outils de calculs ont alors été élaborés afin d'estimer la sonie, ou le niveau d'isonie, à partir de l'analyse du signal physique.

La façon la plus commune d'estimer le niveau sonore perçu d'un son est de calculer son niveau en décibel (dB) pondéré (A, B ou C). Des modèles perceptifs [2, 3, 4, 5, 6] ont été mis au point afin de prendre en considération les mécanismes fins de l'appareil auditif (intégration temporelle, masquage fréquentiel et masquage temporel, sommation de sonie...).

### La sonie des sons non stationnaires

Pour un son stationnaire, n'importe quelle durée du segment de signal peut être utilisée pour effectuer le calcul de la sonie. En revanche, pour un son non stationnaire, l'intégralité du signal est nécessaire à son calcul. Certains modèles [3, 6] adoptent une démarche permettant le calcul de la sonie en fonction du temps (sonie instantanée). Pour calculer la sonie globale d'un son, leurs auteurs préconisent alors d'utiliser différents attributs de la sonie instantanée, comme le maximum ou le niveau RMS [6], le N5 ou le N10 représentant respectivement la sonie dépassée 5 % ou 10 % du temps [7].

### Présentation générale du projet

Le projet «Estimateur de Sonie de Sons Non Stationnaires» dans lequel s'inscrit cette étude est financé par le ministère de l'Écologie et du développement durable (M.E.D.D.). Les quatre étapes clés du projet sont :

#### La comparaison des méthodes de mesure

Il est nécessaire de mener des expériences psycho-acoustiques afin de mesurer la sonie d'un son. Plusieurs méthodes psychophysiques sont possibles. Nous avons débuté le projet par une étude destinée à déterminer la mieux adaptée aux sons non stationnaires [8]. Nous nous sommes intéressés à quatre méthodes de mesure de la sonie fréquemment utilisées : la méthode d'estimation directe [9], la méthode d'ajustement [10], une méthode adaptative à choix forcé sur deux intervalles [11] et une méthode adaptative à poursuites multiples [12].

Il s'est avéré que la méthode d'ajustement présente le meilleur compromis entre la durée de l'expérience et la variabilité des résultats obtenus. Nous l'avons donc adoptée pour le reste de nos travaux.

La mesure de la sonie comporte irrémédiablement une certaine dispersion due aux différences inter-sujets. Cela se manifeste dans nos expériences par un écart type de 4 phones. Un estimateur sera donc acceptable, s'il calcule la sonie à 4 phones près par rapport à la sonie moyenne mesurée.

#### Mesures de la sonie de sons impulsifs

Plusieurs expériences de mesure de la sonie utilisant la méthode d'ajustement ont été menées avec des stimuli différents. Les expériences ont permis, entre autre, de mesurer la sonie de sons complexes, d'étudier l'intégration temporelle [13] et d'obtenir une mesure des lignes isosoniques d'un corpus de sons impulsifs [14].

### L'évaluation des estimateurs

Dans la perspective de l'élaboration d'un estimateur global de la sonie, nous devons être en mesure d'en juger la qualité de façon précise. Nous avons pour cela proposé quatre indices objectifs permettant d'évaluer la précision des calculs de la sonie. Ces indices sont basés sur la comparaison entre la sonie mesurée par des expériences psychoacoustiques et la sonie calculée par les algorithmes. Au moyen de ces indices, une étude comparative des estimateurs de sonie couramment utilisés a pu être menée. Cette partie du projet est développée dans cet article.

### L'élaboration de l'estimateur optimal

La troisième étape du projet nous permet d'identifier les qualités et les limites des estimateurs existants. A partir de ces informations, nous allons travailler à l'élaboration d'un nouvel estimateur de sonie adapté aux sons stationnaires et non stationnaires.

### Les estimateurs de sonie

#### Les estimateurs de type signal : les décibels pondérés (A, B ou C)

La méthode la plus usitée pour estimer en première approche le niveau sonore perçu d'un son consiste à calculer son niveau pondéré (proportionnel au logarithme de la puissance moyenne du spectre pondéré). La pondération tient compte de la variation de la sensibilité de l'oreille en fonction de la fréquence et prend des valeurs adaptées au niveau d'isotonie du signal (40, 70 et 100 phones pour les pondérations A, B et C respectivement).

Afin de calculer la puissance moyenne du signal dans le temps, il est possible de prendre le niveau RMS (Root Mean Square) du signal entier. Toutefois, il est également intéressant de pouvoir calculer la variation du niveau en fonction du temps. Pour cela, une technique possible consiste à découper le signal à l'aide de fenêtres temporelles, et de calculer le niveau à l'intérieur de chacune d'elles. La précision temporelle dépend directement de la taille des fenêtres utilisées. Nous en avons testé trois de 1s, 125 ms et 35 ms. Ces valeurs ont été choisies par analogie avec les constantes de temps proposées dans l'utilisation des sonomètres et qui sont associées aux labels « slow », « fast » et « impulse ». Si le signal traité est long, il est analysé par plusieurs fenêtres. Afin d'obtenir une valeur unique du niveau d'isotonie, nous avons pris le parti de retenir le niveau maximum. Nous avons donc implémenté douze algorithmes (trois pondérations sur le signal entier, plus trois pondérations fois trois tailles de fenêtre) qui calculent les niveaux pondérés d'un corpus de stimuli, premières estimations de la sonie.

Ces estimateurs ne prennent absolument pas en considération les mécanismes fins de l'appareil auditif, comme le masquage fréquentiel, le masquage temporel, la sommation fréquentielle de sonie et l'intégration temporelle.

## Les estimateurs de type modèles perceptifs

Les modèles perceptifs ont été construits dans le but d'imiter le fonctionnement du système auditif. Les deux modèles que nous avons étudiés adoptent une démarche commune : un filtre est appliqué au signal afin de tenir compte de la réponse en fréquence de l'oreille, le signal est passé au travers d'un banc de filtres afin de reproduire la sélectivité fréquentielle de l'oreille (principalement interne), la puissance moyenne à la sortie de chaque filtre est transformée en sonie spécifique. La sonie spécifique est intégrée sur tout l'axe fréquentiel pour obtenir la sonie globale en sones. La sonie peut être convertie en phones selon la formule donnée par la norme ISO 532 B [1].

### Le modèle ISO 532 B

Le modèle de sonie de Zwicker [2] a fait l'objet d'une norme [1]. Nous l'avons implémenté afin de le comparer à d'autres outils de calcul de la sonie. Toutefois, il faut être conscient que ce modèle a été conçu pour estimer la sonie de sons stationnaires. En effet, le calcul de la sonie spécifique est opéré sur la totalité du signal. Puisque notre panel d'expériences comporte des sons non stationnaires, il est important de souligner que le modèle ISO 532 B n'est pas adapté à ce type de signaux. Comme il est souvent utilisé, nous l'avons inclus dans notre étude afin d'estimer l'erreur que ce modèle introduit sur ce type de son.

### Le modèle de Zwicker de sons non stationnaires

Zwicker a adapté le modèle ISO 532 B pour pouvoir calculer la sonie en fonction du temps en tenant compte du masquage temporel [3,7]. Nous avons implémenté ce modèle et fait le choix d'estimer la sonie globale par la valeur maximale de la courbe de sonie instantanée comme le préconise Zwicker [3].

### Les autres modèles

D'autres modèles ont été proposés [15,16,...]. En particulier Moore et coll. [4,5] ont révisé le modèle de calcul de la sonie de Zwicker [2]. Ils se sont basés sur des données physiologiques plus récentes et calculent ainsi différemment la sonie spécifique notamment en ce qui concerne le banc de filtres utilisé. Une amélioration de ce modèle [6] a été proposée afin de l'adapter aux sons non stationnaires. Ce dernier, étant le plus moderne, semble particulièrement prometteur. Toutefois, étant à ce jour en cours d'implémentation, nous n'avons pas pu le tester dans cet article.

Les implémentations des modèles ISO 532 B et Zwicker [3] ont tout d'abord été validées avec les données des articles de l'auteur des modèles.

Les modèles testés sont récapitulés dans le tableau 1.

## Evaluation des estimateurs

### Les différentes expériences

Afin de pouvoir évaluer les estimateurs, il est important de les tester sur différents types de sons. Nous avons pu recueillir de nombreuses données au travers des mesures de sonie effectuées au cours de la deuxième étape du projet. Les stimuli étaient soit synthétiques, soit naturels. Les stimuli synthétiques ont été créés à partir de deux types de forme d'enveloppe temporelle. La première était en forme de créneau avec une attaque courte (demi fenêtre de hanning croissante) suivie d'un plateau d'amplitude constante et de durée variable, et d'une décroissance courte (demi fenêtre de hanning décroissante). La deuxième était en forme d'impulsion

Nom	Descriptions
dBA rms	Calcule le niveau RMS du signal en dB avec la pondération A
dBB rms	Calcule le niveau RMS du signal en dB avec la pondération B
dBC rms	Calcule le niveau RMS du signal en dB avec la pondération C
Max dBA slow	Calcule le niveau en dB(A) sur plusieurs fenêtres consécutives de 1 s du signal. Le niveau maximum est retenu.
Max dBB slow	Idem que dB(A) slow mais avec une pondération B
Max dBC slow	Idem que dB(A) slow mais avec une pondération C
Max dBA fast	Calcule le niveau en dB A sur plusieurs fenêtres consécutives de 125 ms du signal. Le niveau maximum est retenu.
Max dBB fast	Idem que dB(A) fast mais avec une pondération B
Max dBC fast	Idem que dB(A) fast mais avec une pondération C
Max dBA imp.	Calcule le niveau en dB(A) sur plusieurs fenêtres consécutives de 35 ms du signal. Le niveau maximum est retenu.
Max dBB imp.	Idem que dB(A) imp. mais avec une pondération B
Max dBC imp.	Idem que dB(A) imp. mais avec une pondération C
ISO 532 B	Modèle normalisé de calcul de la sonie, selon une méthode décrite par Zwicker [2]. Ce modèle se base sur l'analyse du signal entier.
Zwicker 1977	Modèle amélioré de ISO 532B où l'évolution temporelle du signal est prise en considération [3].

Tabl. 1 : Récapitulatifs des modèles testés

avec une attaque courte et linéaire suivie directement d'une décroissance exponentielle de durée variable. Ces enveloppes temporelles ont servi à fenêtrer soit des sons purs, soit des bruits blancs. A l'intérieur de chaque expérience, le même type de stimulus était utilisé (sons naturels ou synthétiques, créneaux ou impulsions, sons purs ou bruits blancs). Un ou plusieurs paramètres variaient. Par exemple, dans la première expérience, nous avons utilisé des sons synthétiques avec une enveloppe temporelle de type créneau sur un son pur. Nous n'avons fait varier qu'un seul paramètre, la durée du plateau des stimuli. Ainsi, nous avons pu étudier l'effet de la durée sur la sonie [13]. L'ensemble de ces expériences nous a permis d'obtenir neuf corpus de sons différents sur lesquels nous avons pu évaluer les estimateurs de sonie. Le tableau 2 résume les types de stimuli et les variables utilisées dans chaque expérience.

un MAR comparable, celui avec le Rmax le plus faible sera donc préféré.

**Le Rmean (Résiduel moyen)**

Le troisième indice estime l'erreur moyenne, sans la valeur absolue.

$$R_{mean} = \frac{\sum (Scalculée - Smesurée)}{N}$$

Cet indice permet de savoir si les erreurs commises par l'estimateur se répartissent uniformément autour d'une valeur moyenne. Par exemple, si la valeur de Rmean est petite, proche de zéro, alors la répartition des erreurs est uniforme autour de la valeur de la mesure. Par contre, si la valeur du Rmean est grande, positive ou négative, alors la répartition des erreurs est le plus souvent au-

Numéro	Type de sons	Paramètre(s) variable(s)	Référence
1	Synthétiques : créneau, son pur	La durée (de 10 ms à 2 sec)	[13]
2	Synthétiques : impulsion, son pur	La durée (de 10 ms à 2 sec)	[13]
3	Synthétiques : créneau, son pur	La durée (de 10 à 500 ms) et la fréquence (de 125 Hz à 8 kHz)	[13]
4	Synthétiques : impulsion, son pur	La durée (de 10 à 500 ms) et la fréquence (de 125 Hz à 8 kHz)	[13]
5	Synthétiques : impulsion, son pur	La niveau (de 50 à 75 dB SPL) et la fréquence (de 125 Hz à 10 kHz)	[14]
6	Synthétiques : créneau, bruit blanc	La durée (de 10 ms à 2 sec)	[14]
7	Synthétiques : impulsion, bruit blanc	La durée (de 10 ms à 2 sec)	[14]
8	Naturels : type stationnaire	aucun	[14]
9	Naturels : type non stationnaire	aucun	[19]

Tabl. 2 : Récapitulatif des types des stimuli et variables utilisés dans chaque expérience.

**Les indices d'évaluation des estimateurs**

Quatre indices ont été proposés afin de pouvoir évaluer l'efficacité d'un estimateur.

**Le MAR (Mean Average Residual) [17]**

Cet indice calcule simplement la moyenne de la valeur absolue de la différence entre le niveau d'isotonie calculée par l'estimateur, *Scalculée*, et le niveau d'isotonie mesurée par une expérience psychoacoustique sur des sujets, *Smesurée* :

$$MAR = \frac{\sum |Scalculée - Smesurée|}{N}$$

où N représente le nombre de stimuli testés. Cet indice permet de connaître rapidement si l'estimateur donne de bons résultats moyens. Il est le plus intuitif.

**Le Rmax (Résiduel maximum)**

Ce second indice calcule la valeur absolue de l'erreur maximale d'estimation du niveau d'isotonie :

$$R_{max} = \max(|Scalculée - Smesurée|)$$

Il teste donc la robustesse de l'estimateur en délimitant sa marge d'erreur maximale. Si deux estimateurs possèdent

dessus ( $R_{mean} > 0$ ) ou au-dessous ( $R_{mean} < 0$ ) de la valeur de la mesure. En effet, si le modèle avait un biais systématique le MAR et le Rmean auraient la même valeur au signe près.

**Le Rstd (l'écart type du résiduel)**

Le quatrième indice estime l'écart type de l'erreur.

$$R_{std} = \sqrt{\frac{\sum [(Scalculée - Smesurée) - R_{mean}]^2}{N - 1}}$$

Cet indice permet de connaître si l'estimateur comporte un biais systématique. Si le Rstd est nul, il y a effectivement un biais constant. Une fois le biais annulé, l'estimateur doit donc donner des résultats parfaits.

**Les résultats**

**Les tableaux d'indices**

Les valeurs des indices d'évaluation des estimateurs (MAR, Rmax, Rmean et Rstd), pour chaque expérience, sont résumées dans les tableaux 3, 4, 5 et 6. Dans le cas des estimateurs basés sur le décibel, nous avons pris comme estimation du niveau d'isotonie, *Scalculée*, le niveau de pression pondéré.

MAR	Expériences									
Estimateurs :	1	2	3	4	5	6	7	8	9	Moyenne
dBA rms	2	3	14	16	14	25	26	14	20	15
dBB rms	2	3	12	14	13	26	27	13	17	14
dBC rms	2	3	11	13	13	26	27	13	13	13
Max dBA slow	6	6	13	14	13	24	24	14	18	15
Max dBB slow	6	6	11	12	13	25	25	13	15	14
Max dBC slow	6	6	10	11	12	25	25	13	12	13
Max dBA fast	1	2	6	7	6	16	16	13	14	9
Max dBB fast	1	2	3	4	5	17	17	12	11	8
Max dBC fast	1	2	3	3	5	17	17	12	8	8
Max dBA imp.	2	3	4	5	4	13	12	12	11	7
Max dBB imp.	2	3	2	2	4	14	13	12	8	7
Max dBC imp.	2	3	3	2	4	14	13	11	6	6
ISO 532 B	2	4	10	11	9	6	8	3	8	7
Zwicker 1977	1	2	2	2	3	2	2	2	6	2

Tabl. 3 : Récapitulatif des MAR pour chaque estimateur et chaque expérience. Les valeurs en-dessous de 4 sont mises en gras (correspondant à la variance des mesures de sonie)

Rmax	Expériences									
Estimateurs :	1	2	3	4	5	6	7	8	9	Maximum
dBA rms	4	5	24	29	26	28	27	23	30	30
dBB rms	4	5	18	20	21	29	29	18	28	29
dBC rms	4	5	18	18	21	29	28	17	26	29
Max dBA slow	11	12	24	28	26	27	25	23	31	31
Max dBB slow	11	12	17	19	21	29	27	18	26	29
Max dBC slow	11	12	17	17	21	28	27	17	24	28
Max dBA fast	2	3	15	19	17	18	16	22	24	24
Max dBB fast	2	3	8	10	12	20	18	17	20	20
Max dBC fast	2	3	8	8	12	20	18	16	20	20
Max dBA imp.	5	5	11	13	11	15	13	20	23	23
Max dBB imp.	5	5	6	5	9	17	15	17	17	17
Max dBC imp.	5	5	7	8	10	17	15	16	15	17
ISO 532 B	6	9	14	18	16	10	11	8	26	26
Zwicker 1977	3	5	5	4	7	4	4	5	13	13

Tabl. 4 : Récapitulatif des Rmax pour chaque estimateur et chaque expérience. Les valeurs en-dessous de 10 sont mises en gras (limite à laquelle la sonie est doublée)

Rmean	Expériences									
Estimateurs :	1	2	3	4	5	6	7	8	9	Moyenne
dBA rms	1	-2	-14	-16	-14	-25	-26	-14	-20	1
dBB rms	1	-2	-12	-14	-13	-26	-27	-13	-17	1
dBC rms	1	-2	-11	-13	-13	-26	-27	-13	-13	1
Max dBA slow	-6	-6	-13	-14	-13	-24	-24	-14	-18	-6
Max dBB slow	-6	-6	-11	-12	-13	-25	-25	-13	-15	-6
Max dBC slow	-6	-6	-10	-11	-12	-25	-25	-13	-11	-6
Max dBA fast	0	1	-6	-6	-5	-16	-16	-13	-14	1
Max dBB fast	0	1	-3	-4	-5	-17	-17	-12	-11	1
Max dBC fast	0	1	-2	-3	-4	-17	-17	-12	-7	1
Max dBA imp.	2	3	-2	-2	0	-13	-12	-12	-11	3
Max dBB imp.	2	3	0	0	0	-14	-13	-12	-8	3
Max dBC imp.	2	3	1	1	0	-14	-13	-11	-3	3
ISO 532 B	1	0	-10	-11	-9	-6	-8	2	-3	2
Zwicker 1977	1	2	0	0	-1	1	2	1	4	4

Tabl. 5 : Récapitulatif des Rmean pour chaque estimateur et chaque expérience. Les valeurs en-dessous de 4 sont mises en gras

Rstd	Expériences									Moyenne
	1	2	3	4	5	6	7	8	9	
Estimateurs :										
dBA rms	2	3	5	6	6	2	1	4	6	6
dBB rms	2	3	3	2	4	2	1	4	5	5
dBC rms	2	3	3	3	4	2	1	4	7	7
Max dBA slow	4	4	6	6	6	3	2	4	6	6
Max dBB slow	4	4	3	3	4	3	2	4	6	6
Max dBC slow	4	4	3	3	4	3	2	4	7	7
Max dBA fast	1	2	5	6	5	2	1	4	6	6
Max dBB fast	1	2	2	3	4	2	1	4	5	5
Max dBC fast	1	2	3	3	4	2	1	4	6	6
Max dBA imp.	2	1	5	6	6	2	1	4	6	6
Max dBB imp.	2	1	3	2	4	1	1	4	5	5
Max dBC imp.	2	1	3	3	4	1	1	4	7	7
ISO 532 B	3	5	2	2	3	3	2	4	10	10
Zwicker 1977	1	2	3	2	3	2	2	2	7	7

Tabl. 6 : Récapitulatif des Rstd pour chaque estimateur et chaque expérience

### La sonie de sons purs, les expériences 1 à 5

Les cinq premières expériences nous permettent d'évaluer les estimateurs sur des sons purs.

#### Expériences 1 et 2,

##### variation de la durée et de l'enveloppe temporelle

Les signaux de la première expérience étaient des sons purs de 1 kHz avec un niveau crête de 70 dB SPL et ayant une enveloppe temporelle de type créneau de durée variable de 10 ms à 2 s (voir [13] pour plus de détails). Tous les estimateurs, hormis les dB pondérés en modes slow, donnent de bons résultats. Leur MAR est inférieur à 2 phones. Pour ces mêmes estimateurs, le Rmax est également faible (entre 2 et 7 phones). Ces estimateurs ne montrent pas de biais systématique avec des Rmean faibles et un Rstd du même ordre de grandeur que les Rmean.

Pour estimer le niveau perçu à partir du niveau en dB, il est important que le signal soit stationnaire à l'intérieur de la fenêtre d'analyse. Par exemple, si la durée d'un signal est inférieure à celle de la fenêtre d'analyse, le niveau estimé sera inférieur au niveau d'isotonie. C'est le cas lorsqu'on utilise la fenêtre d'analyse « slow » (Rmean négatif), car dans ces expériences, les durées des signaux sont pour la plupart inférieures à 1 seconde. Il est intéressant de noter que la meilleure estimation se fait avec une fenêtre temporelle de 125 ms, valeur proche de la constante d'intégration de l'oreille [18].

Les pondérations A, B, et C n'introduisent pas de différences dans les résultats car les stimuli ne possèdent de l'énergie qu'à 1 kHz, fréquence à laquelle les pondérations n'impliquent aucune atténuation.

Les modèles perceptifs donnent également de bons résultats (MAR de l'ordre de 2 phones). A noter un Rmax plus important tout de même pour le modèle ISO 532 B (6 phones). Cette erreur est commise avec le son le plus court.

L'expérience 2 est semblable à l'expérience 1 mais avec des sons ayant une enveloppe temporelle en forme d'impulsion. Les résultats sont similaires à ceux de l'expérience 1. On note toutefois une légère augmentation du MAR, de l'ordre de 1 phone.

#### Expériences 3, 4 et 5, variation de la durée, de la fréquence et de l'enveloppe temporelle

Ces expériences utilisaient des sons purs brefs dont les durées étaient comprises entre 10 et 500 ms, dont les enveloppes temporelles étaient des créneaux ou des impulsions. Leurs fréquences variaient entre 150 Hz et 10 kHz. La proportion de sons brefs (inférieurs à 100 ms) était plus importante que pour les expériences 1 et 2, augmentant ainsi l'importance d'une analyse du son avec une fenêtre temporelle courte.

Cela se manifeste par une baisse importante de la précision de l'estimation des valeurs en « dB (A, B et C) RMS » et « dB (A, B et C) slow ». De plus, les MAR calculés sont plus faibles pour les pondérations B et C. Ce sont les estimateurs « dB B (imp et fast) » et « dB C (imp et fast) » qui prédisent le mieux les niveaux perçus. Les sons étant restitués avec un niveau moyen de 70 dB SPL, les pondérations B et C étaient donc bien plus adaptées que la pondération A, prévue pour des niveaux d'isotonie de 40 phones.

Le modèle perceptif ISO 532 B n'est effectivement pas adapté au calcul de la sonie de sons brefs (MAR environ égal à 10 phones). Sa version modifiée, Zwicker 1977 [3], donne de meilleurs résultats et estime à moins de 3 phones près la sonie sur ce type de son. De plus, les Rmax de ce modèle sont tout à fait acceptables, car ils restent en-dessous de 7 phones.

On observe également un Rmean faible indiquant l'absence de sur ou sous-estimation systématique confortant ainsi notre choix d'estimer la sonie globale par la valeur maximale de la courbe de sonie en fonction du temps.

## La sonie des signaux complexes, les expériences 6 à 9

### Expériences 6-7, les bruits blancs

Les expériences 6 et 7 ont été effectuées en utilisant comme stimuli des bruits blancs dont les durées étaient comprises entre 10 et 500 ms. L'enveloppe temporelle était de type créneau dans l'expérience 6, et de type impulsion dans l'expérience 7. Les MAR montrent clairement que le niveau en dB, avec une erreur de plus de 12 phones, n'est pas une bonne estimation de la sonie quelles que soit la pondération ou la fenêtre temporelle. Le niveau en dB sous-estime systématiquement le niveau perçu (Rmean négatif et Rstd très faible). Cela s'explique par le fonctionnement du système auditif. En effet, à puissances équivalentes un son pur et un bruit à bande étroite centré sur la fréquence du son pur n'auront la même sonie que si la largeur de bande du bruit est inférieure à une bande critique. Si la largeur de bande du bruit dépasse une bande critique alors le bruit aura, à puissance équivalente, une sonie plus grande. Cette différence peut atteindre plus de 10 phones [19], expliquant ainsi l'erreur systématique (Rstd faible) de l'estimation d'isozonie d'un bruit blanc par son niveau en dB.

Les modèles perceptifs conservent des valeurs de MAR faibles car ils tiennent compte de cet effet. Ces expériences montrent donc, que pour des signaux complexes, les modèles perceptifs sont plus adaptés à l'estimation de la sonie.

### Expériences 8-9, les sons naturels

Dans l'expérience 8 [20], nous avons utilisé des signaux naturels de type stationnaire (perceuse, moto au ralenti, note de flûte...). Comme pour les bruits blancs (expériences 6-7), les niveaux en dB (A B ou C) sous-estiment le niveau

d'isozonie (d'environ 12 phones). Les modèles perceptifs (ISO 532 B et Zwicker 1977) gardent des valeurs faibles de MAR et Rmax. Ce résultat valide clairement ces modèles comme des estimateurs acceptables de la sonie de sons stationnaires environnementaux.

Dans l'expérience 9, nous avons utilisé comme stimuli des enregistrements insitu de scènes sonores d'un chantier (étude réalisée pour le ministère de L'environnement et financée par l'ADEME, voir [14]). Ces stimuli étaient relativement longs (environ cinq secondes) et non stationnaires. Aucun estimateur ne montre un MAR acceptable (en-dessous de 4 phones). Les valeurs de Rmax sont également élevées, avec 13 phones comme valeur la plus faible avec Zwicker 1977. Cette expérience montre donc les limites de ce modèle qui a tendance à surestimer la sonie (Rmean de 4 phones) de façon non constante (Rstd de 7 phones). D'autres paramètres de la sonie en fonction du temps (N5, N10, etc.) ont été testés et ne se sont pas montrés plus performants.

## Conclusions

Le travail présenté ici s'inscrit dans un projet de recherche destiné à élaborer un estimateur de sonie «global» pour les sons stationnaires et non stationnaires. Il consistait à évaluer certains estimateurs de sonie. Une attention particulière a été portée aux méthodes d'évaluation des estimateurs. Quatre indices d'évaluation ont été proposés : le MAR (Mean Average Residual), le Rmax (résiduel maximum), le Rmean (résiduel moyen) et le Rstd (écart type du résiduel). Des estimateurs de type signal et des estimateurs de type modèles perceptifs ont été testés.



Dans l'expérience 9, nous avons utilisé comme stimuli des enregistrements insitu de scènes sonores d'un chantier

Les décibels pondérés sont de bons estimateurs de sonie pour les sons purs, que leur enveloppe temporelle soit de type créneau ou impulsion. La durée de la fenêtre d'analyse est primordiale. Une durée proche de la constante d'intégration de l'oreille permet d'approcher au mieux le niveau perçu. Sur les sons complexes, le niveau en décibel sous-estime le niveau d'isotonie de plus de 10 phons. En effet lorsque le spectre d'un signal s'étend sur plus d'une bande critique, il y a sommation de la sonie, phénomène que la mesure en décibel ne prend pas en considération.

Deux modèles perceptifs ont été testés. Le premier, le modèle ISO 532 B estime la sonie des sons stationnaires de manière acceptable car l'erreur moyenne est inférieure à l'écart type des mesures de sonie. Toutefois, pour les sons non stationnaires (pour lesquelles il n'a pas été conçu), il sous-estime le niveau d'isotonie avec une erreur moyenne d'environ 7 phons. Le second modèle, également conçu par Zwicker en 1977 [3], est adapté aux sons non stationnaires. Il donne des estimations du niveau d'isotonie proches des mesures pour les sons stationnaires (synthétiques et naturels), et pour les sons non stationnaires synthétiques (erreur moyenne inférieure à 2 phons). Par contre, il ne donne pas de bons résultats sur un corpus de sons de l'environnement non stationnaires. Il nous faut maintenant comprendre d'où provient cette erreur par une analyse poussée sur des sons non stationnaires. Nous travaillons ainsi à l'implémentation du modèle de Glasberg et Moore [6] afin de le tester et le comparer à celui de Zwicker [3]. Nous cherchons également si un modèle simple du signal pourrait convenir.

and Exposition on noise control engineering (Internoise), Prague, République Tchèque, (2004)

[14] Boulet I., "La sonie des sons impulsions : Perception, Mesures et Modèles", Université de la Méditerranée d'Aix-Marseille II, Thèse de Doctorat, (2005, en préparation)

[15] Stevens S.S., "Calculation of the loudness of complex noise", J. Acoust. Soc. Am., 28, 807-832, 1956

[16] Blommer M., Otto N., Wakefield G., Feng B.J., Jones C., "Calculating the loudness of impulsive sounds", SAE Transactions, 1995.

[17] Gerken G. M., Bhat V. K. H. et Hutchinson-Clutter M, "Auditory temporal integration and the power function model", J. Acoust. Soc. Am. 88 (2), 767-778, 1990

[18] Scharf B., "Loudness", in Handbook of perception, edited by E. C. Carterette and M. P. Friedman, Academic Press New York, 187-242, 1978.

[19] Zwicker E. et Feldtkeller R., "Psychoacoustique", Masson, Paris, 1981.

[20] Meunier S., Marchioni A. et Rabau G. "Subjective evaluation of loudness models using synthesized and environmental sounds", actes d'Inter Noise 2000, Nice, France, vol. 4, 2205-2209, 2000.

## Références bibliographiques

[1] ISO 532B, "Méthode de calcul du niveau d'isotonie", norme internationale, 1975.

[2] Zwicker E., "Über psychologische und methodische Grundlagen der Lautheit, Acustica", Vol. 8, 237-258, 1958.

[3] Zwicker E., "Procedure for calculating loudness of temporally variable sounds", J. Acoust. Soc. Am., Vol. 62, n°3, p 675-682, 1977.

[4] Moore B.C.J. et Glasberg B.R., "A revision of Zwicker's loudness model", Acustica-acta acustica, 82, 335-345, 1996.

[5] Moore, B. C. J., Glasberg, B. R. et Baer, T. "A model for the prediction of threshold, loudness and partial loudness", J. Audio Eng. Soc., 45, pp 224-240, 1997.

[6] Glasberg B. R. et Moore B. C. J., "A model of loudness applicable to time-varying sounds", J. Audio Eng. Soc., 50, n° 5, 331-342, 2002.

[7] Zwicker E. et Fastl H., "Psychoacoustics: Facts and models", 2nd Edition, Springer-Verlag, Berlin, 1999.

[8] Boulet I., "Comparison between four methods of loudness estimation of stationary and non-stationary sounds", 7ème Congrès Français d'Acoustique (CFA) et 30ème Congrès de la Société Allemande d'Acoustique (DAGA), Strasbourg, France, (2004)

[9] Stevens S.S., "The measurement of loudness", J. Acoust. Soc. Am., 27, 815-820, 1955.

[10] Gescheider, G.A., Psychophysics, Method, Theory, and Application, Second Edition, Lawrence Erlbaum Associates, New Jersey, USA, 1985.

[11] Jestead, W., "An adaptive procedure for subjective judgements", Percept. Psychophys. 28, 85-88, 1980.

[12] Buus, S. Florentine, M., and Poulsen, T., "Temporal integration of loudness, loudness discrimination, and the form of the loudness function", J. Acoust. Soc. Am., 101, 669-680, (1997)

[13] Boulet, I., Marozeau J. et Meunier S., "Temporal summation of loudness as a function of frequency and temporal pattern", 33rd International Congress