



Ministère de l'Ecologie et du Développement Durable

Programme « Bruit et nuisances sonores »

Collaboration GENESIS / CNRS-LMA

Estimateur de sonie de sons non stationnaires

Auteurs : Isabelle Boulet
Fabrice Gagneux
Sabine Meunier
Patrick Boussard
Jeremy Marozeau
Guy Rabau

Document FG/RP/5278
9 décembre 2005

Table des matières

FICHE RECAPITULATIVE DU PROJET	3
INTRODUCTION	5
I- MODELES ET METHODES DE MESURE DE LA SONIE DE SONS NON STATIONNAIRES	7
I-1 LES MODELES DE SONIE DE ZWICKER ET FASTL (1999) ET DE GLASBERG ET MOORE (2002)	7
I-2 METHODES DE MESURE DE LA SONIE ET DU NIVEAU D'ISOSONIE.....	9
I-3 INDICES POUR JUGER DE LA PRECISION DES MODELES DE SONIE.....	10
II- INDICATEURS DE SONIE ET LIMITES D'APPLICATION	11
II-1 BASE DE DONNEES EXPERIMENTALES : LIGNES ISOSONIQUES ET INTEGRATION TEMPORELLE ..	11
II-2 DOMAINES D'APPLICATION ET LIMITES D'INDICATEURS DE SONIE.....	11
III- ETABLISSEMENT D'UN NOUVEAU MODELE DE SONIE DE SONS IMPULSIONNELS	13
III-1 AMELIORATION DE L'ESI : PRINCIPE ET AJUSTEMENT.....	13
<i>III-1.1 Principe de l'estimateur de sonie d'impulsion modifié (ESI_{bc}).....</i>	<i>13</i>
<i>III-1.2 Mesure des exposants des onze fonctions de sonie.....</i>	<i>15</i>
III-2 PROCESSUS DE VALIDATION.....	15
CONCLUSION.....	16
VALORISATION.....	18
ANNEXE 1 : LES MODELES DE SONIE DE SONS NON STATIONNAIRES.....	20
ANNEXE 2 : METHODES DE MESURE DE LA SONIE ET DU NIVEAU D'ISOSONIE....	30
ANNEXE 3 : BASE DONNEES EXPERIMENTALES	42
ANNEXE 4 : L'ESTIMATEUR DE SONIE D'IMPULSION (ESI)	51
ANNEXE 5 : DOMAINE D'APPLICATION DES MODELES DE SONIE ET DE L'ESI	59
ANNEXE 6 : MESURE DES FONCTIONS DE SONIE POUR L'ESI_{BC}	75
ANNEXE 7 : PROCESSUS DE VALIDATION	77
GLOSSAIRE	82
RÉFÉRENCES.....	84

Fiche récapitulative du projet

Le projet a débuté le 10 novembre 2003 pour une durée de deux ans. Ce rapport relate sous la forme d'un document principal et de manière synthétique les résultats des différentes études réalisées dans le cadre du projet. Les détails de la méthodologie et des résultats de chaque phase de l'étude sont présentés en annexes.

Partenaires scientifiques

Responsable et coordinateur scientifique

Entreprise : **GENESIS S.A.**

Prénom : **Patrick** Nom : **Boussard** Fonction : **Directeur Général**

Adresse : **Bât. Gérard Mégie - Domaine du petit Arbois BP 69 - 13545 Aix-en-Provence cedex 04**

Tél : **04.42.90.40.00** Fax : **04.42.97.12.88** Mél : **patrick.boussard@genesis.fr**

Partenaire

Organisme : **LMA - CNRS**

Prénom : **Sabine** Nom : **Meunier** Fonction : **Chargé de recherche**

Adresse : **31 chemin Joseph Aiguier – 13402 Marseille cedex 20**

Tél : **04.91.16.41.75** Fax : **04.91.77.55.65** Mél : **meunier@lma.cnrs-mrs.fr**

Résumé

La sonie, intensité subjective d'un son, est un paramètre psychoacoustique primordial dans l'évaluation de la gêne et dans les études sur la qualité sonore. Aujourd'hui, le seul moyen de connaître précisément la sonie est de mettre en œuvre des tests psychoacoustiques, longs et onéreux. Il existe cependant à l'heure actuelle des modèles qui estiment correctement la sonie de sons stationnaires, tels ceux de Zwicker (1958) et de Moore et coll. (1997). D'autres modèles (Zwicker et Fastl (1999) et Glasberg et Moore (2002)) permettent de calculer la sonie en fonction du temps de sons non stationnaires mais ils présentent certaines limites pour évaluer la sonie globale, en particulier pour des sons impulsionnels. L'objet de nos recherches, financées par le MEDD dans le cadre du programme « Bruits et nuisances sonores », est donc d'étudier la sonie de sons impulsionnels et de trains d'impulsions. Ces deux types de sons, cas particulier de sons non stationnaires, suscitent en effet un intérêt grandissant d'un point de vue industriel et scientifique. Un grand nombre de données expérimentales ont été établies sur ce type de sons (lignes isosoniques¹ et intégration temporelle² de sons impulsionnels, mesure de sonie de trains d'impulsions) pour, non seulement déterminer les relations entre la sonie et les paramètres physiques du signal (fréquence, niveau et durée), mais aussi tester les différents indicateurs³ proposés dans la littérature et en déterminer les domaines d'application. Toutes ces données ont alors contribué à définir

¹ **Ligne isosonique** : Représente le niveau que doit avoir un son pur pour conserver une sonie constante en fonction de la fréquence.

² **Intégration temporelle** : La sonie augmente régulièrement lorsque la durée d'un son d'amplitude constante augmente, de quelques millisecondes jusqu'à une valeur critique, appelée durée critique, qui varie de 50 à 400 ms selon les auteurs. Ce phénomène est appelé intégration temporelle.

³ **Indicateur** : Valeur extraite d'un modèle ou d'un estimateur permettant de calculer la sonie globale de son non stationnaire et impulsionnel.

deux nouveaux indicateurs, calculés à partir du modèle de Zwicker et Fastl (1999), pour donner une première approximation de la sonie de sons impulsionnels. Nous avons ensuite procédé à des mesures psychoacoustiques de fonctions de sonie⁴ d'impulsions de sons purs, afin de tenir compte de l'influence de la fréquence porteuse sur la sonie. Ces données expérimentales ont contribué à déterminer un nouvel estimateur (noté ESI_{bc}) qui améliore le premier estimateur de sonie d'impulsion, noté ESI, que notre équipe de recherche avait développé. Nous avons alors entamé un processus de validation de ce nouvel estimateur ESI_{bc} basé sur des mesures de sonie de sons impulsionnels de synthèse (231 sons) et « naturels » (26 sons). Dans le cadre du projet, l' ESI_{bc} est en cours d'intégration dans le module psychoacoustique du logiciel LEA développé et commercialisé par GENESIS.

Summary

Loudness is an important psychoacoustics parameter in the evaluation of annoyance and studies on sound quality. Currently, the only means of knowing loudness precisely is to implement long and expensive psychoacoustics tests. There are, however, models which correctly estimate the loudness of stationary sounds (Zwicker (1958), Moore et coll. (1997)). Other models make it possible to calculate a time varying loudness adapted to non-stationary sounds (Zwicker et Fastl (1999), Glasberg et Moore (2002)), but they are limited in the evaluation of total loudness, in particular for impulsive sounds. The object of our research, financed by the MEDD (within the framework of the program "Bruits et nuisances sonores") has been to study the loudness of impulsive sounds and trains of impulses. From an industrial and scientific point of view, these two types of sounds, particular cases of non-stationary sounds, are very interesting. A large amount of experimental data has been generated on these types of sounds, in order to determine the relations between the loudness and the physical parameters of the signal and to test the various indicators proposed in the literature. This data has subsequently contributed to the definition of two new indicators, derived from the model of Zwicker and Fastl (1999), to give a first approximation of the loudness of impulsive sounds. We then carried out psychoacoustics measurements of loudness functions of pure tone impulsive sounds, in order to take account of the influence of the carrier frequency on loudness. This experimental data has contributed to the determination of a new estimator (ESI_{bc}) which improves the first loudness estimator of impulsive sounds (ESI) that our team of research had developed previously. We then started a validation process of this new estimator based on loudness measurement of impulsive sounds (231 synthesized sounds and 26 "natural" sounds).

⁴ **Fonction de sonie** : La fonction de sonie est la relation entre l'intensité (ou la pression acoustique) et la sonie.

Introduction

La détermination de la sonie, intensité subjective d'un son, est essentielle pour évaluer la gêne de signaux acoustiques. En effet, plusieurs études ont été recensées sur l'importance de la sonie dans l'évaluation de la gêne (Berglund, 1990 ; Parizet et coll., 1996 ; Meunier et Marchioni, 1998 ; Bousard et Boulet, 2001). Parmi elles, celles réalisées sur des types de bruits spécifiques tels que des bruits de voisinage, des bruits de chantier ou encore des bruits de transport (trafic routier ou ferroviaire) ont confirmé que le niveau de gêne s'explique essentiellement par l'intensité perçue du bruit. En ce qui concerne les signaux sonores impulsionnels, la sonie est encore admise comme un des facteurs importants du désagrément (Cops et Myncke, 1976).

La sonie est aussi un indicateur très influent dans les études sur la qualité sonore (Fastl, 1997 ; Ohta et coll., 1999). Par exemple, si on teste la qualité d'enceintes acoustiques, la sonie est une des dimensions mise en évidence lors d'un jugement de dissemblance (Gabrielson et Sjögren, 1979). Pour une analyse plus fine des dimensions reflétant une différence qualitative, il est important de réaliser des tests à sonie égalisée. Il est donc nécessaire de connaître la sonie des signaux testés.

La sonie, en tant que grandeur subjective, est mesurée par des méthodes psychoacoustiques. Celles-ci consistent à recueillir le jugement d'un grand nombre d'auditeurs. Les principales méthodes de mesure de la sonie sont des méthodes d'estimation directe, d'ajustement ou adaptative. Ces méthodes sont coûteuses car longues à mettre en œuvre. Ainsi, plusieurs solutions ont été envisagées pour calculer la sonie à partir de données physiques du signal.

Une première approximation de la sonie, largement utilisée en métrologie, consiste à pondérer le niveau de pression acoustique pour tenir compte des variations de la sensibilité de l'oreille en fonction de la fréquence. Selon la pondération choisie, basée sur les lignes isosoniques à 40, 70 et 100 phons, on utilise le dB(A), le dB(C) ou le dB(D) respectivement. Cependant, ces niveaux pondérés ne tiennent pas compte d'autres phénomènes physiologiques, comme le masquage fréquentiel et temporel ou le mécanisme de filtrage de l'oreille, et restent insuffisants pour estimer correctement la sonie de sons complexes (Marozeau et coll., 2005).

Une autre approche est basée sur la modélisation du système auditif. De nombreuses études ont été alors réalisées, d'une part pour déterminer des procédures expérimentales pour la mesurer, d'autre part pour connaître les processus auditifs de la formation de la sonie et savoir de quels facteurs physiques elle dépend afin de pouvoir calculer la sonie à partir d'algorithmes. Ces études ont, pour la plupart, été réalisées sur des sons stationnaires et quasi stationnaires et des modèles de sonie ont été proposés pour ces types de sons. Ces vingt dernières années, l'intérêt des auteurs s'est porté plus particulièrement sur les sons non stationnaires et les recherches se sont intensifiées sur ce sujet avec la croissance des études sur la qualité sonore. Il existe aujourd'hui des modèles permettant de calculer la sonie en fonction du temps (Zwicker et Fastl (1999) et Galsberg et Moore (2002)) en tenant compte du masquage temporel. Ces auteurs préconisent l'utilisation d'indicateurs statistiques pour approcher la sonie globale, mais ces indicateurs ont aussi montré certaines limites (Meunier et coll., 2000). En ce qui concerne les sons de très courtes durées, cas particulier de sons non stationnaires, les études restent rares et il n'existe pas à ce jour de modèle de sonie validé pour ce type de sons pourtant très présents dans notre environnement sonore.

La première étape a consisté à étudier les modèles et les méthodes de mesure de la sonie de sons non stationnaires (chapitre I). Nous nous sommes ensuite intéressés plus particulièrement aux sons impulsionnels et aux trains d'impulsions pour lesquels nous avons recueillis une base de données expérimentales. Ces données nous ont permis de mettre en évidence les limites de différents indicateurs

de sonie (chapitre II). La troisième étape du projet de recherche a été consacrée à l'amélioration de l'estimateur ESI en considérant les données expérimentales et les résultats de la première étape. Nous avons ainsi entamé une autre campagne de mesures psychoacoustiques de fonctions de sonie d'impulsions de sons purs afin de tenir compte de l'influence de la fréquence porteuse sur la sonie (paragraphe III-1). Ces données expérimentales ont contribué à déterminer un nouvel estimateur (noté ESI_{bc}) qui améliore l'ESI. La dernière partie du projet (paragraphe III-2) présente un processus de validation de ce nouvel estimateur ESI_{bc} basé sur des mesures de sonie de sons impulsionnels de synthèse (231 sons) et « naturels » (26 sons).

I- Modèles et méthodes de mesure de la sonie de sons non stationnaires

I-1 Les modèles de sonie de Zwicker et Fastl (1999) et de Glasberg et Moore (2002)

Deux modèles connus permettent de calculer la sonie en fonction du temps (Zwicker et Fastl, 1999 et Glasberg et Moore, 2002). Ces auteurs préconisent des indicateurs statistiques pour estimer la sonie globale de sons évoluant dans le temps. Ces indicateurs sont selon Zwicker et Fastl (1999) le N4 ou le N5 (sonie dépassée 4 ou 5 % du temps) et selon Glasberg et Moore (2002), le STLmax (maximum de la sonie à court terme).

Ces deux modèles ont été implémentés et validés sur des données expérimentales de la littérature par notre équipe de recherche dans le cadre du projet.

Le principe de base du modèle de Zwicker et Fastl (1999) est le calcul de l'excitation de la membrane basilaire par bandes critiques, en prenant en compte diverses pondérations modélisant les filtrages de l'oreille externe et de l'oreille moyenne. Le modèle tient aussi compte du masquage temporel et fréquentiel, et enfin de l'intégration de la sonie spécifique sur les vingt-quatre bandes critiques toutes les deux millisecondes.

Le principe du modèle de Glasberg et Moore (2002) présente certaines différences par rapport à celui de Zwicker et Fastl (1999) notamment en ce qui concerne le calcul des filtres auditifs, les corrections de champ, le calcul de l'excitation et du masquage temporel. De plus, la sonie est calculée chaque milliseconde.

La figure 1 permet de résumer les principes des modèles de sonie de Zwicker et Fastl (1999) et de Glasberg et Moore (2002) et de mettre en évidence les. La description des deux modèles est détaillée en annexe 1.

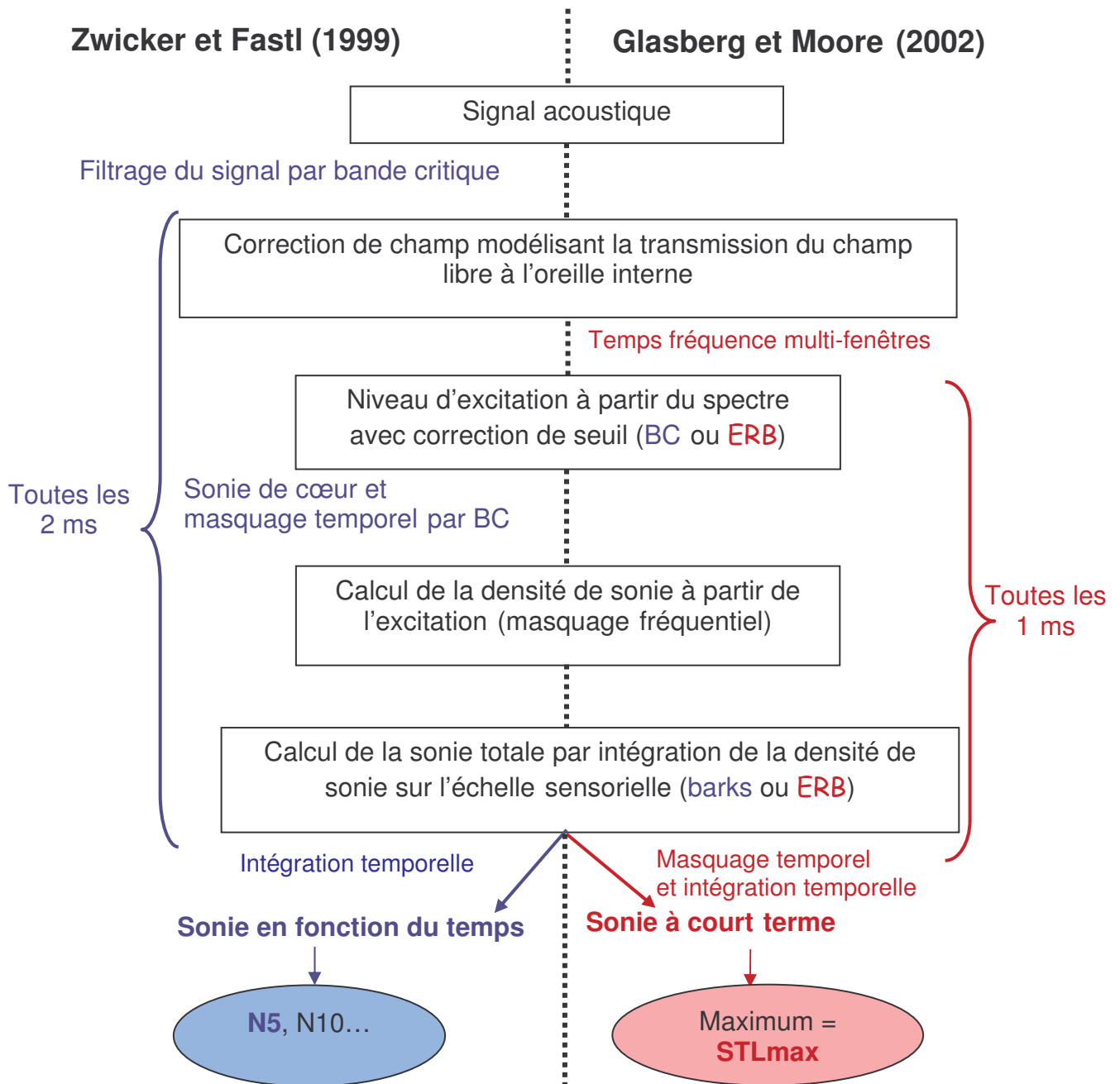


Figure 1 : Ce schéma résume le principe de calcul de chacun des modèles de sonie de sons non stationnaires. L'abréviation BC signifie bande critique⁵ et ERB⁶ signifie Equivalent Rectangular Bandwidth.

⁵ **Bande critique :** À intensité constante, la sonie augmente avec la largeur spectrale. Mais cette augmentation n'apparaît que lorsque la largeur dépasse une certaine valeur dépendant de la fréquence. Cette largeur spectrale est appelée bande critique.

⁶ **ERB :** Equivalent Rectangular Bandwidth. Au lieu d'utiliser le modèle de bandes critiques, Moore et Glasberg (1983) introduisent la notion de filtre rectangulaire équivalent (ERB), dont la largeur dépend aussi de la fréquence.

I-2 Méthodes de mesure de la sonie et du niveau d'isophonie⁷

Nous nous intéressons dans ce paragraphe aux différentes méthodes expérimentales de mesure du niveau d'isophonie pour différents types de sons, dont le détail des mesures est décrit en annexe 2.

A notre connaissance, peu d'études ont été réalisées sur la comparaison des méthodes dans le but de déterminer celle qui présentera le meilleur compromis entre précision, rapidité et fiabilité pour mesurer la sonie ou le niveau d'isophonie. Des recherches sur les méthodes psychophysiques ont été réalisées dans d'autres domaines de la psychoacoustique (Marvit et coll., 2003). Il est important, pour la métrologie, de connaître la méthode de mesure du niveau d'isophonie la mieux adaptée au type de sons étudiés.

Cette nouvelle étude a trois objectifs : le premier est de connaître les avantages et les inconvénients de chaque méthode ; le deuxième est de déterminer la méthode la plus appropriée en fonction du type de sons et de la précision désirée afin d'obtenir une référence fiable autour de laquelle construire l'estimateur de sonie ; le troisième est de définir l'erreur maximale acceptable par le modèle de sonie.

Lors de cette étude, nous avons comparé quatre méthodes de mesure de la sonie. Les deux premières sont fréquemment utilisées. Il s'agit de la méthode d'estimation de grandeur sans référence (Stevens, 1956) et de la méthode d'ajustement (Gescheider, 1985). La troisième procédure testée est la méthode adaptative à choix forcé sur deux intervalles (2AFC-2I) du type 2down-1up et 2up-1down (Jestead, 1980). La dernière méthode est également une méthode adaptative proposée par Buus et ses collaborateurs pour éviter certains biais présents dans d'autres méthodes adaptatives. C'est la méthode adaptative à poursuites multiples (Buus et coll., 1997).

Les sons ont été classés en trois groupes, décrits en annexe 2. Chaque groupe représente une catégorie de sons. Nous avons étudié neuf sons stationnaires, dix sons impulsionnels et dix trains d'impulsions. La moitié des signaux constituant notre banque d'échantillons sonores a été enregistrée et l'autre moitié synthétisée. Les enregistrements ont été effectués dans la salle anéchoïque du laboratoire.

L'ensemble des résultats a permis de conclure que la méthode d'ajustement présente le meilleur compromis avec une précision de 4,7 phons (valeur moyenne des écarts types, cf. tableau A-2.6) et une fiabilité de l'ordre de 2 phons (valeur de l'erreur intra-individuelle, cf. paragraphe 2.3 de l'annexe 2). Les méthodes adaptative et adaptative à poursuites multiples donnent des écarts types comparables à la méthode d'ajustement mais durent quatre fois plus longtemps. C'est donc la méthode d'ajustement que nous avons retenue pour mesurer les niveaux d'isophonie dans la suite de l'étude.

La méthode d'estimation directe, plus rapide, présente cependant une variabilité plus importante (écart type d'environ 6,7 phons). C'est aussi la seule des méthodes étudiées qui permet de mesurer directement la sonie.

Au vu des valeurs moyennes des écarts types, et en vue de la construction d'un modèle de sonie pour des sons non stationnaires, une mesure du niveau d'isophonie avec une précision inférieure à 4 phons sera retenue comme précise.

⁷ **Niveau d'isophonie** : Niveau en dB SPL que doit avoir un son pur de référence de 1 kHz pour qu'il ait la même sonie que le son test. Le niveau d'isophonie s'exprime en phons.

I-3 Indices pour juger de la précision des modèles de sonie

Afin de pouvoir comparer les différents modèles de sonie et de juger de leur précision, plusieurs indices ont été proposés (Marozeau et col., 2005). Ils sont décrits dans ce qui suit.

Le **premier indice est le MAR** (Mean Absolute Residual), c'est l'écart moyen. Il a été utilisé par Gerken et coll. (1990) pour comparer différentes fonctions mathématiques permettant de mieux caractériser l'intégration temporelle. Le MAR est défini comme la moyenne des valeurs absolues de la différence entre le niveau d'isotonie calculé ($L_{\text{calculé}}$) et le niveau d'isotonie mesuré ($L_{\text{mesuré}}$) :

$$MAR = \frac{\sum_{i=1}^n |L_{i \text{ calculé}} - L_{i \text{ mesuré}}|}{n} \quad \text{où } n \text{ représente le nombre de stimuli testés.}$$

Cet indice permet de déterminer si l'indicateur donne, en moyenne, de bonnes prédictions du niveau d'isotonie.

Le **deuxième indice est le Rmax**. C'est le maximum de la valeur absolue des différences entre le niveau d'isotonie calculé ($L_{\text{calculé}}$) et le niveau d'isotonie mesuré ($L_{\text{mesuré}}$) :

$$Rmax = \max(|L_{i \text{ calculé}} - L_{i \text{ mesuré}}|)$$

Cet indice permet de déterminer la plus grande erreur effectuée par le modèle ou l'indicateur pour prédire le niveau d'isotonie. Il permet en quelque sorte de tester la robustesse de l'indicateur.

Le **troisième indice est noté Rmean**. C'est la moyenne des différences entre le niveau d'isotonie calculé ($L_{\text{calculé}}$) et le niveau d'isotonie mesuré ($L_{\text{mesuré}}$) :

$$Rmean = \frac{\sum_{i=1}^n (L_{i \text{ calculé}} - L_{i \text{ mesuré}})}{n}$$

Cet indice permet de savoir si les erreurs commises par l'indicateur 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-dessus ($Rmean > 0$) ou au-dessous ($Rmean < 0$) de la valeur de la mesure. Il permet de vérifier que le modèle ou l'estimateur ne possède pas de biais systématique. En effet, si le modèle avait un biais systématique le MAR et le Rmean auraient la même valeur au signe près.

Enfin, le **dernier indice est le Rstd**. C'est l'écart type des différences entre le niveau d'isotonie calculé et le niveau d'isotonie mesuré.

$$Rstd = std(L_{i \text{ calculé}} - L_{i \text{ mesuré}})$$

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

II- Indicateurs de sonie et limites d'application

Parmi les sons non stationnaires, les sons impulsionnels et les trains d'impulsions suscitent un grand intérêt auprès des industriels et des scientifiques. Nous nous intéressons dans cette partie plus particulièrement aux sons impulsionnels isolés dont la répétition constitue un train d'impulsions.

II-1 Base de données expérimentales : Lignes isosoniques et intégration temporelle

Nous décrivons ici comment nous avons obtenu une base de données expérimentales concernant des sons impulsionnels dont la forme est décrite sur la figure 2. Ces données sont comparées à celles de la littérature concernant les sons brefs afin de déterminer les améliorations envisageables des modèles de sonie et de l'estimateur de sonie d'impulsion. Enfin, les données permettront de définir la précision des modèles et de l'estimateur de sonie.

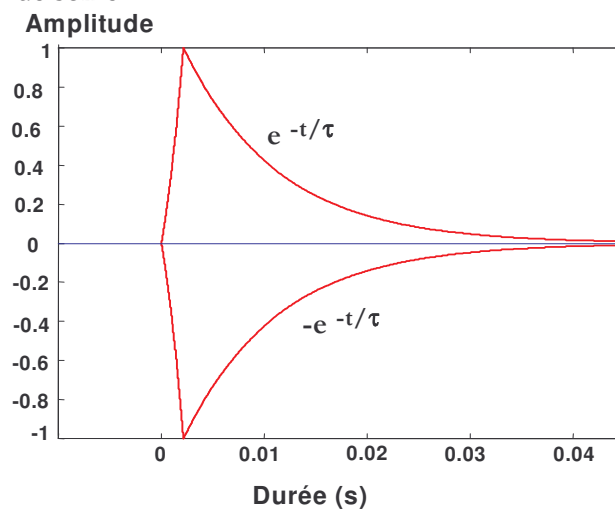


Figure 2 : Modèle d'enveloppe temporelle utilisée dans cette étude pour synthétiser des sons impulsionnels de forme « naturelle ».

Nous avons mesuré des lignes isosoniques afin d'obtenir des données expérimentales permettant d'étudier l'influence de la fréquence et du niveau sur la précision des indicateurs de sonie. Nous avons aussi étudié l'intégration temporelle non seulement pour des signaux en forme de créneau, les plus souvent étudiés, mais aussi pour des signaux en forme d'impulsion, plus proches de sons « naturels » afin de tenir compte de l'influence de la durée. Les expériences et leurs résultats sont décrites en annexe 3.

Ces expériences ont permis de recueillir un grand nombre de données expérimentales qui nous permettent d'étudier les domaines d'application de différents indicateurs. Les comparaisons sont présentées au paragraphe II-2.

II-2 Domaines d'application et limites d'indicateurs de sonie

Pour estimer la sonie de sons impulsionnels, nous avons recensé plusieurs indicateurs tels le N5 issu du modèle de Zwicker et Fastl (1999) et le STLmax issu du modèle de Glasberg et Moore (2002). Notre équipe de recherche a déterminé deux autres indicateurs (Nmax et N30ms) calculés à partir du modèle

de Zwicker et Fastl (1999). Enfin, le dernier estimateur testé (ESI) avait été développé dans le cadre d'une collaboration entre la société GENESIS et le Laboratoire de Mécanique et d'Acoustique (Meunier et coll., 2001).

Nous comparons dans ce paragraphe les niveaux d'isophonie calculés par les différents modèles de sonie et l'ESI aux données expérimentales que nous avons obtenues lors de ce projet de recherche. A partir des données expérimentales que nous avons obtenues lors de ce travail, nous avons testé les modèles sur des sons non stationnaires, en particulier sur des sons impulsionnels. Nous rappelons ci-dessous les indices de calcul de la sonie que nous avons étudiés :

- **N5**

Le modèle de Zwicker et Fastl (1999) permet de calculer la sonie et le niveau d'isophonie en fonction du temps de sons non stationnaires. Différents indicateurs, préconisés par les auteurs pour approcher le niveau d'isophonie global, sont calculés à partir du niveau d'isophonie en fonction du temps. Il s'agit ici du niveau d'isophonie dépassé 5% du temps (N5). Il est noté Zwicker 1999 (N5).

- **Nmax, N30ms**

L'équipe de recherche a été amenée à définir d'autres indicateurs qui ont semblé mieux adaptés au calcul de la sonie de sons impulsionnels. Ils sont notés Zwicker 1999 (Nmax) et Zwicker 1999 (N30ms). Ces deux derniers indicateurs nous semblent mieux adaptés au calcul du niveau d'isophonie global de sons non stationnaires et impulsionnels.

- **STLmax**

Le modèle de Glasberg et Moore (2002) permet de calculer la sonie et le niveau d'isophonie en fonction du temps de sons variant dans le temps. Pour approcher le niveau d'isophonie global de sons variant rapidement dans le temps, les auteurs préconisent de calculer la valeur maximale du niveau d'isophonie « à court terme » (cf. annexe 1), noté Moore 2002 (STLmax).

- **Estimateur de sonie de sons impulsionnels (ESI)**

L'estimateur de sonie d'impulsion, noté ESI, est basé sur un principe simple. En effet, des études antérieures aux projets (Boullet, 2005) ont montré que la sonie de sons impulsionnels est principalement corrélée à l'énergie. Des études complémentaires ont montré qu'il était nécessaire de tenir compte de la durée du signal impulsionnel pour améliorer le calcul de la sonie. Le détail de l'élaboration de l'ESI est décrit en annexe 4. Cet estimateur, résultat de nos propres recherches, permet de calculer le niveau d'isophonie de sons impulsionnels.

Tout d'abord, nous avons étudié les limites des modèles sur les lignes isophoniques de sons impulsionnels que nous avons établies (voir figures A-5.2, A-5.3 et A-5.4 de l'annexe 5). Tous les modèles prévus pour des sons non stationnaires (Nmax, N30ms, STLmax) prédisent à moins de 5 phons près le niveau d'isophonie mesuré (voir tableau A-5.1). Nous avons aussi observé que plus le niveau d'isophonie mesuré augmente, plus le niveau d'isophonie calculé est inférieur à ce dernier. Enfin, seul l'estimateur de sonie d'impulsion (ESI) présente une déficience pour des fréquences très éloignées de 1 kHz.

Nous avons ensuite comparé les calculs obtenus par les modèles aux mesures du niveau d'isophonie de créneaux et d'impulsions sur nos données concernant l'intégration temporelle (voir figures A-5.5 et A-5.6). Les sons testés étaient des sons purs de 1 kHz dont les durées allaient de 10 millisecondes à 3 secondes. Les modèles ne donnent plus de bons résultats pour des impulsions de durée supérieures à

300 ms. Pour ces durées, les indicateurs utilisés (N_{max} , N_{30ms} , STL_{max}) donnent des valeurs supérieures aux niveaux d'isophonie mesurés de plus de 2 phones (voir tableau A-5.2). L'estimateur de sonie d'impulsion ne prédit plus correctement le niveau d'isophonie pour des durées supérieures à 300 ms.

Nous avons ensuite comparé les calculs du niveau d'isophonie aux mesures de sons impulsionnels de durée et de fréquence différentes (figures A-5.7 à A-5.16). Les durées des sons variaient de quelques millisecondes à quelques centaines de millisecondes et les fréquences étaient comprises entre 125 Hz et 8 kHz. On a ainsi pu montrer les limites du STL_{max} et du N_5 . En effet, pour ces deux indicateurs, les écarts dépassent 4 phones. Pour le N_5 , les écarts sont d'autant plus élevés que les fréquences sont plus élevées. Quant au STL_{max} , il prédit moins bien le niveau d'isophonie de sons impulsionnels aux basses et hautes fréquences. L'estimateur de sonie d'impulsion (ESI) prédit assez bien les niveaux d'isophonie mesurés, à moins de 4 phones près en moyenne, de sons dont la durée est inférieure à 300 ms. Les écarts sont d'autant plus grands que les fréquences sont plus éloignées de 1 kHz et que la durée est supérieure à 300 ms pour les créneaux et à 1 s pour les impulsions (voir tableau A-5.3).

Pour les sons impulsionnels que nous avons testés (figure A-5.17), il semble que le modèle le mieux adapté au calcul du niveau d'isophonie est le modèle de Zwicker et Fastl (1999). Plus précisément, ce sont les indicateurs N_{max} et N_{30ms} qui estiment en moyenne le mieux le niveau d'isophonie mesuré avec un $MAR=4$ phones (voir tableau A-5.4). Par contre on notera que l'écart maximum (R_{max}) est de l'ordre d'une dizaine de phones ce qui reste une valeur élevée pour estimer correctement le niveau d'isophonie de tout type de son. Il semble que ces indicateurs soient plus fiables sur des sons de synthèse ($MAR=2$ phones). En effet, sur les expériences comportant des sons réels de l'environnement (stationnaires, non stationnaires et impulsionnels) le MAR moyen s'élève à 7 phones.

Nous avons aussi montré les limites des modèles sur les trains d'impulsions (figure A-5.18). En effet, pour ce type de son, aucun modèle ne donne un niveau d'isophonie proche de celui mesuré (voir tableau A-5.4).

Pour résumer, les indicateurs tels que le N_{max} et le N_{30ms} , issus du modèle de Zwicker et Fastl (1999), sont en moyenne les plus précis et les plus stables (MAR compris entre 2 et 5 phones), avec tout de même un R_{max} de 10 phones, pour calculer le niveau d'isophonie de sons impulsionnels.

III- Etablissement d'un nouveau modèle de sonie de sons impulsionnels

Les indicateurs présentés jusqu'alors ne permettent pas encore de calculer systématiquement de manière précise le niveau d'isophonie global de sons impulsionnels. Nous présentons donc dans cette partie l'élaboration d'un nouvel estimateur mieux adapté au calcul de la sonie de sons impulsionnels.

III-1 Amélioration de l'ESI : principe et ajustement

III-1.1 Principe de l'estimateur de sonie d'impulsion modifié (ESI_{bc})

Le principe de l'estimateur de sonie d'impulsion modifié ESI_{bc} est décrit ci-dessous et schématisé figure 3.

Le signal acoustique est filtré pour obtenir un signal temporel dans chacune des bandes critiques. L'étape suivante du modèle est de calculer par bande l'énergie et le temps de descente. La sonie de

cœur, basée sur l'estimateur de sonie d'impulsion initial (ESI), est ensuite calculée comme le produit de l'énergie, à l'exposant alpha, et de la durée d'extinction de l'impulsion, à l'exposant bêta. Enfin, la densité de sonie est intégrée sur les vingt-quatre bandes critiques pour prendre en compte le masquage fréquentiel.

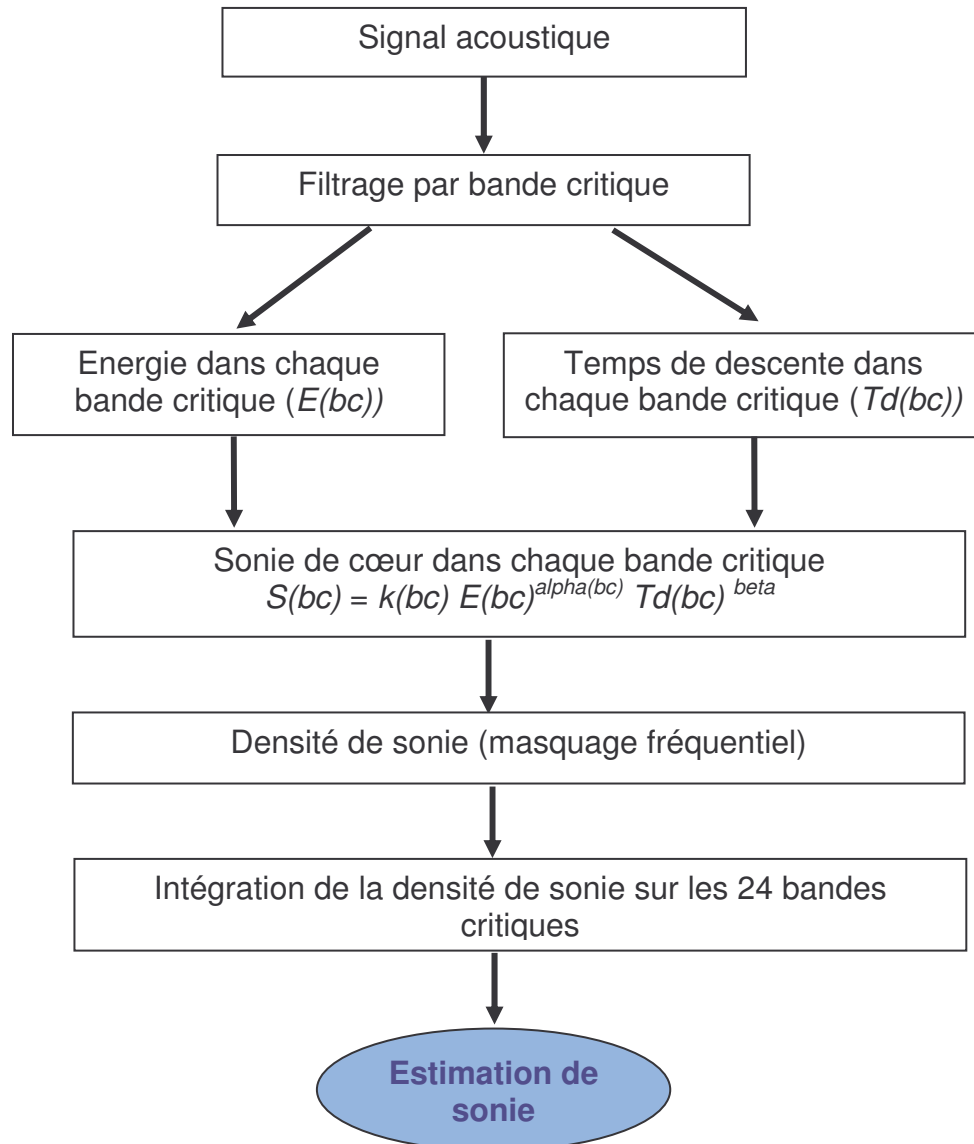


Figure 3 : Principe de calcul de l'estimateur de sonie d'impulsion modifié noté ESI_{bc} . E est l'énergie, Td la durée de décroissance de l'impulsion filtrée dans une bande critique et bc représente un indice de bande critique (allant de 1 à 24).

L'exposant alpha est déterminé à partir du calcul de l'exposant de la fonction de sonie d'impulsions de sons purs dont les fréquences des porteuses correspondent aux fréquences centrales des bandes critiques. La mesure de ces fonctions de sonie est présentée dans le paragraphe suivant. L'exposant bêta, quant à lui, a été déterminé lors de recherches antérieures au projet (voir annexe 4).

III-1.2 Mesure des exposants des onze fonctions de sonie

La mesure des exposants des onze fonctions de sonie, pour plusieurs fréquences de porteuse, permet d'affiner l'estimateur de sonie d'impulsion (ESI) et de tenir compte de l'influence de la fréquence.

Les sons impulsionnels étudiés ont tous été synthétisés. Leur enveloppe temporelle avait une attaque modélisée par une fonction linéaire croissante sur 5 ms, immédiatement suivie d'une décroissance d'amplitude modélisée par une fonction exponentielle décroissante. Le signal était coupé à -60 dB par une extinction linéaire. La structure fine de remplissage des enveloppes, était composée de sons purs de 250, 450, 700, 1 000, 1 370, 1 850, 2 500, 3400, 4800, 7 000 et 10500 Hz. Chaque son, pour une fréquence était joué à 11 niveaux crêtes compris entre 50 et 100 dB SPL par pas de 5 dB. Le détail des mesures est décrit en annexe 6.

Les valeurs des exposants des onze fonctions de sonie sont reportées dans le nouvel estimateur de sonie d'impulsion (ESI_{bc}). Pour les fréquences centrales des bandes critiques non étudiées, les valeurs des exposants ont été interpolées à partir des valeurs mesurées.

III-2 Processus de validation

Le processus de validation de l' ESI_{bc} est basé sur la comparaison de la sonie mesurée de différents sons impulsionnels et calculée par ce dernier et par les autres modèles cités précédemment.

Les sons étudiés étaient soit synthétisés (231 sons) soit enregistrés (26 sons). Les tableaux récapitulant la description et les caractéristiques physiques des signaux étudiés sont reportés en annexe 7 ainsi qu'une description des mesures de la sonie.

La sonie mesurée a été comparée à la sonie calculée par l'ESI, l' ESI_{bc} et le Zwicker 1999 (N30ms).

Les valeurs des indices permettant de juger de la précision des modèles sont indiquées dans les tableaux A-7.3, A-7.4, A-7.5 et A-7.6. Il en ressort que l' ESI_{bc} permet d'estimer avec le plus de précision la sonie mesurée des 257 sons avec un MAR moyen de 3 phones. La dispersion des mesures étant de l'ordre de 4 phones, l' ESI_{bc} est un estimateur précis de sonie de sons impulsionnels.

Conclusion

Les avancées scientifiques que ce projet a permis dans le domaine de la psychoacoustique ont fait l'objet de publications et de communications orales lors de congrès (Boullet (2004), Boullet et coll. (2004) et Marozeau et coll. (2005)). Trois articles correspondent aux études réalisées dans la première étape du projet, et un article est en prévision pour le 8^{ème} Congrès Français d'Acoustique sur la dernière étape du projet (« Un nouvel estimateur de sonie d'impulsion : élaboration et validation »).

Quatre méthodes de mesure de la sonie avaient été comparées fin 2002 sur des sons stationnaires et non stationnaires. Cette étude, qui depuis a fait l'objet d'une publication au 7^{ème} Congrès Français d'Acoustique (Boullet, 2004) a permis de conclure que la méthode d'ajustement présentait le meilleur compromis entre la durée de l'expérience et la variabilité des résultats.

Une base de données importante de mesure de la sonie a été recueillie sur une grande variété de sons (naturels et synthétiques). L'objectif était double : 1) recueillir des informations qui nous permettent de construire un estimateur de sonie adapté aux sons impulsionnels, 2) pouvoir par la suite, comparer la sonie calculée par notre estimateur à la sonie réelle issue de la mesure.

Nous avons ensuite analysé les résultats des mesures de sonie réalisées sur des sons impulsionnels afin de comprendre plus en détail la manière dont ils sont perçus par le système auditif. Les analyses nous ont par exemple permis de tracer les lignes isosoniques et d'étudier l'intégration temporelle de la sonie de sons impulsionnels. Les résultats obtenus sur l'intégration temporelle ont été publiés pour le congrès international Internoise (Boullet et col., 2004). De plus, cette étape a permis d'établir un grand nombre de données expérimentales sur lesquelles nous avons pu étudier et comparer les différents modèles ou estimateur.

Afin de pouvoir comparer les différents modèles existants d'estimation de la sonie de sons non stationnaires, les modèles de Zwicker et Fastl (1999) et de Glasberg et Moore (2002) ont été implémentés. Ces algorithmes modélisent le fonctionnement du système auditif, en particulier en ce qui concerne les phénomènes de masquage (temporel et fréquentiel) et d'intégration. Ils permettent d'obtenir une courbe de sonie en fonction du temps mais ne permettent de calculer la sonie globale de sons non stationnaires. En revanche, plusieurs indicateurs sont proposés dans la littérature, comme le N4, N5 ou N10 (sonie dépassée 4%, 5% ou 10 % du temps).

Pour juger de la qualité des indicateurs de sonie, des indices ont été déterminés. Les quatre indices retenus sont : mean average residual (MAR), résiduel maximum (Rmax), résiduel moyen (Rmean) et écart type du résiduel (Rstd).

Nous avons ainsi déterminé les domaines d'applications et les limites des indicateurs existants ainsi que de l'ESI. Nous avons montré que deux indicateurs calculés à partir du niveau d'isotonie en fonction du temps selon le modèle de Zwicker et Fastl (1999) donnent de bons résultats avec un MAR moyen de 4 phones. Ces indicateurs, déterminés par notre équipe de recherche, sont le niveau d'isotonie maximum (Nmax) et le N30ms. Ces indicateurs semblent être mieux adaptés pour des sons non stationnaires de synthèse (MAR = 2 phones) que pour des sons non stationnaires naturels (MAR = 6 phones). L'écart maximum avec les mesures obtenu avec ces indicateurs pour calculer le niveau d'isotonie peut être de 14 ou 15 phones pour les sons impulsionnels, ce qui reste une valeur élevée par rapport aux valeurs de la dispersion des mesures (de l'ordre de 4 phones). Notons tout de même que les écarts les plus importants sont ceux concernant les trains d'impulsions, sons pour lesquels aucun modèle ne donne un niveau d'isotonie proche de celui mesuré.

Enfin, nous avons établi onze fonctions de sonie afin de construire l'ESI_{bc} en tenant compte de l'influence de la fréquence. Cette étude a aussi permis d'établir d'autres données expérimentales afin de commencer un processus de validation complété par des mesures de sonie de sons impulsionnels « naturels ».

Ainsi, nous pouvons conclure que sur plus de deux cent sons impulsionnels, l'ESI_{bc} permet d'estimer le niveau d'isotonie rapidement et avec une erreur moyenne (MAR moyen de l'ordre de 3 phones) inférieur à la dispersion des mesures (Ecart type de l'ordre de 4 phones).

Valorisation

Les travaux réalisés dans le cadre de ce projet ont conduit GENESIS à implanter, dans le logiciel d'expertise acoustique LEA qu'elle développe, un module psychoacoustique incluant une première version de l'estimateur de sonie pour des sons impulsionnels. Ce logiciel, destiné à l'analyse de signaux acoustiques et au design sonore industriel, dispose de fonctions classiques d'analyse temporelle et fréquentielle des signaux. Il propose de plus des fonctions innovantes d'analyse et de sculpture de représentations temps-fréquences qui ont été développées en collaboration avec le CNRS-LATP (Laboratoire d'analyse, topologie et probabilités). Comme le montre la figure 2, la conception s'appuie sur une collaboration représentée, pour le module psychoacoustique, par le Laboratoire de Mécanique et d'Acoustique.



Figure 2 : fenêtre « à propos » de LEA

La figure 3 donne un aperçu de l'intégration du module psychoacoustique dans LEA. On peut voir le calcul des décibels et des décibels pondérés, souvent considérés dans le monde industriel comme une première approximation de l'intensité perceptive d'un son. On distingue aussi la courbe de sonie en fonction du temps que l'implémentation du modèle de Zwicker et Fastl permet d'obtenir. On accède aussi aux attributs de la courbe de sonie en fonction du temps : N5, N10 et maximum. Enfin, une première version de notre algorithme d'estimation de la sonie d'un son impulsionnel N30ms est présente dans le module.

A court terme, le calcul de la sonie en fonction du temps par le modèle de Glasberg et Moore (2002) ainsi que le calcul de la sonie de sons impulsionnels par l'ESI_{bc} seront intégrés dans le module de psychoacoustique.

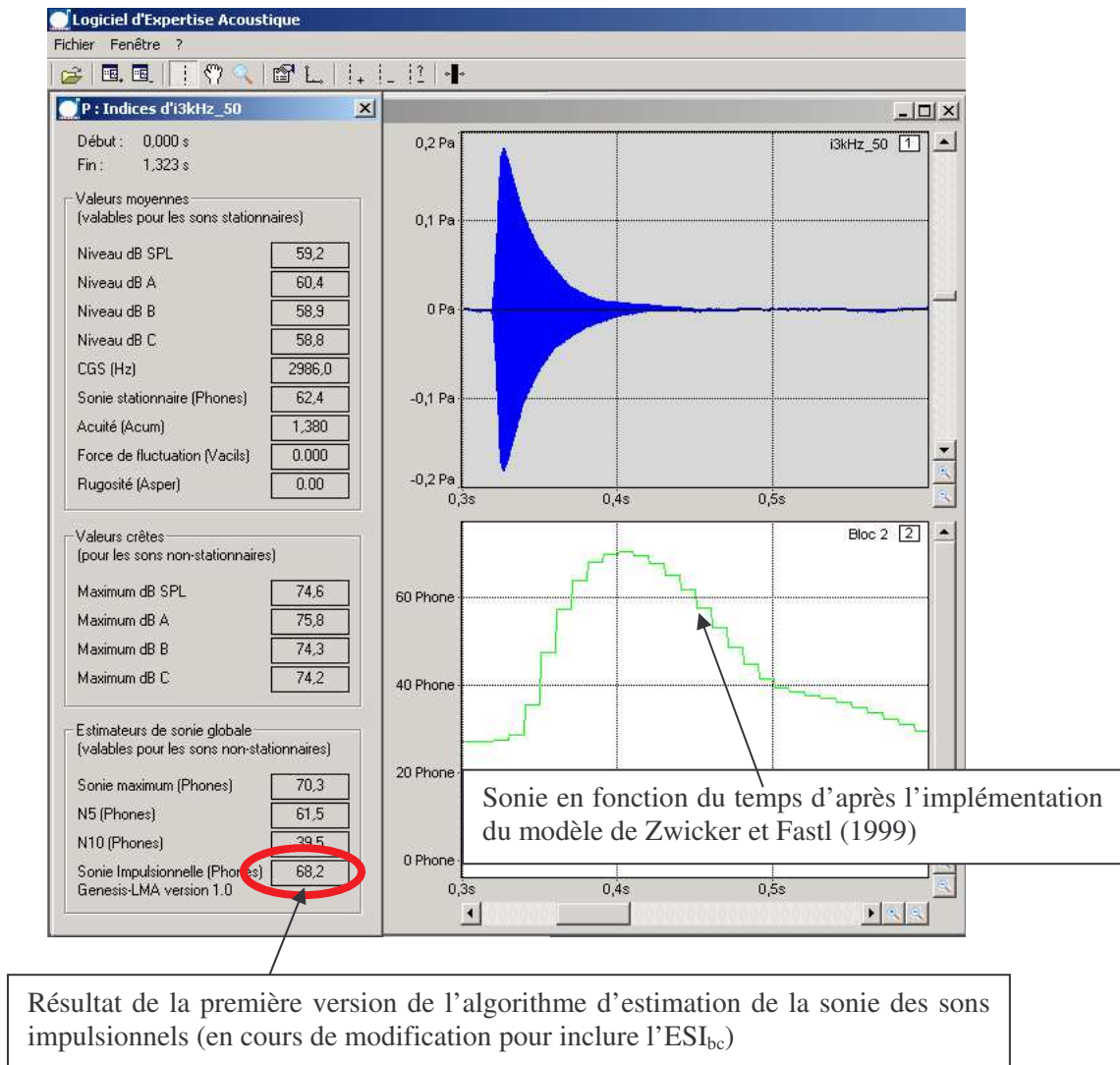


Figure 3 : aperçu du module psychoacoustique de LEA

Annexe 1 : Les modèles de sonie de sons non stationnaires

A l'heure actuelle, deux méthodes -celle de Zwicker et Fastl (1999) et de Glasberg et Moore (2002)-, permettent le calcul de la sonie de sons non stationnaires. Nous décrivons dans cette annexe le principe de chacun des deux modèles.

1- Le modèle de sonie de sons non stationnaires de Zwicker et Fastl (1999)

Le principe de base du modèle est le calcul de l'excitation de la membrane basilaire par bandes critiques, en prenant en compte diverses pondérations modélisant les filtrages de l'oreille externe et de l'oreille moyenne, le calcul du masquage fréquentiel et temporel.

Le signal acoustique est filtré pour obtenir un signal temporel dans chacune des bandes critiques (paragraphe 1-1). L'étape suivante du modèle est de calculer toutes les 2 ms dans chacune des vingt-quatre bandes critiques le niveau d'excitation. Pour cela les auteurs tiennent compte dans leur modèle de la transmission à travers l'oreille externe et l'oreille moyenne (paragraphe 1-2). La sonie de cœur est ensuite calculée à partir du niveau d'excitation et tenant compte d'une correction de seuil (paragraphe 1-3). On obtient à ce stade du modèle vingt-quatre sonies de cœur en fonction du temps. C'est alors qu'est pris en compte le masquage temporel. Ce phénomène est modélisé grâce à une analogie électrique par un quadripôle (paragraphe 1-4). Ensuite, toutes les 2 ms est calculée la densité de sonie qui est intégrée sur les vingt-quatre bandes critiques pour prendre en compte le masquage fréquentiel (paragraphe 1-5). Après un filtrage passe-bas de la sonie en fonction du temps pour tenir compte de l'intégration temporelle on obtient la sonie finale en fonction du temps.

1-1 Modélisation et calcul des filtres auditifs

La sonie est liée à la répartition de l'excitation, provoquée par un stimulus, le long de la membrane basilaire (Zwicker, 1958). Par conséquent, on représente le fonctionnement de l'oreille ou plus précisément de la membrane basilaire, comme celui d'un banc de filtres.

Ces filtres « physiologiques » sont modélisés selon Zwicker par des bandes critiques ; ce sont des bandes spectrales adjacentes dont l'unité est le bark. La largeur des bandes critiques est constante (100 Hz) en dessous de 500 Hz. Au-dessus de 500 Hz, cette largeur augmente avec la fréquence (voir tableau A-1.1).

Numéro de la bande	Fréquence centrale en Hz	Largeur de la bande critique en Hz	Fréquence de coupure supérieure en Hz
1	50	80	100
2	150	100	200
3	250	100	300
4	350	100	400
5	450	110	510
6	570	120	630
7	700	140	770
8	840	150	920
9	1000	160	1080
10	1170	190	1270
11	1370	210	1480
12	1600	240	1720
13	1850	280	2000
14	2150	320	2320
15	2500	380	2700
16	2900	450	3150
17	3400	550	3700
18	4000	700	4400
19	4800	900	5300
20	5800	1100	6400
21	7000	1300	7700
22	8500	1800	9500
23	10500	2500	12000
24	13500	3500	15500

Tableau A-1.1 : Tableau définissant les 24 bandes critiques selon Zwicker (Zwicker et Feldtkeller (1981), p. 71). Pour chaque bande est recensée la fréquence centrale (Hz), la largeur de la bande critique (Hz) et la fréquence de coupure supérieure (Hz).

1-2 Corrections de champ : transmission de l'oreille externe et de l'oreille moyenne

Le modèle de sonie prend en compte la fonction de transfert de l'oreille externe et de l'oreille moyenne en introduisant un facteur de correction (exprimé en dB/Hz à la figure A-1.1) qui est à soustraire au niveau du signal enregistré. Ce facteur dépend de la fréquence. Au-dessus de 1000 Hz, la fonction de transfert a la forme de la courbe de seuil absolu mais inversée. En dessous de 1000 Hz, Zwicker suppose que la transmission du champ libre à l'oreille interne est uniforme et de gain 1. Il suppose aussi que l'augmentation du seuil pour les basses fréquences est uniquement due au bruit interne de l'oreille.

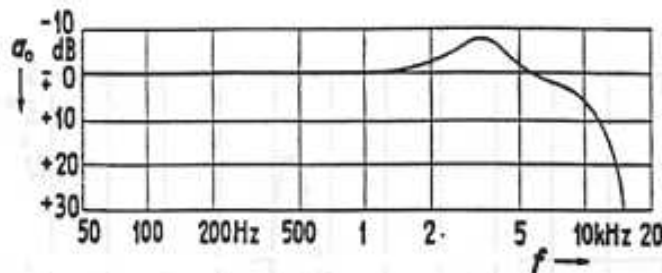


Figure A-1.1 : Variations avec la fréquence du facteur de transmission a_0 entre le champ libre et l'oreille interne (Zwicker et Feldtkeller, 1981, p. 111). En champ diffus, ce facteur est modifié et il faut introduire une autre correction.

1-3 Calcul de la sonie de coeur

La sonie de cœur est calculée à partir de l'équation ci-dessous, tirée de Zwicker et Fastl (1999):

$$\text{Sonie de coeur} = K \cdot [((0,5) + 0,5 \cdot (10^{(0,1 \cdot (Le - Les))})^{\alpha - 1})]$$

$$\text{avec } K = 0,068 \cdot 0,5^{-\alpha} \cdot [10^{(\alpha \cdot Les / 10)}]$$

L'exposant α (égal à 0,23) est celui de la droite correspondant à la fonction de sonie d'un bruit uniformément excitant (courbe du haut en tirets-points de la figure A-1.2). «Le» est le niveau d'excitation produit par le stimulus et «Les» est le niveau d'excitation au seuil dans le silence, produit par le bruit physiologique interne.

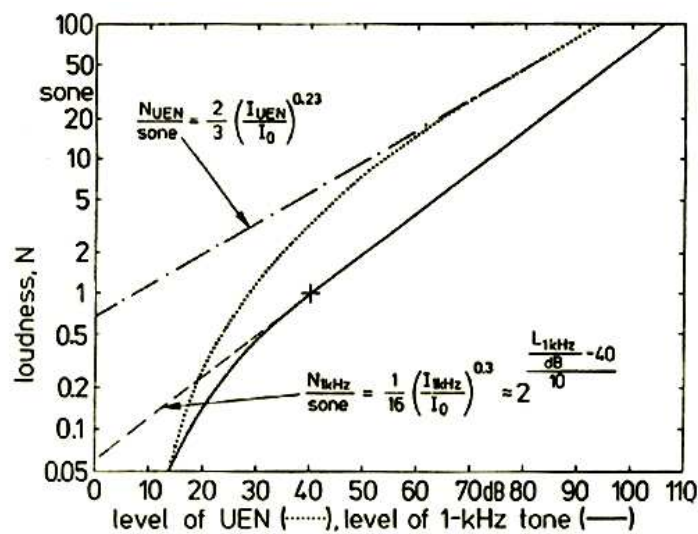


Figure A-1.2 : Fonctions de sonie d'un son pur à 1 kHz (courbe du bas, en trait plein) et d'un bruit uniformément excitant ou UEN (courbe du haut, en pointillés). Les droites approchent les fonctions de sonie, suivant une loi de puissance, d'un son pur à 1 kHz (en tirets) et d'un bruit uniformément excitant (tirets-points). La sonie est représentée en fonction du niveau de pression des signaux (Zwicker et Fastl, 1999).

De plus, si le niveau d'excitation est inférieur au niveau d'excitation correspondant au seuil d'audition absolu alors la sonie de cœur est nulle.

1-4 Calcul du masquage temporel

Un signal peut être masqué (ou difficile à détecter) s'il est précédé de près dans le temps par un autre signal (masquage postérieur). Un signal peut aussi être masqué par un son qui le suit de près dans le temps (masquage antérieur ou rétroactif).

Zwicker et Fastl utilisent un quadripôle (figure A-1.3) pour modéliser le masquage temporel postérieur en fonction de l'intensité du signal et de sa durée. En effet, selon la durée du signal, le masquage s'étendra sur une portion plus ou moins longue après la fin du signal. Le masquage temporel est modélisé par une décharge de condensateurs.

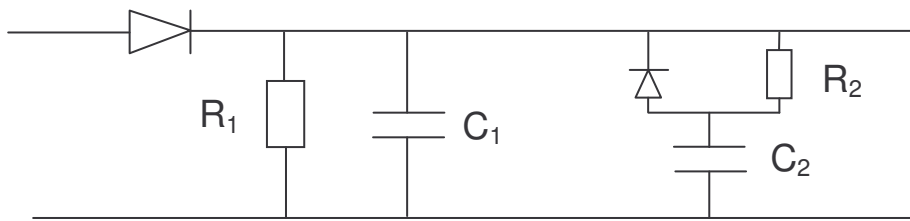


Figure A-1.3 : Circuit RC du quadripôle servant à modéliser le masquage temporel postérieur. $R_1=35\text{k}\Omega$; $C_1=0,7\ \mu\text{F}$; $R_2= 20\ \text{k}\Omega$; $C_2=1\ \mu\text{F}$ (Zwicker, 1984).

Les charges et décharges des condensateurs dépendent des tensions d'entrée et de sortie. Supposons que la tension du signal d'entrée soit un créneau. C_1 est chargé instantanément et C_2 se charge avec une constante de temps $T_2= R_2 * C_2$ égale à 20 ms. C_2 est considéré comme chargé après une durée comprise entre $3 * T_2=60\text{ms}$ (condensateur chargé à 95%) et $5 * T_2=100\ \text{ms}$ (condensateur chargé à 99%).

Lorsque le signal est interrompu, les condensateurs se déchargent selon deux cas :

- 1) Le signal d'entrée dure moins de 100 ms :

Le condensateur C_2 n'a pas eu le temps de se charger complètement. C_1 se décharge alors dans R_1 et charge C_2 par R_2 .

- 2) Le signal d'entrée dure plus de 100 ms :

Les condensateurs C_1 et C_2 sont complètement chargés, le système est en équilibre. C_1 et C_2 se déchargent alors dans R_1 . La décharge, se faisant avec une constante de temps égale à $R_1(C_1 + C_2)$, est plus lente que dans le cas (1). Par exemple, sur la figue A-1.4, on peut observer le masquage temporel occasionné par des créneaux de sons purs de 5 kHz et de durée 10 ms et 100 ms.

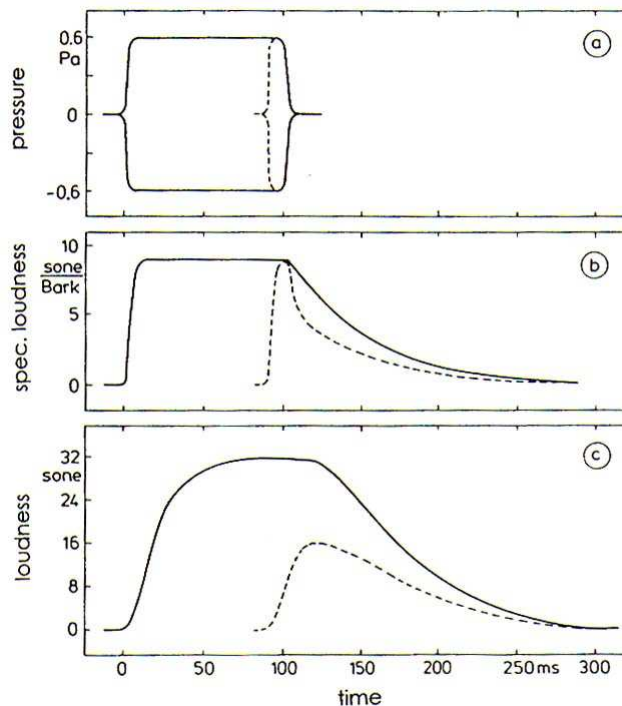


Figure A-1.4 : Etablissement de la sonie de créneaux à 5 kHz de 100 ms (trait plein) et 10 ms (pointillés). (a) Enveloppe temporelle des deux créneaux, (b) Densité de sonie correspondant aux deux créneaux dans la bande de bark 19, (c) Sonie en fonction du temps (Zwicker et Fastl 1999).

1-5 Calcul de la densité de sonie et de la sonie totale

Dans chaque bande critique et toutes les 2 ms est calculé le niveau d'excitation puis la sonie de cœur. Zwicker calcule la densité de sonie à partir du diagramme d'excitation. Selon Zwicker, le diagramme d'excitation reflète le diagramme de masquage d'un son pur masqué par un bruit à spectre étroit. Zwicker suppose en effet que les diagrammes d'excitation ont la même forme que les courbes de seuils de détection masquée. Les courbes de masquage traduisent la sélectivité de l'oreille et dépendent de la fréquence centrale et du niveau du signal d'entrée (figure A-1.5).

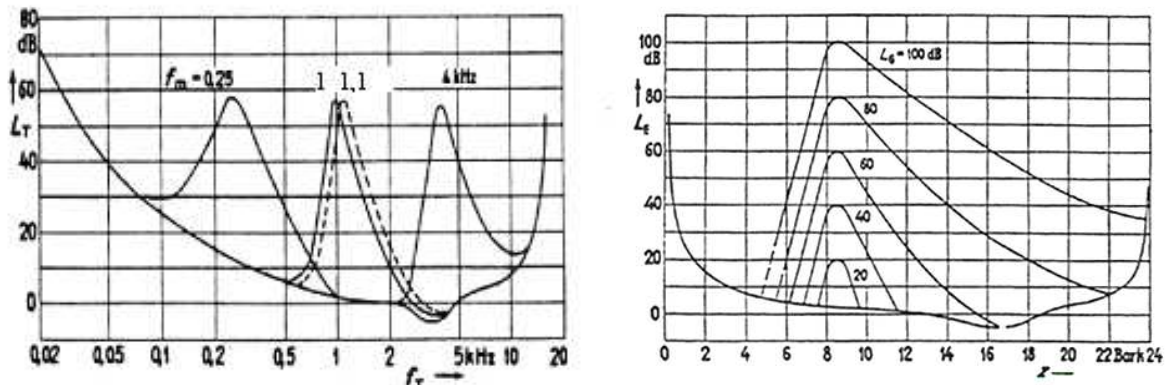


Figure A-1.5 : A gauche : courbes de masquage par des bandes de bruit à quatre fréquences centrales différentes (250 Hz, 1 kHz, 1,1 kHz et 4 kHz). Les courbes correspondent aux seuils de détection mesurés en présence de ces bandes de bruits (de largeur spectrale égale à une bande critique) émises à un niveau de pression de 60 dB ; A droite : diagramme d'excitation produit par un bruit centré sur 1 kHz et de 160 Hz de largeur spectrale. Le niveau de pression de la bande de bruit est noté L_g et vaut successivement 20, 40, 60, 80 et 100 dB (Zwicker et Feldtkeller, 1981).

L'allure du flanc supérieur est globalement conservé et dépend du niveau et de la bande critique. L'allure du flanc inférieur sera transformée en une verticale de façon à ce que l'aire sous les courbes reste la même. Le calcul de la densité de sonie est donc décomposé en un calcul, à partir du niveau d'excitation de cœur, de la sonie de cœur et de la sonie de flanc (figure A-1.6).

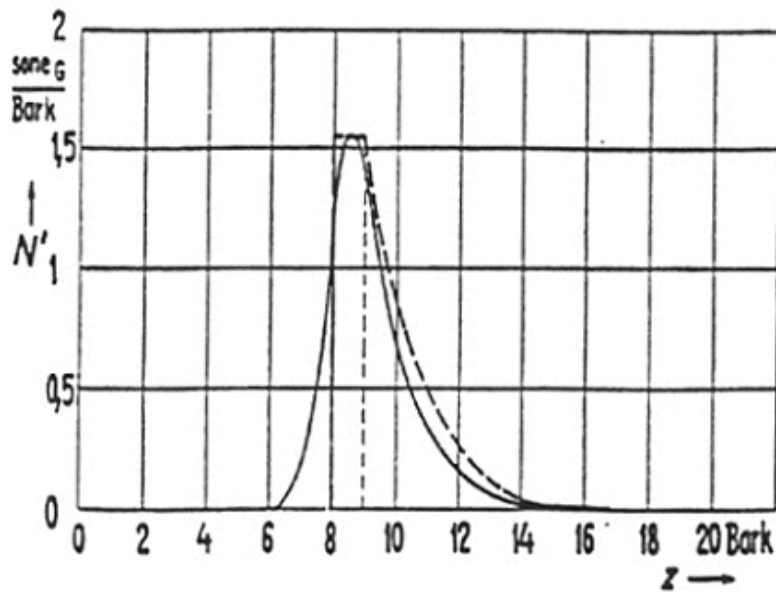


Figure A-1.6 : Courbe de densité de sonie, en fonction de la tonie en barks, d'un son pur de 60 dB à 1 kHz. La courbe en traits-tirets représente l'approximation du modèle, décomposée en une partie rectangle (sonie de cœur) et une partie courbe décroissante (sonie de flanc), utilisée pour le calcul de la sonie globale (Zwicker et Feldtkeller, 1981).

Le calcul de la densité de sonie est décomposée en deux partie : le calcul de la sonie de cœur (voir paragraphe 1-3) et de la calcul de la sonie de flanc.

Le calcul de la sonie de flanc est assez laborieux, car il n'existe pas de relation mathématique pour les courbes de seuils masqués. Les courbes des sonies de flanc des diagrammes patrons (figure A-1.7), utilisés autrefois, ont été approximées pour les programmes de calcul par des lignes brisées (figure A-1.8). Les segments de droite de la ligne brisée sont délimités par des lignes horizontales. Ces lignes horizontales correspondent à différents échelons de densité de sonie dépendant du niveau d'excitation. En effet, plus le niveau d'excitation est élevé, et donc la densité de sonie, plus la pente du premier segment de la ligne brisée est raide.

La densité de sonie est alors la succession dans chaque bande critique des sonies de cœur et/ou des sonies de flanc. En effet, si la sonie de cœur d'une bande critique n se situe au-dessous de la sonie de flanc de la bande critique précédente $n-1$, alors la sonie de cœur de la bande critique n est masquée. Ce sera alors la sonie de flanc de la bande critique précédente $n-1$ qui sera prise en compte pour déterminer la densité de sonie dans la bande critique n (voir figure A-1.7).

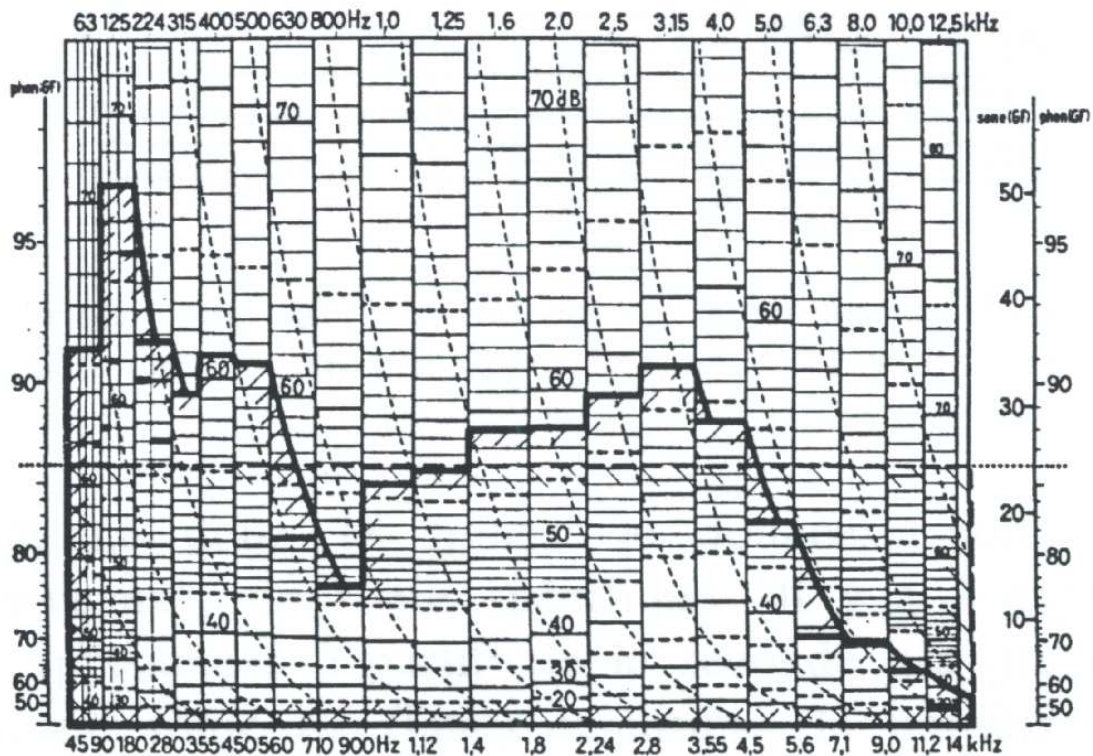


Figure A-1.7 : Exemple de distribution des niveaux par bandes critiques en fonction de la fréquence d'un bruit de machine. La sonie de cœur calculée dans chaque bande critique est représentée par les traits épais horizontaux. La sonie de flanc, modélisant le masquage fréquentiel, est représentée sur la même figure par les courbes descendantes en trait épais déterminées d'après les diagrammes patrons (courbes en pointillés). La sonie globale est repérée par la ligne horizontale en traits-tirets. Le niveau d'isonie correspondant est de 86 phones.

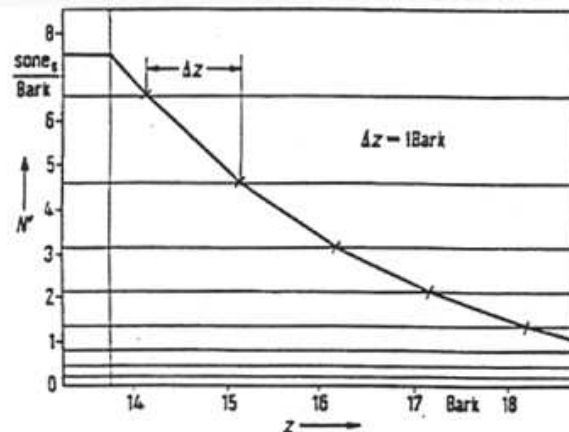


Figure A-1.8 : Exemple de ligne brisée utilisée pour approximer la sonie de flanc induite par une bande située à 2 kHz ayant une sonie de cœur de 7,5 sones/Bark.

Pour calculer la sonie totale, la densité de sonie est alors intégrée sur les 24 bandes critiques (somme de sonie) : c'est l'aire représentée sous la ligne en trait plein épais de la figure A-1.7. Cette étape est répétée toutes les 2 ms afin de calculer une sonie en fonction du temps, ici sans tenir compte de l'intégration temporelle.

Le modèle de Zwicker et Fastl (1999) décrit ci-dessus, permet de calculer la sonie en fonction du temps (par exemple voir la figure A-1.4 (c)), par contre on ne sait pas comment les auditeurs jugent la sonie

globale de sons variant dans le temps. Zwicker et Fastl (1999) préconisent le calcul des indices statistiques N10, N5 ou N4 (sonie dépassée 10, 5 ou 4% du temps) pour estimer la sonie globale d'un son non stationnaire.

2- Le modèle de sonie de sons non stationnaires de Glasberg et Moore (2002)

Glasberg et Moore (2002) ont aussi développé un modèle de sonie pour des sons variant dans le temps. Le modèle s'affranchit de la transmission à travers l'oreille externe et l'oreille moyenne en utilisant un filtre à réponse impulsionnelle finie (FIR). La fonction de transfert du filtre (FIR) simulant la transmission du champ libre à l'oreille interne est donnée sur la figure A-1.9.

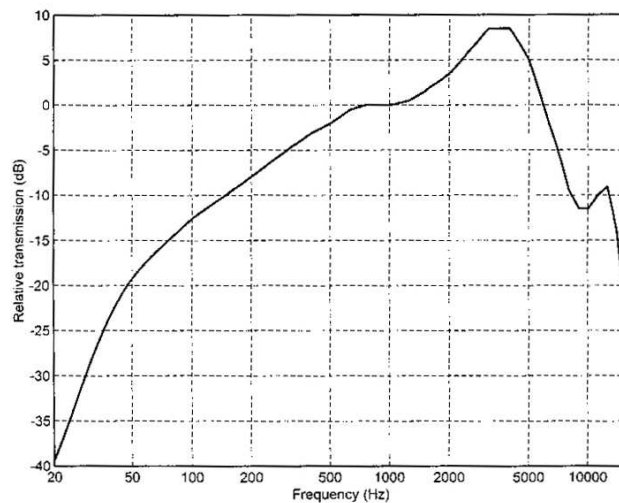


Figure A-1.9 : Fonction de transfert du filtre numérique à réponse impulsionnelle finie utilisée pour simuler les effets de l'oreille externe et de l'oreille moyenne. Le gain est calé pour être de 0 dB à 1000 Hz (Glasberg et Moore, 2002).

La cochlée est censée fonctionner comme un banc de filtres passe-bande dont les fréquences centrales sont étalées de 50 à 15000 Hz. La largeur des filtres augmente avec leurs fréquences centrales (voir figure A-1.10). Par exemple, pour un filtre centré sur 100 Hz, la largeur de bande rectangulaire équivalente (ERB) est de 35 Hz. Pour une fréquence centrale de 10 kHz, la largeur de bande rectangulaire est de 1100Hz.

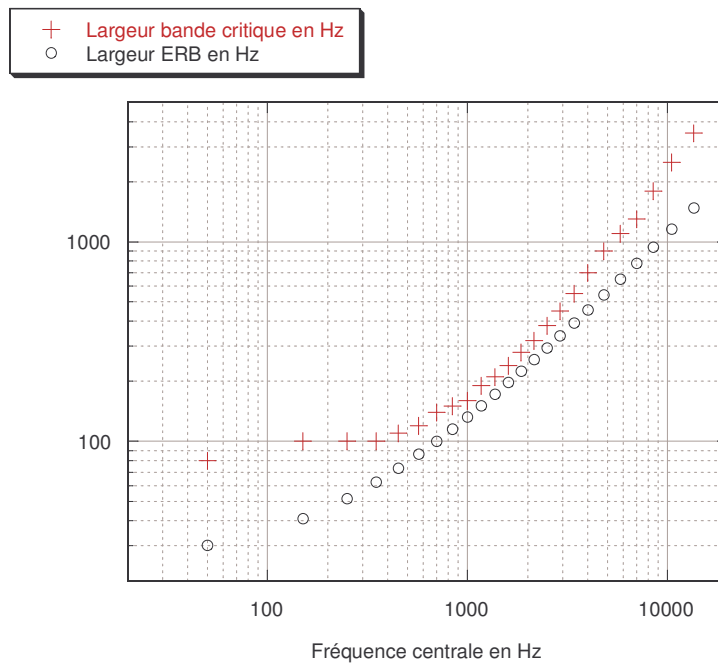


Figure A-1.10 : Largeur des bandes critiques (croix) et des ERB (cercles) en fonction de la fréquence centrale en Hz. Les valeurs des largeurs des bandes critiques sont tirées d'un tableau donné par Zwicker (1961). La largeur des ERB en Hz est égale à $24,7(4,37f+1)$ où f est la fréquence centrale du filtre exprimée en kHz (Moore, 1997).

La forme des filtres dépend du niveau. En effet, les pentes des filtres à basse fréquence deviennent moins raides quand le niveau augmente. La courbe donnant l'amplitude de sortie de chaque filtre, pour un son donné, est censée modéliser le diagramme d'excitation.

La méthode utilisée par Moore pour calculer le diagramme d'excitation est basée sur une analyse temps fréquence. Pour obtenir une résolution à basse fréquence comparable à celle du système auditif, l'analyse se fait sur des segments relativement longs (64 ms). Cependant pour les hautes fréquences les fenêtres d'analyse seront plus courtes (2 ms). Ainsi, 6 FFT sont effectuées en parallèle dont les durées sont 64, 32, 16, 8, 4, et 2 ms pour des intervalles de fréquences qui sont respectivement 20-80 Hz, 80-500 Hz, 500-1250 Hz, 1250-2540 Hz, 2540-4050 Hz et 4050-15000 Hz. Le diagramme d'excitation est calculé à partir du spectre toutes les 1 ms, il y a donc recouvrement.

Le diagramme d'excitation est calculé à partir de la sortie des filtres auditifs centrés sur les fréquences composant le son. Prenons par exemple, un son pur excitant à 1 kHz. Le diagramme supérieur de la figure A-1.11 montre les courbes de réponse de filtres centrés dans le voisinage de 1 kHz. Le trait vertical en pointillé représente la fréquence du son pur excitant. Sur le diagramme inférieur, les abscisses correspondent aux fréquences centrales des filtres voisins (cf. diagramme supérieur), et les ordonnées représentent la réponse à 1 kHz de ces filtres. Le diagramme d'excitation est ainsi constitué par la ligne qui joint les points, pour chaque fréquence centrale des filtres voisins.

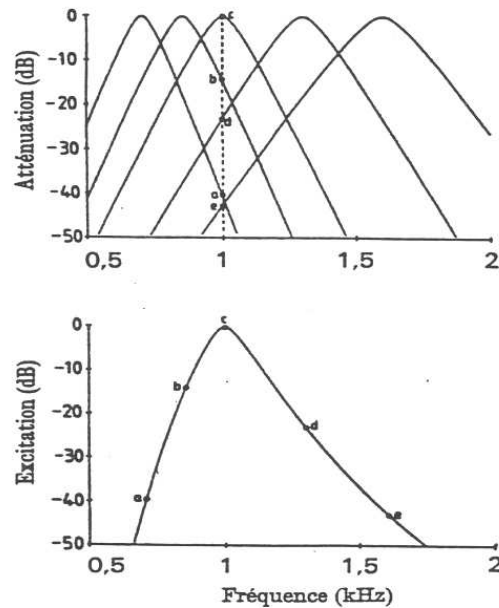


Figure A-1.11 : Principe de la détermination du diagramme d'excitation d'un son pur à 1 kHz selon le modèle de Moore.

Pour un son complexe, le diagramme d'excitation est obtenu en répétant l'opération décrite ci-dessus pour toutes les fréquences du signal. Le diagramme d'excitation final est la somme des excitations partielles.

L'étape suivante du modèle est de calculer ce que Glasberg et Moore (2002) appellent la sonie « instantanée ». Les auteurs supposent que la sonie instantanée n'est pas quelque chose de conscient dans la perception. Elle correspondrait à l'activité totale dans le nerf auditif mesurée sur un intervalle de temps très court, de l'ordre de la milliseconde. Le calcul de la sonie instantanée se fait à partir du diagramme d'excitation, de la même manière que dans le modèle stationnaire (Moore et coll., 1997). Le diagramme d'excitation est transformé en sonie spécifique et l'aire sous le diagramme de sonie donne la sonie instantanée. Ce calcul vaut pour une écoute monaurale. Pour une écoute binaurale, la sonie instantanée est sommée sur les 2 voies pour donner la sonie instantanée globale.

Glasberg et Moore (2002) calculent ensuite la sonie à court terme à partir de la sonie instantanée. La sonie instantanée est lissée, par un filtrage passe-bas, pour tenir compte de l'intégration temporelle (la constante de temps n'est pas donnée par l'auteur mais elle semble être de l'ordre de 100 ms) et du masquage temporel (décroissance de la sonie après la fin du stimulus). La sonie à court terme représente la sonie perçue à chaque instant. Les auteurs préconisent de prendre le maximum de la sonie à court terme comme valeur de la sonie globale d'un son court.

A partir de la sonie à court terme, la sonie à long terme peut être calculée en utilisant aussi une sorte d'intégration temporelle. La sonie à long terme traduirait le fait que la sonie globale d'un son est gardée en mémoire après la fin de ce dernier pendant quelques secondes et en l'absence de nouveau stimulus.

ANNEXE 2 : Méthodes de mesure de la sonie et du niveau d'isophonie

Cette annexe a pour but de présenter la méthodologie et les résultats en complément du document principal sur la comparaison des méthodes de mesures du niveau d'isophonie pour des sons stationnaires, des sons impulsionnels et des trains d'impulsions.

1- Méthode

1-1 Stimuli

Les sons ont été classés en trois groupes, décrits dans les tableaux A-2.1, A-2.2 et A-2.3. Chaque groupe représente une catégorie de sons. Nous avons étudié neuf sons stationnaires, dix sons impulsionnels et dix trains d'impulsions. La moitié des signaux constituant notre banque d'échantillons sonores a été enregistrée et l'autre moitié synthétisée. Les enregistrements ont été effectués dans la salle anéchoïque du laboratoire.

a) Sons stationnaires

Les sons stationnaires ont été synthétisés à partir de sons purs et d'une bande de bruit centrée sur 1 kHz (lignes 1 à 3 du tableau A-2.1) ou enregistrés (six dernières lignes du tableau A-2.1).

Nom du son	Description	Niveau en dB SPL	Durée totale
StaS1	bande de bruit centrée sur 1kHz de largeur 40Hz	60,5	1 s
StaS2	Son pur 200 Hz	77,9	1 s
StaS3	Son pur 4000 Hz	70,9	1 s
StaE1	Cailloux versés	59,5	1 s
StaE2	Moto au ralenti	61,6	1 s
StaE3	Perceuse	59,3	1 s
StaE4	Vélo en roue libre	62,3	1 s
StaE5	Note jouée à la flûte	66,8	1 s
StaE6	Moto au ralenti	41,7	1 s

Tableau A-2.1 : Description des neuf sons stationnaires utilisés dans l'étude.

b) Sons impulsionnels

Les sons impulsionnels que nous avons utilisés présentent tous la même forme, indiquée à la figure 2 du document principal, c'est-à-dire un temps de montée rapide (< 5 ms), pas de palier, et un temps de descente plus ou moins long (de 3 à 71 ms). Les impulsions ont été soit synthétisées à partir d'un bruit blanc et d'un son pur de 1 kHz, soit enregistrées dans une salle anéchoïque.

Nom du son	Description	Niveau crête en dB SPL	Durée totale
ImpS1	Sinus de 1 kHz	89,9	95 ms
ImpS2	bruit blanc	84,6	285 ms
ImpS3	bruit blanc	79,2	40 ms
ImpS4	bruit blanc	69,6	100 ms
ImpS5	bruit blanc	79,6	285 ms
ImpE1	Choc entre 2 pierres	77,7	30 ms
ImpE2	Claquement de doigt dans la bouche	83,5	45 ms
ImpE3	Claquement de langue	70,0	10 ms
ImpE4	Bouchon de champagne	76,0	100 ms
ImpE5	jouet clic-clac	88,2	3 ms

Tableau A-2.2 : Description des dix sons impulsionnels utilisés dans l'étude.

c) *Trains d'impulsions*

Les trains d'impulsions de synthèse, notés TImpS sont construits à partir d'une impulsion répétée (ImpE1) à son identique à une fréquence de n impulsions par seconde (quatre impulsions par seconde pour le son TImpS1, six pour le son TImpS2, dix pour les sons TImpS3 et TImpS5 et vingt pour les sons TImpS4 et TImpS6). Trois sons de l'environnement (brise-roche hydraulique, coups de marteau sur du fer et coups de marteau sur du béton) ont été enregistrés. Dans le tableau A-2.3 sont présentés les différents sons étudiés avec leur nom, leur niveau crête, leur description et leur durée.

Le son TImpE1 est constitué de vingt-cinq coups de brise-roche hydraulique (à la fréquence de huit par seconde). Le son TImpE4 est constitué du même nombre de coups et avec la même fréquence de répétition, mais il est restitué à un autre niveau. Dans le son TImpE2 deux impulsions sont présentées avec une fréquence d'une impulsion par seconde. Cinq impulsions constituent le son TImpE3 dont la fréquence de répétition est de trois impulsions par secondes.

Nom du son	Description	Niveau en dB SPL	Durée totale
TImpS1	4 impulsions / s	62,7	1 s
TImpS2	6 impulsions / s	67,7	2 s
TImpS3	10 impulsions / s	82,7	2 s
TImpS4	20 impulsions / s	67,7	2,5 s
TImpS5	10 impulsions / s	59,7	2 s
TImpS6	20 impulsions / s	77,7	2,5 s
TImpE1	Brise roche hydraulique	75	3,8 s
TImpE2	Coups de marteau sur fer	84,5	1,5 s
TImpE3	Coups de marteau sur béton	65	2,3 s
TImpE4	Brise roche hydraulique	65	3,8 s

Tableau A-2.3 : Description des dix trains d'impulsions utilisés dans l'étude.

1-2 Auditeurs

Les tests ont été réalisés par quatorze auditeurs (cinq femmes et neuf hommes) âgés de 25 à 58 ans dans la salle anéchoïque du laboratoire. La moitié des auditeurs n'avaient pas d'expérience particulière en matière de tests psychoacoustiques. Tous les auditeurs avaient une audition normale.

1-3 Dispositif expérimental

Le principe du montage expérimental pour les tests de mesure de la sonie était le même pour toutes les méthodes testées. Les fichiers sons étaient convertis en signal analogique à l'aide d'une carte son Oros. Le niveau des sons variait à l'aide d'un atténuateur programmable Wilsonics (model PAT1). Les sons étaient restitués par une enceinte Génélec (type 1031A) dans la salle anéchoïque du laboratoire.

1-4 Procédures

Dans ce paragraphe, nous décrivons les tâches des auditeurs lors des mesures du niveau d'isophonie. Chaque test est divisé en trois parties, chacune consacrée à un type de son. Ainsi, chaque auditeur a passé neuf tests (trois méthodes et trois types de sons). Tous les tests ont été réalisés dans la salle anéchoïque du laboratoire.

Avant de continuer, il est important de définir deux termes, « son test » et « son de comparaison », qui

seront utiles dans la description des tests psychoacoustiques mis en place à l'exception de la méthode d'estimation de grandeur. Le son test est le son dont on veut connaître la sonie. Le son de comparaison est celui dont on connaît a priori la sonie, et dont le niveau peut être modifié. En pratique, nous avons choisi une bande de bruit centrée sur 1 kHz, de largeur 120 Hz, donc inférieure à la largeur d'une bande critique (160 Hz) et d'un ERB (130 Hz) à 1 kHz. Lorsqu'il y a égalité des sonies, le niveau du son de comparaison correspond au niveau d'isotonie du son test. La durée du son de comparaison est de 1 seconde pour les sons stationnaires et les trains d'impulsions. Pour les sons impulsionnels, nous avons choisi un son de comparaison de 500 ms car il est plus facile d'ajuster le son de comparaison lorsque la durée de celui-ci est plus courte (Pedersen et coll., 1977).

a) Méthode d'estimation de grandeur sans référence (notée EST)

La méthode d'estimation de grandeur sans référence consiste à demander aux auditeurs d'estimer la sonie d'un son en donnant un nombre (réel positif) de son choix proportionnel à la sonie. Le sujet se fixe alors sa propre échelle, souvent liée à la sonie du premier son présenté.

Plusieurs estimations d'un même son, présenté aléatoirement dans une séquence de plusieurs sons, peuvent être demandées. Cela permet de vérifier la cohérence des réponses. Le fait de présenter les sons dans un ordre aléatoire différent pour chaque sujet permet d'atténuer un biais dû à l'effet du son qui précède le son jugé. C'est-à-dire que si le son n-1 est plus fort que le son n, le sujet aura tendance à juger la sonie du son n plus grande que si le son n-1 était moins fort.

Dans le test que nous avons réalisé, les sons, présentés dans un ordre aléatoire différent pour chaque auditeur, étaient répétés 3 fois chacun. L'auditeur estimait la sonie de chaque son en donnant un nombre de son choix proportionnel à la sonie. Pour chaque série de sons, un son de sonie connue (son pur de 1kHz à 80 dB SPL soit 16 sonies) était présent pour pouvoir normaliser les estimations de sonie des quatorze auditeurs.

Ce test durait environ 3 minutes pour 10 sons.

b) Méthode d'ajustement (notée AJUST)

La méthode d'ajustement consiste à présenter en alternance un son de comparaison et le son dont on cherche à mesurer la sonie (son test). L'auditeur doit ajuster le niveau du son de comparaison, à l'aide d'un potentiomètre par exemple, de manière à ce qu'il ait la même sonie que le son test. Le son de comparaison est fréquemment un son pur de 1 kHz ou une bande de bruit centrée sur 1 kHz de largeur inférieure à 160 Hz car leur niveau en dB SPL est égal au niveau d'isotonie en phones.

Dans notre test, l'auditeur entendait successivement le son test et le son de comparaison. Un silence de 300 ms séparait les deux sons. Chaque paire de sons était séparée de 400 ms. La tâche de l'auditeur consistait à ajuster le niveau du son de comparaison de manière à ce qu'il ait la même sonie que le son test. L'auditeur pouvait faire varier, à l'aide d'un boîtier de commande, le niveau du son de comparaison de plus ou moins 5, 2 ou 1 dB. La paire de sons était présentée sans arrêt jusqu'à ce que le sujet ait terminé l'égalisation. Une nouvelle paire lui était alors présentée. Les différents sons étaient présentés dans un ordre aléatoire, différent pour chaque auditeur. Le niveau du son de comparaison, au départ du test, était fixé aléatoirement au-dessus (environ + 15 dB) ou au-dessous (environ -15 dB) du niveau d'isotonie probable du son test (des tests ayant été faits au préalable pour avoir une première évaluation des niveaux). Ceci évite un biais dû à une approche ascendante ou descendante systématique du niveau d'isotonie du son test.

Le même test était répété, mais le son de comparaison était cette fois présenté en premier. Réaliser le test d'égalisation dans les deux sens permet de compenser le biais qui est de juger le plus souvent le son

présenté en second plus fort.

Ce test durait environ 20 minutes pour 10 sons pour obtenir le niveau d'isotonie final.

c) Méthode adaptative à choix forcé sur deux intervalles de type 2down-1up et 2up-1down (notée 2I - 2AFC)

Lors d'une mesure du niveau d'isotonie par une méthode adaptative, l'auditeur entend le son test suivi du son de comparaison ou l'inverse, avec une probabilité égale. La tâche est alors d'indiquer lequel des deux sons est le plus fort. La variation du niveau du son de comparaison dépend de la réponse de l'auditeur lors de l'essai précédent. Une méthode adaptative se caractérise par des inversions successives de variation du niveau du son de comparaison, variations dites up-down dans la littérature de langue anglaise. L'un des critères d'inversion le plus fréquemment utilisé pour la méthode adaptative (critère « 2down-1up ») consiste à réduire le niveau du son de comparaison après deux réponses consécutives identiques (par exemple, son de comparaison plus fort) et à le relever après que l'auditeur ait répondu que le son test est plus fort. Les pas de réduction ou d'augmentation du niveau sont au départ assez grand (en général 5 dB) puis diminuent (2 dB) après la deuxième inversion du niveau.

Dans ce test, l'auditeur entendait le son test suivi du son de comparaison ou l'inverse, avec la même probabilité. Les deux sons étaient séparés par 300 ms de silence. La tâche de l'auditeur était d'indiquer lequel des deux sons était le plus fort. Le niveau du son de comparaison, au départ du test, était présenté environ 15 dB au-dessus ou au-dessous du niveau d'isotonie probable du son test (respectivement pour 2down-1up et 2up-1down).

La procédure 2down-1up fait converger le niveau du son de comparaison vers la valeur pour laquelle il est jugé plus fort que le son test dans 71 % des essais (Levitt, 1971). Ainsi, la procédure fournit une estimation du niveau pour lequel le son de comparaison est perçu « juste plus fort » que le son test (Schlauch et Wier, 1987). Une séquence s'arrêtait à la sixième inversion et on calculait le seuil en faisant la moyenne des niveaux aux quatre dernières inversions. Pour chaque son, trois mesures de seuil étaient réalisées. Le niveau d'isotonie « juste plus fort » était alors la moyenne des trois valeurs des seuils obtenus. Cette opération a été effectuée pour tous les sons, présentés dans un ordre aléatoire.

Le principe de la procédure 1down-2up est le même que celui de la méthode adaptative 2 down-1up. Cette méthode permet d'obtenir le niveau d'isotonie « juste moins fort ».

En faisant la moyenne des valeurs du niveau d'isotonie obtenues avec les méthodes 2down-1up et 1down-2up on obtient le niveau d'isotonie. Obtenir le niveau d'isotonie par cette méthode permet d'obtenir un résultat plus stable qu'avec la méthode adaptative 1down-1up (Gelfand, 1998).

Ce test durait environ deux fois 40 minutes (en moyenne 4 minutes par son) pour 10 sons.

d) Méthode adaptative à poursuites multiples (notée Mult 2I-2AFC)

Le principe de cette méthode est le même que celui de la méthode adaptative à choix forcé sur deux intervalles du type 2down-1up et 2up-1down mais au lieu de suivre une seule séquence, plusieurs sons sont mélangés et donc plusieurs séquences sont alors simultanées. Pour les sons stationnaires ainsi que pour les trains d'impulsions nous avons mélangé quatre sons. Pour les sons impulsionnels, leurs durées étant plus courtes, nous avons mélangé cinq sons. Nous avons alors fait deux tests de cinq sons dont un son était commun aux deux tests.

A chaque essai, la paire de sons était choisie aléatoirement parmi les séquences qui n'étaient pas déjà finies. La fin d'une séquence avait lieu après la sixième inversion. Nous calculions alors une moyenne sur les quatre dernières inversions. Cette opération avait lieu trois fois comme pour la méthode 2AFC-

2I.

De même que pour la méthode adaptative, la moyenne des valeurs du niveau d'isotonie obtenues avec les méthodes 2 down – 1 up et 1 down – 2 up permet d'obtenir le niveau d'isotonie.

Une série de quatre sons se terminait au bout de 40 minutes environ. Il fallait donc 80 minutes (2down-1up et 2up-1down) pour déterminer le niveau d'isotonie pour 4 sons.

Le tableau A-2.4 permet une vision plus globale des conditions expérimentales en détaillant les quatre méthodes de mesure de la sonie.

Méthodes	Estimation	Ajustement 1	Ajustement 2	2down 1up	2up 1down	Multitracking 2down 1up	Multitracking 2up 1down
Son test (T)	9 sons stationnaires, 10 sons impulsionnels et 10 trains d'impulsions					4 sons stationnaires, 9 sons impulsionnels et 4 trains d'impulsions	
Ordre de présentation de T	Aléatoirement, différent pour chaque auditeur						
Son de comparaison (C)	Pas de son de comparaison	Spectre : Bande de bruit centrée à 1 kHz de largeur 120 Hz < une bande critique					
		Durée : 1 s pour les sons stationnaires et les trains d'impulsions 500 ms pour les sons impulsionnels					
Ordre de présentation de T et C	T seul	T & C	C & T	T & C ou C & T aléatoirement			
Silence entre T et C	500 ms						
Niveaux de départ de C par rapport à T		Aléatoirement 15 dB au-dessus ou dessous		15 dB au-dessus	15 dB au-dessous	15 dB au-dessus	15 dB au-dessous
Tâche	Donner un nombre de son choix proportionnel à la sonie	Ajuster le niveau de C pour que T et C aient la même sonie		Indiquer lequel des 2 sons est le plus fort			
Durée du test en minutes pour 10 sons	3	10	10	40	40	40	40
Sujets	14 (5 femmes et 9 hommes)						
Age	25 à 58 ans						

Tableau A-2.4 : Conditions expérimentales pour les différentes méthodes de mesure de la sonie

1-5 Analyse des données

a) Estimation directe

Nous avons calculé la moyenne géométrique des trois estimations données par chacun des quatorze sujets pour chaque son. Un son pur de 1000 Hz à un niveau de 80 dB SPL, de sonie connue (16 sonies) faisait partie des échantillons sonores et servait de référence.

La sonie moyenne pour chacun des auditeurs et de chaque son était normalisée par rapport à la sonie du son de référence :

$$\text{sonie du son test} = \text{estimation donnée au son test} * 16 / \text{estimation donnée au son de référence}$$

La sonie de chaque son est la moyenne géométrique des valeurs obtenues pour les quatorze auditeurs. Nous supposons notre groupe d'auditeurs « normal » du point de vue de la fonction de sonie normale (exposant de la pente de la fonction de sonie égal à 0,6). On peut donc déduire le niveau d'isotonie (L) à partir de la sonie (S) en utilisant la relation :

$$L = 40 + 10 \log_2(S)$$

b) Ajustement

Théoriquement, le niveau du son de comparaison donne directement le niveau d'isotonie du son test en phones. Le niveau d'isotonie final est la moyenne des valeurs obtenues lors des ajustements dans les deux sens.

c) Adaptative à choix forcé sur deux intervalles de type 2down-1up et 2up-1down

Le niveau d'isotonie est obtenu par la moyenne des deux valeurs obtenues par les procédures 2 down-1 up et 2 up-1 down.

d) Adaptative à poursuites multiples

Le niveau d'isotonie est obtenu par la moyenne des deux valeurs obtenues par les procédures 2 down-1 up et 2up-1 down. Le principe est le même que pour la méthode adaptative à choix forcé sur deux intervalles de type 2down-1up et 2up-1down.

2- Résultats

Pour choisir la méthode la mieux adaptée à la mesure d'un niveau d'isotonie, il est nécessaire de considérer, non seulement la cohérence des résultats et la durée de la procédure, mais aussi la précision. Les niveaux d'isotonie obtenus ne peuvent pas être comparés à des résultats de la littérature ou être considérés comme juste ou faux. La cohérence des résultats sera alors relative. Un des points importants dans cette étude est donc la comparaison des résultats obtenus avec les quatre méthodes de mesure. Un autre aspect intéressant, abordé plus loin dans le paragraphe 2-2, est l'étude de la dispersion des résultats, traduite ici par le calcul des écarts types, pour déterminer la précision des méthodes.

2-1 Cohérence des méthodes

Les figures A-2.1, A-2.2 et A-2.3 montrent les niveaux d'isotonie obtenus par les différentes méthodes psychoacoustiques pour les sons décrits précédemment. On notera que la méthode adaptative à poursuites multiples (Mult 2I-2AFC) n'a été utilisée que sur quatre sons stationnaires, neuf sons impulsionnels et quatre trains d'impulsions. Cette méthode impose une mesure très longue et c'est pour cela que nous avons volontairement limité le nombre de sons. Chaque point indique le niveau d'isotonie moyen des quatorze auditeurs, pour un son et une méthode donnés. Les valeurs des niveaux d'isotonie, pour un même son, ont été légèrement décalées suivant l'axe des abscisses pour permettre une meilleure lisibilité des écarts types. Les écarts types, représentés par les barres verticales, sont calculés sur les mesures du niveau d'isotonie final des quatorze auditeurs.

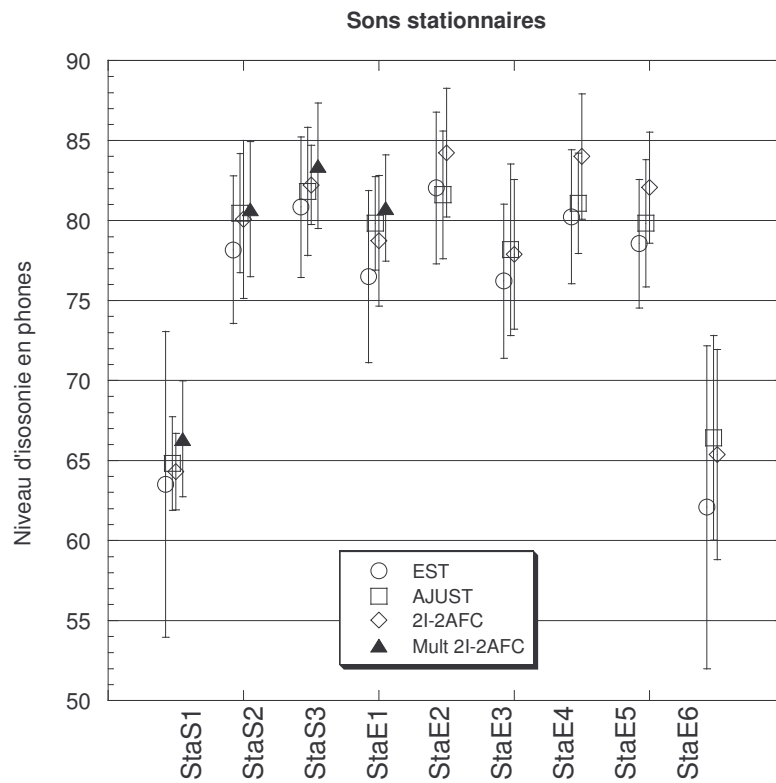


Figure A-2.1 : Niveaux d'isophonie des sons stationnaires obtenus par les différentes procédures. Les symboles sont légèrement décalés horizontalement pour une meilleure visibilité. Les barres verticales représentent les écarts types à la moyenne des mesures.

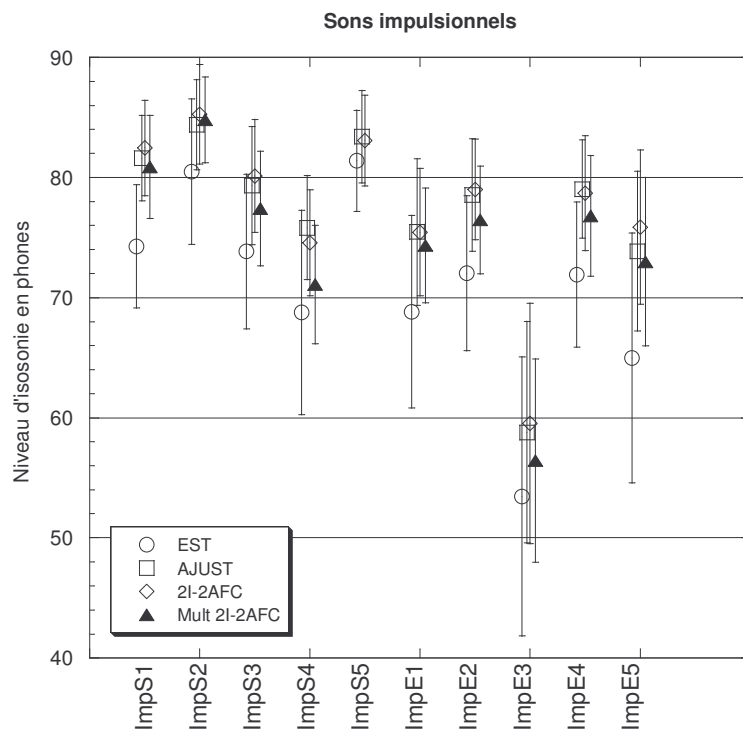


Figure A-2.2 : Niveaux d'isophonie des sons impulsionnels obtenus par les différentes procédures. Les symboles sont légèrement décalés horizontalement pour une meilleure visibilité. Les barres verticales représentent les écarts types à la moyenne des mesures.

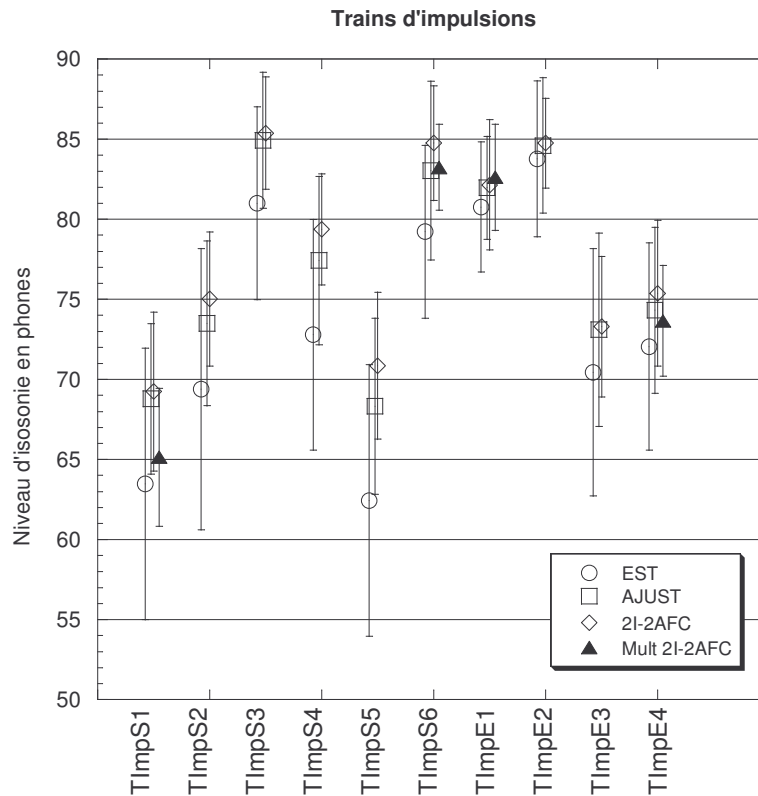


Figure A-2.3 : Niveaux d’isotonie des trains d’impulsions obtenus par les différentes procédures. Les symboles sont légèrement décalés horizontalement pour une meilleure visibilité. Les barres verticales représentent les écarts types à la moyenne des mesures.

Les valeurs du niveau d’isotonie obtenues par la méthode d’estimation de grandeur sont toujours plus faibles que celles obtenues par les trois autres méthodes. En effet, les niveaux d’isotonie mesurés par la méthode d’estimation directe sont en moyenne 4 phones en dessous des mesures faites avec les trois autres méthodes. Les différences des niveaux d’isotonie entre les trois autres méthodes (AJUST, 2I-2AFC et Mult 2I-2AFC) sont en moyenne égales à 1,5 phones.

On peut aussi observer que les différences entre les méthodes sont plus marquées pour les sons impulsionnels. En effet, la moyenne des différences entre les quatre méthodes est de 3,7 phones pour les sons impulsionnels, alors qu’elle est de 2,4 phones pour les trains d’impulsions et de 1,8 phones pour les sons stationnaires. Cependant, les différences des niveaux d’isotonie obtenus avec les quatre méthodes ne sont pas toujours significatives.

Globalement, il y a un effet significatif de la méthode sur le niveau d’isotonie ($F(3)=11,08$; $p<0,001$). Le tableau A-2.5 représente les résultats de l’analyse de la variance (ANOVA) pour étudier l’influence de la méthode en fonction du type de son. Il ressort de cette analyse statistique que, pour chaque type de son, il y a une influence de la méthode sur le niveau d’isotonie.

Sons stationnaires				
	dl	F	p	R ²
Influence de la méthode	3*	3,71*	0,0192*	1,86*
Sons impulsionnels				
	DI	F	p	R ²
Influence de la méthode	3*	12,06*	0 *	8,93*
Trains d’impulsions				
	DI	F	p	R ²
Influence de la méthode	3*	5,51*	0,003*	3,41*

Tableau A-2.5 : Analyse de la variance du niveau d’isotonie (variable dépendante). La méthode de mesure du niveau d’isotonie et le son sont les variables indépendantes (facteurs fixés).

Nous avons aussi déterminé, pour chaque type de son, les méthodes donnant des niveaux d'isophonie statistiquement différents. Les résultats de l'analyse post-hoc (Test de Scheffé) montre que, pour les sons stationnaires, la seule différence significative est entre EST et Mult 2I-2AFC. Pour les sons impulsionnels, EST est significativement différente des trois autres méthodes. Enfin, pour les trains d'impulsions, on a trouvé uniquement une différence significative entre EST et les deux méthodes AJUST et 2I-2AFC.

Pour résumer, lorsqu'il y a un effet significatif de la méthode sur la mesure du niveau d'isophonie, il est principalement dû à la différence entre la méthode d'estimation directe et les trois autres méthodes. Il est important de rappeler ici que cette méthode permet de mesurer la sonie. Pour connaître le niveau d'isophonie, nous avons utilisé la relation $\text{phones} = 40 + 10 \log_2(\text{sones})$. Cette relation part de l'hypothèse que la pente de la fonction de sonie à 1 000 Hz est de 0,6. Or la pente de la fonction de sonie du groupe d'auditeurs est de 0,49. L'écart entre le niveau d'isophonie obtenu avec la méthode d'estimation directe et les trois peut s'expliquer par cette approximation.

Mais globalement, les quatre méthodes de mesure du niveau d'isophonie sont cohérentes. Il est donc important, à ce stade de l'étude, de définir avec quelle précision les méthodes permettent de mesurer le niveau d'isophonie.

2-2 Précision des méthodes

Nous avons représenté sur les figures A-2.4, A-2.5 et A-2.6 les valeurs des écarts types obtenues par les quatre méthodes de mesure de la sonie pour les sons stationnaires, les sons impulsionnels et les trains d'impulsions respectivement. Chaque point indique la valeur de l'écart type à la moyenne des niveaux d'isophonie des quatorze auditeurs, pour une méthode donnée en fonction du niveau d'isophonie du son.

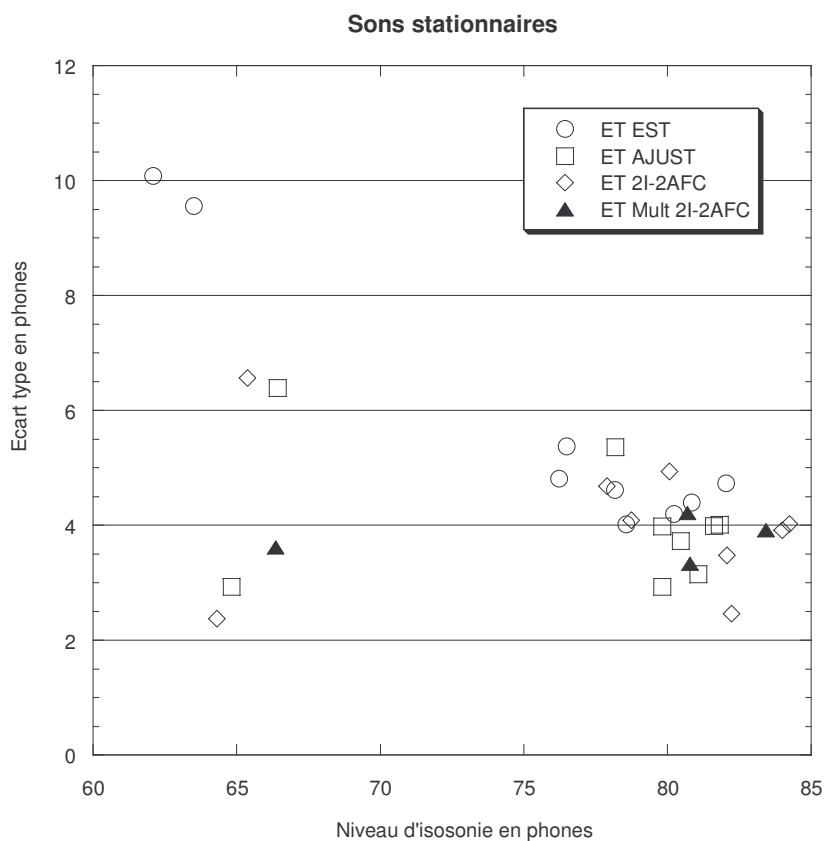


Figure A-2.4 : Ecart type en fonction des méthodes de mesure de la sonie pour neuf sons stationnaires.

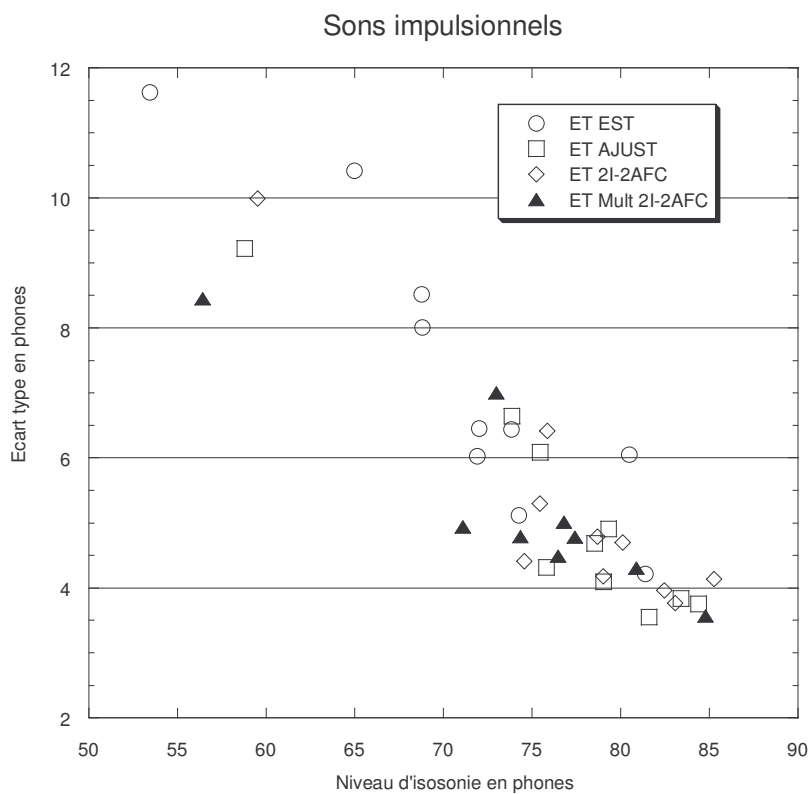


Figure A-2.5 : Ecart type en fonction des méthodes de mesure de la sonie pour dix sons impulsionnels.

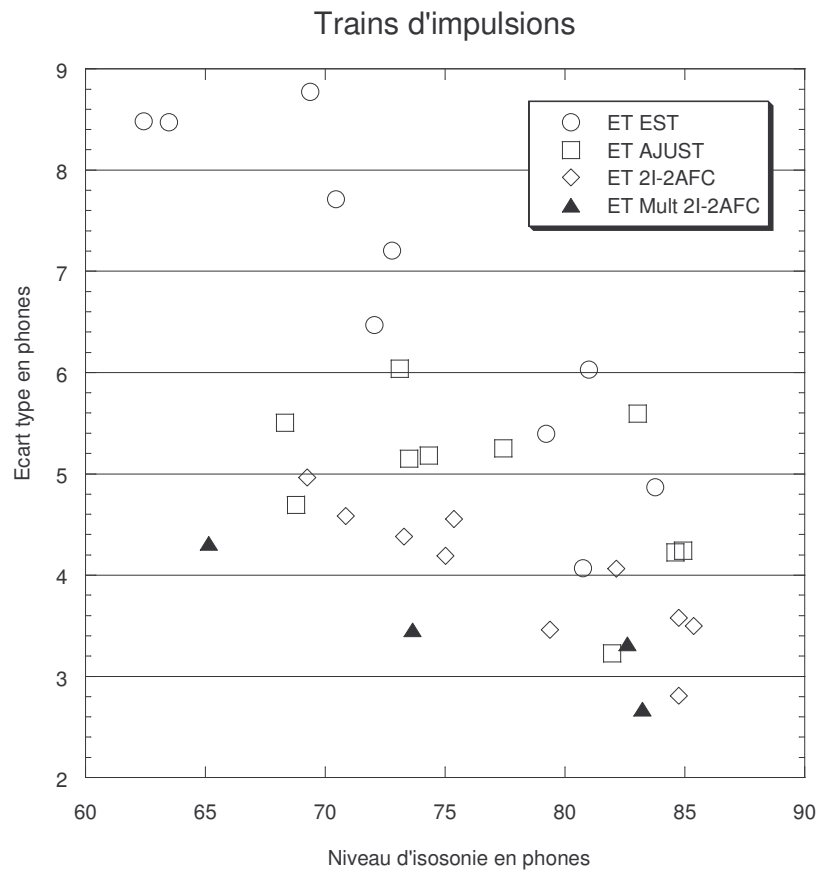


Figure A-2.6 : Ecart type en fonction des méthodes de mesure de la sonie pour 4 trains d'impulsions.

Globalement, on observe que les écarts types sont plus élevés pour les sons de niveaux faibles que pour ceux de niveaux forts. Il faut noter qu'à bas niveau les seuils différentiels de l'oreille sont plus élevés qu'à fort niveau. Ainsi, à fort niveau, on perçoit des différences de niveau plus faibles qu'à bas niveau, ce qui explique une dispersion plus faible des mesures du niveau d'isotonie. On peut aussi observer que la dispersion des données dépend des méthodes. D'après les figures A-2.4, A-2.5 et A-2.6, on remarque que la méthode EST présente les écarts types les plus élevés.

Le tableau A-2.6 représente les valeurs moyennes des écarts types pour chaque méthode et chaque type de son. Les valeurs de ce tableau permettent d'étudier la précision de chaque méthode en fonction du type de son et de façon générale.

Écarts types en phones	Estimation	Ajustement	Multitracking	Adaptative
Sons stationnaires	5,75	4,05	3,8	4,1
Sons impulsionnels	7,6	5,3	5,7	5,3
Trains d'impulsions	6,75	4,9	3,44	4
Moyenne	6,7	4,7	4,3	4,5

Tableau A-2.6: Moyenne des écarts types en fonction de chaque méthode et de chaque type de son. La dernière ligne est la moyenne des écarts types sur tous les sons en fonction des quatre méthodes.

En ce qui concerne les sons stationnaires et les trains d'impulsions, la méthode Mult 2I-2AFC donne la variabilité la plus faible. Pour les sons impulsionnels, c'est la méthode AJUST qui permet de mesurer le niveau d'isotonie avec la plus faible dispersion.

De manière générale, la méthode d'estimation directe produit les écarts types les plus élevés. A moins de 0,5 phones près, les trois autres méthodes ont en moyenne les mêmes écarts types quel que soit le type de son. Un autre point intéressant est que les écarts types sont en moyenne inférieurs à 5 phones, sauf pour la méthode d'estimation. Plus précisément, pour les méthodes AJUST, 2I-2AFC et Mult 2I-2AFC les écarts types sont de l'ordre de 4,5 phones.

Si l'on considère, pour déterminer la méthode la mieux adaptée, non seulement la précision mais aussi la durée, la méthode d'ajustement, une des plus rapides (20 minutes pour dix sons) présente le meilleur compromis. Elle permet de mesurer directement le niveau d'isophonie en phones.

Au vu de ces résultats on ne peut pas s'attendre à avoir une précision du modèle de sonie inférieure à 6 phones pour des niveaux d'isophonie faibles (< 65 phones). On peut, par contre, pour des niveaux d'isophonie supérieurs à 65 phones espérer une précision de l'ordre de 4 phones pour le modèle de sonie.

2-3 Fiabilité de la méthode d'ajustement

Dans le paragraphe précédent, nous avons déterminé la méthode d'ajustement comme celle présentant le meilleur compromis entre précision et durée. Nous avons donc voulu connaître la différence intra-individuelle obtenue avec cette méthode. Pour cela, nous avons réalisé un test avec six stimuli (2 sons impulsionnels, 2 sons stationnaires et 2 trains d'impulsions) issus des expériences précédentes. Le protocole et le dispositif expérimental étaient les mêmes que précédemment. Le test, réalisé par sept auditeurs, était répété quatre fois.

La différence intra-individuelle que nous obtenons est ici d'environ 2 phones.

ANNEXE 3 : Base données expérimentales

Nous détaillons ici les expériences que nous avons réalisées afin d'établir une base de données expérimentales pour étudier les limites des modèles de sonie et proposer un outil mieux adapté à la mesure de la sonie de sons impulsionnels.

1- Mesure des lignes isosoniques

1-1 Stimuli

Les sons utilisés pour cette expérience ont été synthétisés. Leur enveloppe est caractérisée par :

- une attaque linéaire en amplitude sur 2 ms,
- pas de palier,
- une décroissance exponentielle coupée à -60 dB du niveau crête (durée mise pour passer de 90 à 10 % de l'amplitude maximum égale à 20 ms).

Le choix de ces durées est basé sur l'ordre de grandeur des durées des sons en créneau étudiés par les auteurs précédemment cités. Leur durée totale, égale à la durée de l'attaque plus la durée de la décroissance, est de 57 ms.

La structure fine de remplissage des enveloppes était composée de sons purs : 150 Hz, 450 Hz, 700 Hz, 1000 Hz, 2150 Hz, 2900 Hz, 4000 Hz, 5800 Hz, 8500 Hz et 10500 Hz. Ces fréquences correspondent à des fréquences centrales de bandes critiques. Le choix de ces fréquences a été fait pour tenir compte des variations caractéristiques des lignes isosoniques : par exemple, autour de 3 kHz la sensibilité de l'oreille est plus importante (fréquence de résonance de l'oreille externe), puis elle diminue jusqu'à 8 kHz, augmente à nouveau jusqu'à 10 kHz et diminue au-delà de 10 kHz. Le tableau A-3.1 récapitule les caractéristiques des signaux que nous étudions.

Nom	Temps de montée (ms)	Temps de descente (ms)	Durée totale des impulsions synthétisées (ms)	Fréquence (Hz)
sin150	2	20	57	150
sin450	2	20	57	450
sin700	2	20	57	700
sin1k	2	20	57	1000
sin2k15	2	20	57	2150
sin2k9	2	20	57	2900
sin4k	2	20	57	4000
sin5k8	2	20	57	5800
sin8k5	2	20	57	8500
sin10k5	2	20	57	10500

Tableau A-3.1 : Tableau récapitulatif des sons sélectionnés pour la mise en œuvre des tests

Nous nous sommes intéressés à trois lignes isosoniques (50, 60 et 75 phones) pour des sons impulsionnels de 57 ms dont l'enveloppe temporelle est celle que nous avons déterminée précédemment (figure 3 du document principal) pour des sons impulsionnels « naturels ».

1-2 Dispositif expérimental

Le signal numérique est converti en signal analogique grâce à une carte son Oros. Le signal est ensuite atténué par un atténuateur programmable Wilsonics (model PATT) avant d'être restitué par une enceinte amplifiée Génélec (type 1031A) située dans la salle anéchoïque du laboratoire.

La mesure des niveaux est faite par un microphone Bruel & Kjaer (type 2669) situé à la place de la tête de l'auditeur en l'absence de celui-ci.

1-3 Auditeurs

L'ensemble des tests a été réalisé par quinze auditeurs âgés de 21 à 40 ans. Les auditeurs ne présentaient aucune déficience auditive.

1-4 Procédure

Le principe de la procédure était ici d'ajuster le niveau du son impulsionnel pour qu'il ait la même sonie qu'un son stationnaire de comparaison dont le niveau était fixé à 50, 60 ou 75 dB SPL. Le niveau d'isotonie du son impulsionnel a été mesuré par une méthode d'ajustement dans les deux sens de présentation (son impulsionnel suivi du son de comparaison et son de comparaison suivi du son impulsionnel). Le son de comparaison était une bande de bruit centrée sur 1 kHz, de largeur 160 Hz. Son niveau en dB SPL correspond au niveau d'isotonie en phones. Pour chaque son, six tests ont été réalisés (trois niveaux d'isotonie \times deux ordres de présentation du son test et du son de comparaison). Les niveaux reportés pour les sons impulsionnels sont les niveaux crêtes.

Le niveau crête du son impulsionnel, au départ de la session, était présenté à l'auditeur environ 20 dB au-dessus ou au-dessous du niveau d'isotonie du son de comparaison (des tests ayant été faits au préalable pour avoir une première estimation des niveaux d'égalité subjective). Le son de comparaison, de 500 ms, et le son impulsionnel étaient séparés de 300 ms. Chaque paire était séparée de 400 ms. La paire de sons était présentée à l'auditeur sans arrêt, il pouvait alors faire varier l'intensité du son impulsionnel par pas de plus ou moins 5, 2 ou 1 dB, et validait son jugement quand il estimait égales les sonies des deux sons. L'ordre de présentation des paires à égaliser (dix au total) était aléatoire et différent pour chaque auditeur. Pour un niveau donné du son de comparaison et pour chaque son impulsionnel, on fait la moyenne des deux niveaux crêtes ajustés (correspondant aux deux ordres de présentation) par chaque auditeur. On calcule ensuite la moyenne de ces valeurs sur les quinze auditeurs. On obtient ainsi le niveau crête que doit avoir un son impulsionnel pour avoir la même sonie que le son de comparaison soit un niveau d'isotonie de 50, 60 et 75 phones.

1-5 Résultats

Nous avons tracé sur la figure A-3.1 les trois courbes isotoniques à 75, 60 et 50 phones ainsi que les écarts types. Globalement, nous pouvons remarquer un certain parallélisme entre les trois courbes avec un "pic" relatif à 1 kHz et un minimum vers 3 kHz. Les courbes présentent aussi un point d'inflexion vers 6 kHz.

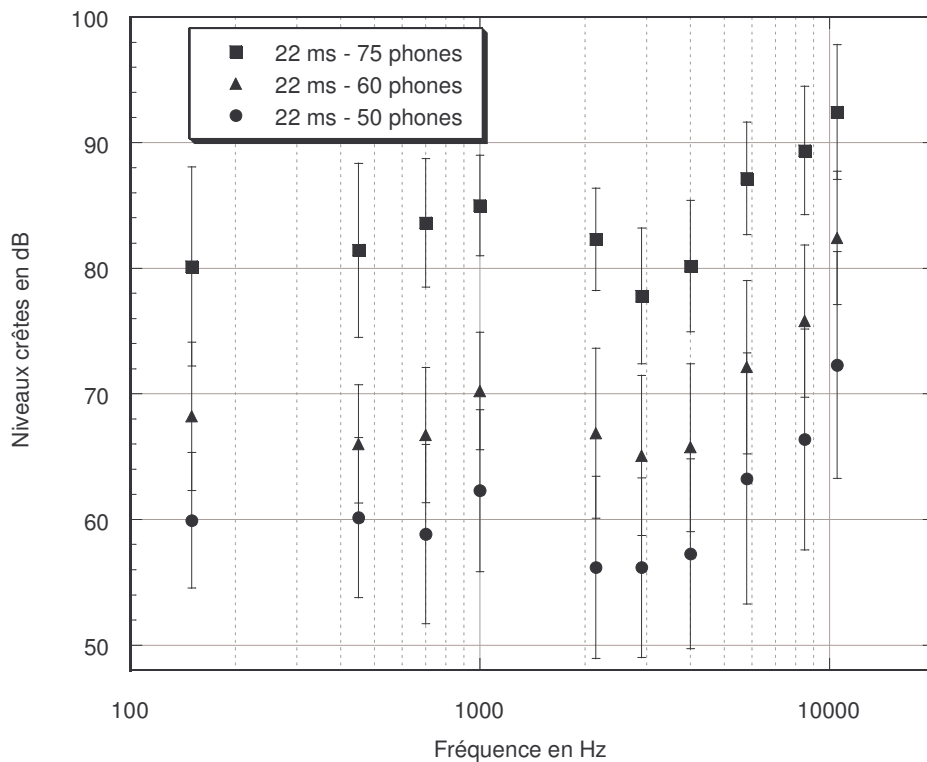


Figure A-3.1 : Lignes isononiques pour des sons impulsionnels à 75 phons (carrés), 60 phons (triangles) et 50 phons (ronds).

2- Mesure de l'intégration temporelle

2-1 Première expérience : Influence de la fréquence et du décours temporel

Cette première expérience consiste à mesurer le niveau d'isononie de sons de courte durée en forme de créneau et en forme d'impulsion, de portées différentes.

a) Stimuli

Les signaux étudiés ont été construits avec une enveloppe temporelle en forme de créneau (voir figure A-3.2) ou d'impulsion (voir figure A-3.3). La structure fine de remplissage des enveloppes était composée de sons purs (125 Hz, 350 Hz, 1 kHz, 3 kHz et 8 kHz) et d'un bruit blanc. Tous les niveaux crêtes des signaux ont été fixés à 72 dB SPL lors de la restitution.

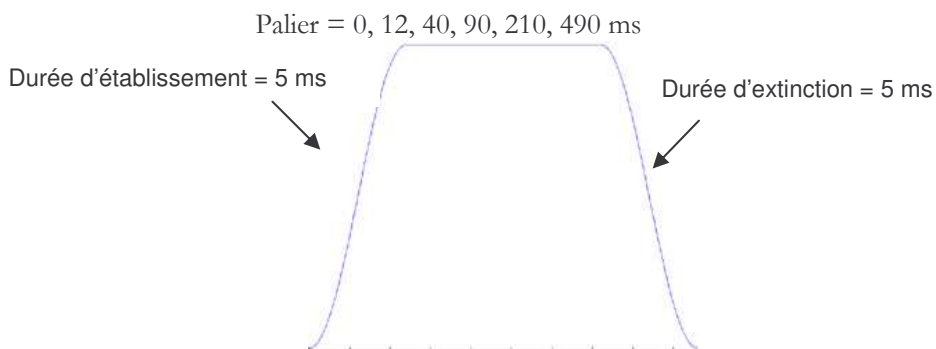


Figure A-3.2 : Modèle d'enveloppe pour la construction des créneaux.

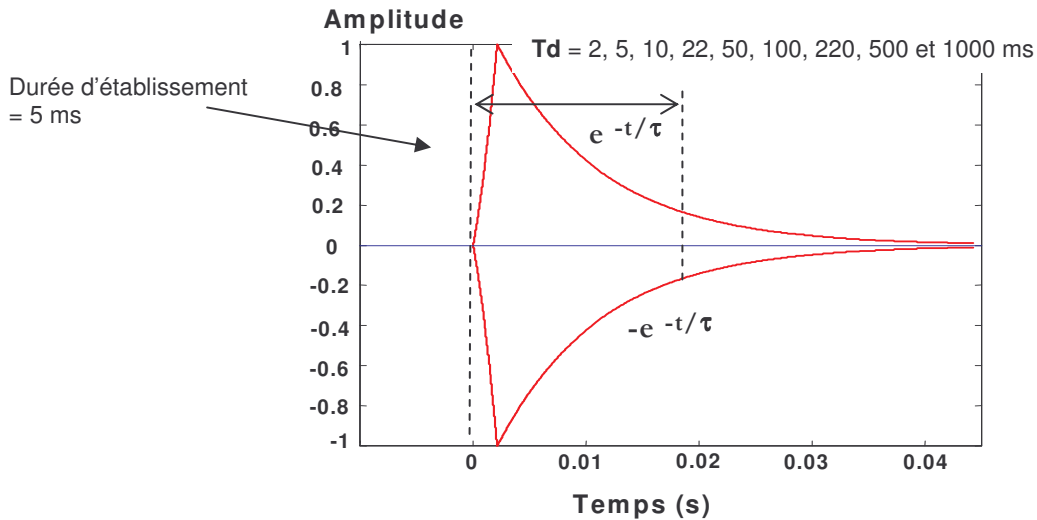


Figure A-3.3 : Modèle d'enveloppe pour la construction des impulsions.

Durée totale des sons synthétisés (ms)
Créneaux
10
22
50
100
220
500
Impulsions
10
20
35
70
155
310
670
1520
3155

Tableau A-3.2 : Durées en ms des signaux synthétisés.

L'établissement et l'extinction des stimuli en créneaux ont pour forme des demi-fenêtres de hanning et durent 5 ms. Pour les créneaux, la durée totale du son est la somme des durées d'établissement, du palier et d'extinction. Les durées des créneaux sont donc de 10, 22, 50, 100, 220 et 500 ms. Le choix des durées a été effectué afin d'obtenir une répartition régulière sur une échelle logarithmique.

On caractérise ici une impulsion par une montée rapide et linéaire de l'amplitude, pas de palier et une extinction de forme exponentielle décroissante ($e^{-\alpha(t-T_d)}$) ; où T_d est le temps de descente : durée mise pour passer de 90% à 10 % de l'amplitude). La durée d'établissement est de 5 ms et les temps de descente sont : 2, 5, 10, 22, 50, 100, 220, 500 et 1000 ms. La construction des signaux, avec la décroissance exponentielle, entraîne une durée totale infinie. Le signal était alors coupé à -60 dB du niveau crête. La durée totale des impulsions est donnée dans le tableau A-3.2.

b) Auditeurs

Les tests ont été réalisés par dix auditeurs, six femmes et quatre hommes. Les auditeurs étaient des étudiants et des personnels du laboratoire âgés de 24 à 42 ans et avaient tous un audiogramme normal.

c) Procédure

Les mesures du niveau d'isophonie ont été faites avec une méthode d'ajustement (voir annexe 2). Elles se pratiquaient en deux sessions. Dans la première, le son test était présenté avant le son de comparaison. Dans la deuxième session, le son test suivait le son de comparaison. Le son de comparaison était une bande de bruit centrée sur 1 kHz, de largeur 120 Hz et de durée 500 ms. Un silence de 300 ms séparait les deux sons. Deux paires étaient séparées de 400 ms. La tâche de l'auditeur consistait à ajuster le niveau du son de comparaison de manière à ce qu'il ait la même sonie que le son test. L'auditeur pouvait faire varier, à l'aide d'un boîtier de commande, le niveau du son de comparaison de plus ou moins 5, 2 ou 1 dB. La paire de sons était présentée sans arrêt jusqu'à ce que le sujet ait jugé égales les sonies des deux sons. Les différents sons étaient présentés dans un ordre aléatoire, différent pour chaque auditeur. Le niveau du son de comparaison, au départ du test, était présenté aléatoirement au-dessus (environ + 10 dB) ou au-dessous (environ -15 dB) du niveau d'isophonie probable du son test (des tests ayant été faits au préalable pour avoir une première estimation des niveaux d'isophonie).

Le niveau d'isophonie retenu est la moyenne des deux jugements des dix auditeurs.

d) Dispositif expérimental

Les fichiers sons étaient convertis en signal analogique à l'aide d'une carte son Oros. Le niveau des sons était réglé à l'aide d'un atténuateur programmable Wilsonics (model PAT1). Les sons étaient émis par une enceinte amplifiée Génélec 1031A.

e) Résultats

Les niveaux d'isophonie en fonction de la durée sont représentés sur les figures A-3.4 et A-3.5 pour les créneaux et pour les impulsions respectivement. La fréquence est le paramètre des graphes.

On observe sur la figure A-3.4 que pour les créneaux, le niveau d'isophonie augmente rapidement jusqu'à 100 ms (en moyenne 3 phones par doublement de durée) puis, au-delà de cette durée il continue à augmenter, mais plus doucement (en moyenne 1,5 phones par doublement de durée).

Sur la figure A-3.5, on peut observer pour toutes les impulsions une dépendance linéaire du niveau d'isophonie avec le logarithme de la durée. En effet, la sonie augmente globalement d'environ 1,5 phones (en moyenne sur toutes les fréquences) par doublement de durée.

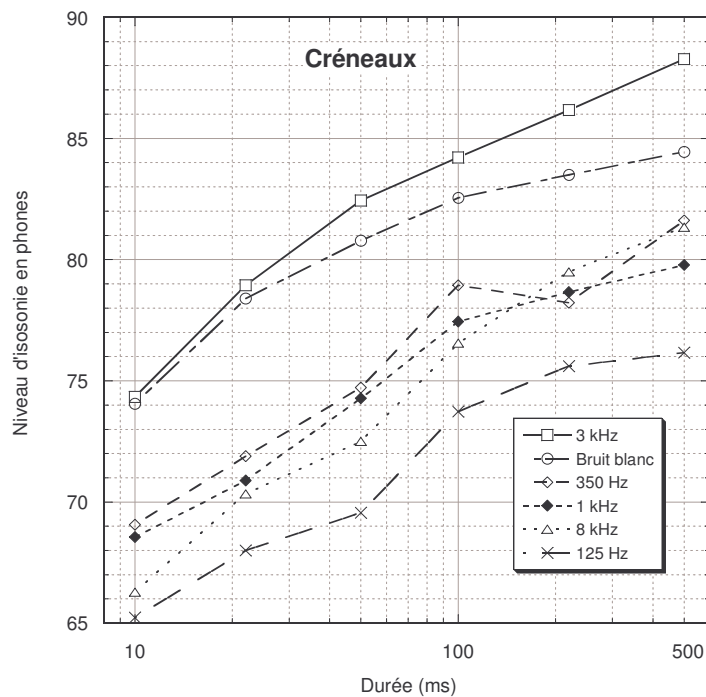


Figure A-3.4 : Niveaux d'isotonie des créneaux (bruit blanc, 125 Hz, 350 Hz, 1 kHz, 3 kHz et 8 kHz) en fonction de la durée.

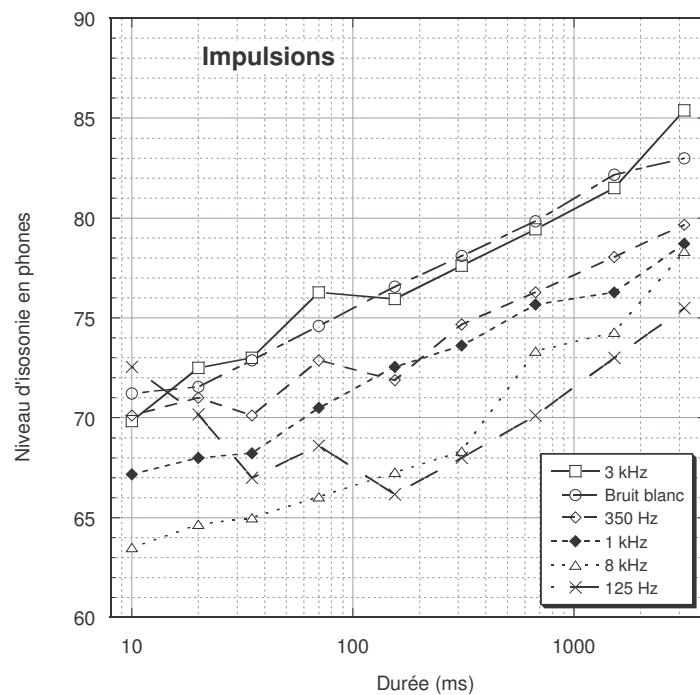


Figure A-3.5 : Niveaux d'isotonie des impulsions (bruit blanc, 125 Hz, 350 Hz, 1 kHz, 3 kHz et 8 kHz) en fonction de la durée.

Globalement, on observe bien une augmentation du niveau d'isotonie avec la durée. Cependant, ceci n'est pas confirmé pour les impulsions dont la structure fine est de 125 Hz. En effet, pour ces stimuli le niveau d'isotonie diminue avec la durée pour les durées inférieures à 155 ms. Ces stimuli, pour des durées très courtes, ont un étalement spectral de plus de 400 Hz. L'énergie du signal est donc répartie

sur plusieurs bandes critiques, ce qui entraîne une augmentation du niveau d'isophonie par rapport aux stimuli de durées plus longues mais pour lesquels l'énergie est répartie dans une seule bande critique.

Cette première expérience n'a pas permis de mettre en évidence une durée critique, durée à partir de laquelle le niveau d'isophonie ne varie plus. Pour les créneaux, on observe un début de palier à partir de 100 ms (figure A-3.4). En effet, à partir de cette durée, le niveau d'isophonie évolue peu avec la durée (environ 1,5 phons par doublement de durée) mais n'est pas stabilisé. En ce qui concerne les impulsions, le niveau d'isophonie augmente constamment avec la durée d'environ 1,5 phons par doublement de durée (figure A-3.5).

Nous nous sommes intéressés aussi à la dispersion des mesures : les écarts types augmentent lorsque la durée des sons diminue et lorsque les fréquences du son test et du son de comparaison sont éloignées. Lorsque la durée des sons diminue, la différence qualitative entre le son test et le son de comparaison est plus marquée. De même, pour les sons plus longs mais de fréquences différentes, la différence de timbre entre son test et son de comparaison entraîne une augmentation de la variabilité. Lorsque la durée des sons diminue, le niveau d'isophonie diminue. Le seuil différentiel de l'oreille est alors plus élevé ce qui entraîne une dispersion des mesures plus grande.

2-2 Deuxième expérience : Détermination expérimentale de la durée critique de créneaux et d'impulsions

La première expérience ne nous a pas permis de mettre en évidence une durée critique. Les résultats de cette expérience ont permis de montrer qu'il n'y avait pas d'effet de la fréquence mesurable par cette procédure. Une seule fréquence a alors été retenue pour cette deuxième expérience. Afin de pouvoir déterminer la valeur de la durée critique pour les créneaux, une nouvelle expérience a été réalisée avec des stimuli de durées plus longues. Pour les impulsions, nous avons uniquement rajouté une impulsion de 1060 ms. Nous n'avons pas étudié des impulsions de durée supérieure à 2 s car, ne possédant pas de palier, nous ne nous attendons pas à observer de durée critique compte tenu des résultats de la première expérience. De plus, au delà de cette durée d'autres phénomènes entrent en jeu tel que le décrutement.

a) Stimuli

La construction des signaux s'est effectuée de la même manière que pour la première expérience de cette annexe. L'établissement et l'extinction des créneaux ont pour forme des demi-fenêtres de hanning de 5 ms. Leurs durées totales sont de 10, 22, 50, 100, 220, 360, 500, 1000 et 1500 ms. Les temps de descente des impulsions sont de 2, 5, 10, 22, 50, 100, 220, 360, 500 et 1000 ms. La fréquence des signaux est de 1 kHz. Les durées des sons sont récapitulées dans le tableau A-3.3. Tous les niveaux crêtes sont fixés à 72 dB SPL lors de la restitution.

Durée totale des créneaux (ms)	Durée totale des impulsions (ms)
10	10
22	20
50	35
100	70
220	155
360	310
500	670
1000	1060
1500	1520
	2000

Tableau A-3.3: Durée en ms des signaux de 1 kHz dont le niveau crête est de 72 dB SPL.

b) Auditeurs

Les tests ont été réalisés par douze sujets, cinq femmes et sept hommes. Ils avaient entre 19 et 42 ans et avaient tous un audiogramme normal.

c) Procédure

La procédure utilisée est la même que pour la première expérience. Elle diffère uniquement par le fait que le son de comparaison était de la même fréquence que le son test, pour faciliter la tâche de l'auditeur et ainsi diminuer les valeurs des écarts types. C'est-à-dire que le son de comparaison était un son pur de 1 kHz de 1 s. Le niveau d'isotonie est la moyenne des deux jugements des douze auditeurs.

d) Dispositif expérimental

Le dispositif expérimental est le même que dans la première expérience.

e) Résultats

Les niveaux d'isotonie, obtenus pour les créneaux et les impulsions, sont représentés en fonction de la durée sur la figure A-3.6.

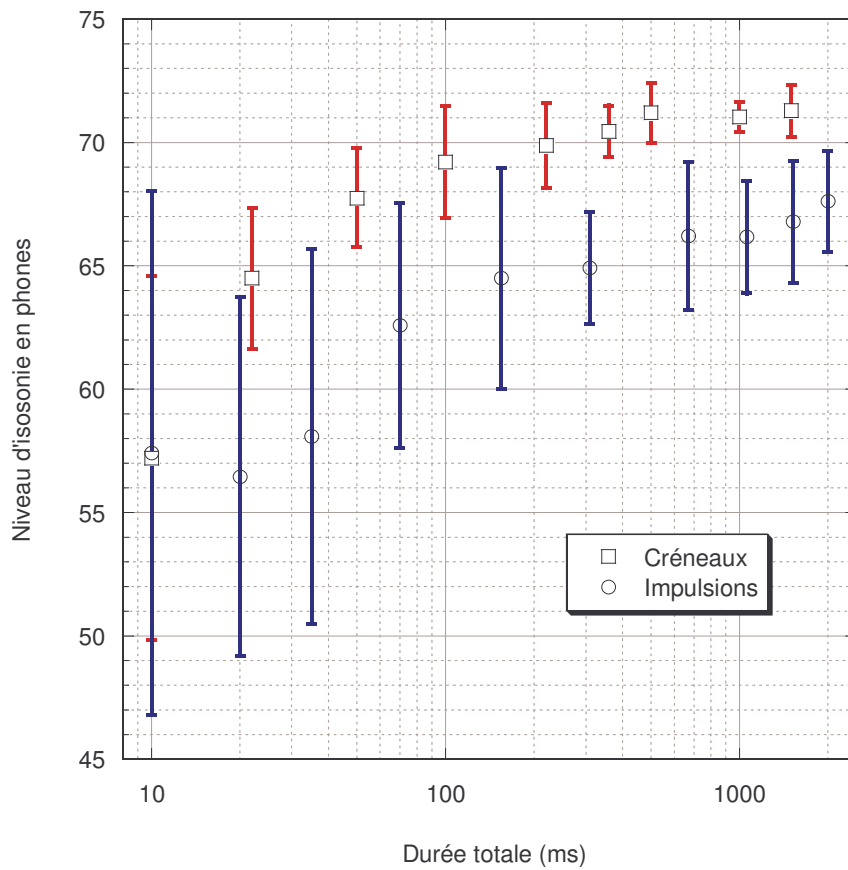


Figure A-3.6 : Niveaux d'isotonie de créneaux et d'impulsions en fonction de leur durée.

On observe sur la figure A-3.6 que, pour les créneaux, le niveau d'isotonie augmente avec la durée jusqu'à 360 ms ; après cette durée, il ne varie que très légèrement (moins de 1 phone). Pour les impulsions, on observe aussi une augmentation du niveau d'isotonie avec la durée mais le niveau d'isotonie ne se stabilise pas même s'il n'augmente que très peu à partir de 100 ms (environ 1 phones par doublement de durée).

Pour les très courtes durées, les valeurs des écarts types augmentent lorsque la durée des stimuli diminue. L'influence du son de comparaison sur la variabilité ne dépend plus, pour des durées très courtes, de la différence des structures fines mais de la différence des durées. Le fait d'avoir pris un son de comparaison de même structure fine que le son test, dans cette expérience, a permis d'éviter les différences de timbre entre le son test et le son de comparaison et a ainsi permis de réduire la variabilité des résultats pour des sons de durées supérieures à 200 ms. En dessous de cette durée, pour diminuer la variabilité, il aurait fallu utiliser un son de comparaison de durée plus courte. Par exemple, comme le préconise Pedersen (1977), le son de comparaison aurait été de 22 ms pour le son test de 10 ms, 50 ms pour le son test de 22 ms et ainsi de suite jusqu'à avoir un son de comparaison dont la durée permette l'établissement de la sonie (environ 360 ms pour les créneaux). Mais cette procédure aurait été plus complexe pour les impulsions.

Cette étude permet aussi d'étudier l'influence de la durée sur la précision des modèles. C'est ce qui est détaillé dans l'annexe 5.

Annexe 4 : L'Estimateur de sonie d'impulsion (ESI)

Cette annexe présente comment a été élaboré l'estimateur de sonie d'impulsion (ESI). Ce travail, antérieur au projet de recherche, présente tout d'abord les caractéristiques physiques de sons impulsionnels tels que nous les avons définis. En effet, les sons construits sur la base d'une enveloppe temporelle rectangulaire ne sont pas totalement représentatifs des sons que l'on peut trouver dans notre environnement. Ceux-ci présentent plutôt des enveloppes à décroissance exponentielle sans partie entretenue. Nous présenterons ensuite les corrélations obtenues entre les caractéristiques physiques et la sonie des sons impulsionnels, afin de proposer un estimateur de sonie d'impulsion noté ESI. Nous utilisons le terme d'estimateur, et non celui de modèle, car le calcul de la sonie ne se base pas ici sur une modélisation du système auditif, mais sur la recherche de paramètres physiques qui influent sur la sonie.

1- Caractéristiques physiques des sons impulsionnels

La première étape de notre étude a été d'enregistrer un panel de sons représentatifs des sons impulsionnels afin d'en déterminer les caractéristiques physiques types. C'est donc avec des percussions, des pierres, des barres de fer, des ballons, des pétards et autres objets que nous avons généré vingt-quatre sons impulsionnels « réels » (voir tableau A-4.1). L'enregistrement des sons a eu lieu en salle anéchoïque avec le matériel suivant :

- Un microphone Bruel & Kjaer type 2669 ;
- Un amplificateur Bruel & Kjaer « Nexus » ;
- Un magnétophone numérique DAT DAP1 Tascam.

Le déroulement temporel de chacun des sons impulsionnels enregistrés peut être caractérisé par une attaque rapide ne dépassant pas 5 ms, pas de palier au maximum d'amplitude et un temps de descente plus ou moins long (voir figure A-4.1).

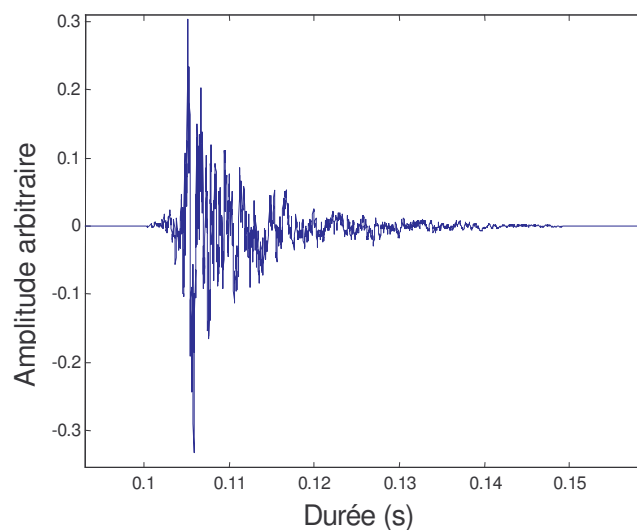


Figure A-4.1 : Allure d'un son impulsionnel représentatif des signaux étudiés.

La figure A-4.2 montre une façon de modéliser le déroulement temporel de ce type de signaux impulsionnels. L'attaque est modélisée par une fonction linéaire croissante et la décroissance par une fonction exponentielle décroissante. L'enveloppe temporelle ainsi définie permet de synthétiser des

sons impulsionnels, dont chaque caractéristique sera contrôlée (voir tableau A-4.2), pour une meilleure analyse de l'influence de chaque paramètre physique sur la sonie.

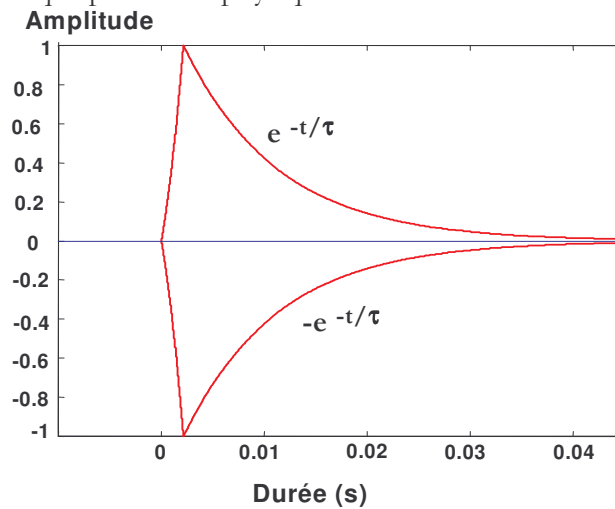


Figure A-4.2 : Modèle d'enveloppe temporelle utilisée dans cette étude pour synthétiser des sons impulsionnels.

Les caractéristiques physiques d'un son impulsionnel que nous pouvons extraire sont, le temps de montée, le niveau crête, le temps de descente et l'énergie (tableaux A-4.1 et A-4.2).

Nous avons défini le temps de montée par le temps mis pour passer de 10 à 90 % du maximum d'amplitude du signal. Le temps de descente est le temps mis pour passer de 90 à 10 % du maximum d'amplitude. Cette définition, inspirée de l'électronique, est aussi utilisée par Scharf et Buus (1986) pour définir ces paramètres temporels. Le calcul du temps de montée et du temps de descente est basé sur le calcul de l'enveloppe temporelle du signal.

Il faut noter que la difficulté ici était de faire un choix d'extraction d'enveloppe valable pour n'importe quel type de signaux complexes. Nous n'avons pas utilisé la méthode de la transformée de Hilbert, mieux adaptée pour extraire l'enveloppe de signaux harmoniques. Notons aussi que le but de l'extraction d'enveloppe temporelle était de pouvoir calculer le temps de montée et le temps de descente. La méthode que nous avons choisie pour déterminer l'enveloppe est basée sur le filtrage passe-bas sur le signal redressé. C'est-à-dire que tant que l'amplitude du signal augmente l'enveloppe suit exactement l'évolution temporelle du signal redressé. Lorsque l'amplitude du signal décroît, l'enveloppe décroît de manière exponentielle mais avec une décroissance moins raide que celle du signal (Zölzer, 1997). Cette décroissance est posée de telle manière que l'on respecte un bon compromis entre la durée du signal et l'obtention d'une enveloppe lissée.

Le niveau crête est le niveau en dB SPL de l'échantillon correspondant à la valeur maximale prise par l'amplitude (correspondant à la valeur maximale de la pression instantanée) pendant la durée totale du signal.

L'énergie est définie comme l'intégrale, sur toute la durée du signal, de l'amplitude au carré du signal.

2- Estimateur de sonie d'impulsion

Après avoir déterminé les différentes caractéristiques physiques, nous avons mis en place une expérience pour mesurer la sonie de quarante sons impulsionnels, aussi bien enregistrés que synthétisés. Dans le paragraphe suivant, nous allons décrire les stimuli étudiés.

2-1 Stimuli

Nom	CGS (Hz)	Niveau crête (dB)	Temps de montée (ms)	Temps de descente (ms)	Energie (W.m ⁻² .s)	Description
bimp01	2123	90	0,1	5,4	9,5	2 pierres
bimp02	1829	79	0,1	3,2	0,9	Pétard
bimp03	601	88,9	0,45	7,2	26,5	Ballon crevé
bimp04	600	78,1	0,1	3,7	1,1	Beachball
bimp05	764	78,5	1,1	18,3	6,1	Claquement de doigt dans la bouche
bimp06	2541	71,2	0,4	3,5	0,3	Claquement de langue
bimp07	895	91,6	1,7	8,9	64,5	Débouchage d'un bouchon de vin
bimp08	751	84,55	0,1	20,6	21,3	Débouchage d'un bouchon de champagne
bimp09	5665	86,25	0,1	3,2	2,5	Jouet clic-clac
bimp10	6420	87,4	0,4	9,3	17,2	Fer sur fer
bimp11	1615	81	2,2	3,5	2,7	Claquement de mains
bimp12	1932	85	0,2	5,5	6,3	Marteau sur bois
bimp13	798	71,7	1,4	54,4	2,2	Marteau sur carreau
bimp14	6535	88,8	0,9	10,9	18,8	Marteau sur fer
bimp15	2409	78,6	0,3	3,3	1,1	Marteau sur pierre
bimp16	2345	83,3	0,2	5,3	3,1	Ping pong1
bimp17	1461	86	0,2	8,3	10,5	Ping pong2
bimp18	643	90,2	4,2	7,1	58,7	Sac éclaté
bimp19	1449	87,15	1,45	5,7	11,4	Tambourin
bimp20	13717	88,9	0,1	42,1	37,0	Triangle
bimp21	3959	90,3	0,1	3,8	12,1	Verre cassé
bimp22	2445	86,2	0,51	9,7	16,7	Wood-block manche
bimp23	2675	83,75	0,7	12,7	10,6	Wood-block
bimp24	2445	82,3	0,6	20,8	5,3	Xylophone

Tableau A-4.1 : Caractéristiques des vingt-quatre sons impulsionnels enregistrés.

Une partie des sons étudiés étaient les sons que nous avons enregistrés. Nous avons reporté dans le tableau A-4.1 certaines de leurs caractéristiques physiques, telles que le centre de gravité spectral en Hz (CGS), le niveau crête en dB SPL, les temps de montée et de descente en ms et l'énergie. La nature de la source est indiquée à la dernière colonne.

Les sons que nous avons synthétisés, pour une analyse plus fine, ont été construits sur le modèle de la figure A-4.2. L'enveloppe était ensuite multipliée par la structure fine du signal (figure A-4.3). La structure fine était soit un son pur de 1 kHz (sons 25 à 28 du tableau A-4.2) soit un bruit blanc (sons 29 à 40 du tableau A-4.2).

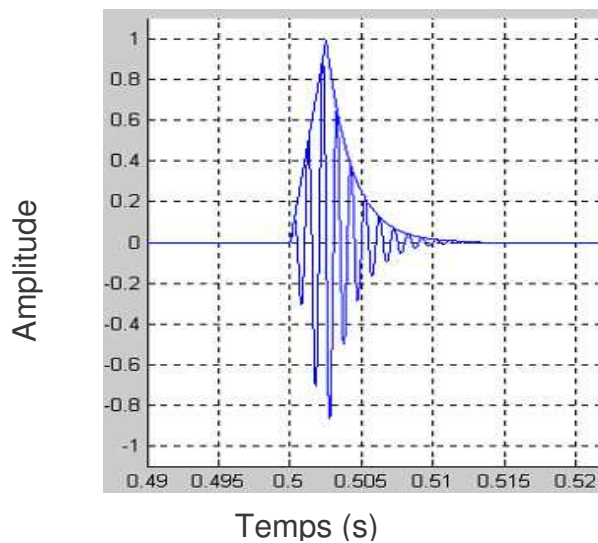


Figure A-4.3 : Exemple d'enveloppe et d'un son synthétisé (enveloppe multipliée par la structure fine)

Le tableau A-4.2 regroupe certaines des caractéristiques physiques des sons de synthèse.

	Description	Niveau crête (dB)	Temps de montée (ms)	Temps de descente (ms)	Energie ($W \cdot m^{-2} \cdot s$)
bimp25	Son pur 1 kHz	88,7	0,5	3,2	12,4
bimp26	Son pur 1 kHz	92,0	0,5	24,7	119,7
bimp27	Son pur 1 kHz	91,3	2,0	24,7	125,6
bimp28	Son pur 1 kHz	90,8	2,0	5,2	30,3
bimp29	Bruit blanc	91,5	0,2	24,4	68,5
bimp30	Bruit blanc	88,2	0,2	4,0	9,3
bimp31	Bruit blanc	90,6	0,3	66,8	208,3
bimp32	Bruit blanc	88,9	0,3	9,6	30,4
bimp33	Bruit blanc	89,7	0,5	4,3	13,4
bimp34	Bruit blanc	91,7	0,4	71,0	208,2
bimp35	Bruit blanc	89,4	0,3	9,7	30,8
bimp36	Bruit blanc	93,1	0,5	24,6	83,6
bimp37	Bruit blanc	91,3	1,5	24,1	75,3
bimp38	Bruit blanc	90,5	1,9	3,8	13,8
bimp39	Bruit blanc	92,3	1,8	71,0	210,0
bimp40	Bruit blanc	89,9	1,6	10,0	32,7

Tableau A-4.2 : Caractéristiques physiques des seize sons impulsionnels synthétisés.

2-2 Auditeurs

Quinze auditeurs otologiquement normaux ont participé aux tests. Les auditeurs étaient des étudiants ou faisaient partie du personnel du Laboratoire de Mécanique et d'Acoustique de Marseille. Ils étaient âgés de 20 à 50 ans.

2-3 Dispositif expérimental

L'interface numérique/analogique était une carte son Oros. Différentes atténuations pouvaient être appliquées aux signaux avec des atténuateurs programmables Wilsonics (model PAT1). Les signaux étaient restitués via une enceinte amplifiée Genelec 1031A dans la salle anéchoïque du laboratoire.

La mesure des niveaux se faisait par l'intermédiaire d'un microphone Bruel & Kjaer (type 2669) situé à la place de la tête de l'auditeur en l'absence de celui-ci.

2-4 Procédure

La sonie a été mesurée en utilisant une méthode d'estimation de grandeur sans référence (voir annexe 2). La tâche de l'auditeur consistait à estimer la sonie d'un son en donnant un nombre (réel positif) de son choix proportionnel à la sonie. Les sons étaient répétés 3 fois chacun et étaient présentés dans un ordre aléatoire, différent pour chaque auditeur. La sonie d'un son pour un auditeur est la moyenne géométrique de ses trois estimations.

La valeur finale de la sonie est la moyenne géométrique des estimations des 15 auditeurs.

2-5 Résultats

Les valeurs de la sonie mesurée par la méthode d'estimation de grandeur pour des sons impulsionnels enregistrés et synthétisés, sont données dans le tableau A-4.3.

Sons réels	Sonie	bimp21	14,6
bimp01	14,3	bimp22	10,6
bimp02	6,1	bimp23	8,4
bimp03	17,5	bimp24	8,6
bimp04	8,6	Sons de synthèse	Sonie
bimp05	10,3	bimp25	9,4
bimp06	4,3	bimp26	18,7
bimp07	16,4	bimp27	14,6
bimp08	16,8	bimp28	10,1
bimp09	4	bimp29	24,2
bimp10	19,3	bimp30	16,6
bimp11	9,3	bimp31	42,4
bimp12	9,4	bimp32	14,7
bimp13	9,9	bimp33	13,6
bimp14	14,2	bimp34	43,4
bimp15	6,5	bimp35	21,9
bimp16	9,1	bimp36	29,0
bimp17	15,1	bimp37	25,4
bimp18	21,2	bimp38	10,9
bimp19	13,05	bimp39	40,9
bimp20	16,1	bimp40	18,7

Tableau A-4.3 : Tableau récapitulatif des valeurs de la sonie mesurée avec la méthode d'estimation de grandeur par quinze auditeurs pour des sons impulsionnels enregistrés et des sons impulsionnels de synthèse.

D'après la loi psychophysique de Stevens (1957), on s'attend à obtenir une relation de puissance entre la sonie (intensité subjective) et l'intensité physique (Energie / temps). On notera que pour les sons impulsionnels que nous étudions, l'intensité varie dans le temps, la sonie donc aussi. Nous avons donc réalisé une régression linéaire sur les valeurs logarithmiques des caractéristiques physiques et de la sonie pour calculer les corrélations entre les différents paramètres (tableau A-4.4).

Matrice de corrélations	Nc	Log(Tm)	Log(Td)	Log(E)	Log(S)
Nc	1	0,14	0,24	0,86	0,70
Log(Tm)	0,14	1	0,16	0,29	0,15
Log(Td)	0,24	0,16	1	0,66	0,68
Log(E)	0,86	0,29	0,66	1	0,88
Log(S)	0,70	0,15	0,68	0,88	1

Tableau A-4.4 : Matrice de corrélations entre le niveau crête Nc, le logarithme du temps de montée (Tm), le logarithme du temps de descente (Td), le logarithme de l'énergie (E) et le logarithme de la sonie (S).

On constate tout d'abord une bonne corrélation entre la sonie et l'énergie (R=0,88). Sur la figure A-4.4, nous avons donc tracé le logarithme de la sonie (S) en fonction du logarithme de l'énergie (E). En pointillés est représentée la droite de régression entre ces deux paramètres (R=0,88). L'équation de cette droite est :

$$\text{Log}(S) = 0,3 \text{ Log}(E) + C \quad (\text{A-4.1})$$

soit :

$$S = K \cdot E^{0,3} \quad (\text{A-4.2})$$

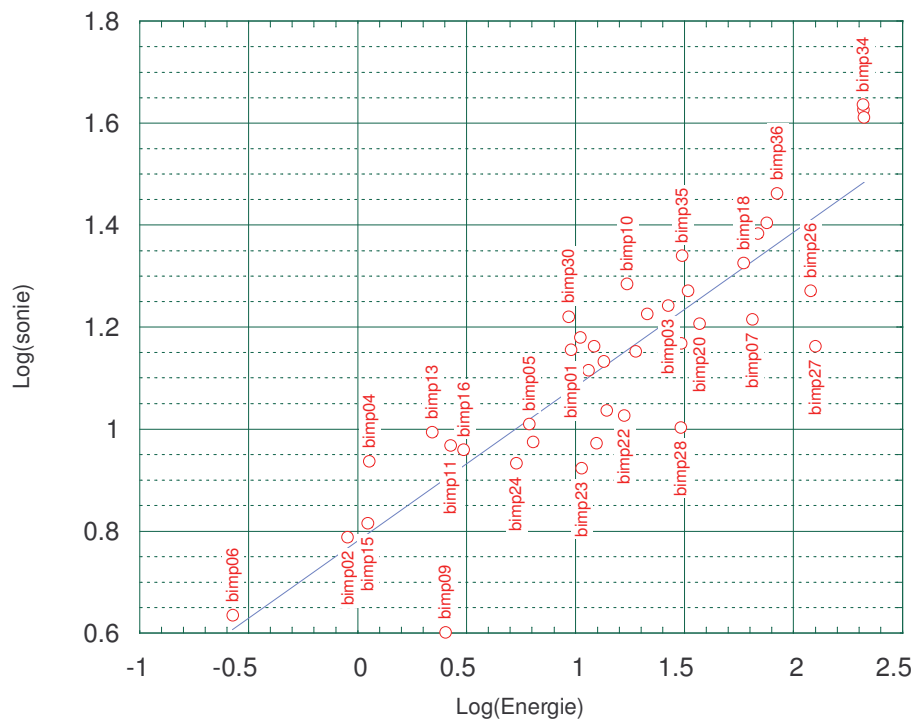


Figure A-4.4 : Logarithme de la sonie en fonction du logarithme de l'énergie. La droite correspond à la fonction puissance (IV-2) sur une échelle logarithmique.

On notera que l'on trouve expérimentalement, pour des sons impulsionnels, ce que Stevens (1957) a montré concernant les sons stationnaires : la sonie est une fonction de l'énergie à la puissance 0,3.

Les valeurs des corrélations du tableau A-4.4 permettent de montrer qu'il y a aussi une bonne corrélation entre la sonie et le niveau crête (R=0,70), ainsi qu'entre la sonie et le temps de descente (R=0,68). Cependant, ces trois paramètres (énergie, niveau crête et temps de descente) ne sont pas tous indépendants : l'énergie est corrélée au niveau crête (R=0,86) et au temps de descente (R=0,66). Afin de

prendre en compte cette dépendance, nous avons réalisé une régression linéaire multiple incrémentielle pour trouver la relation linéaire entre le logarithme de la sonie et, le logarithme de l'énergie, le logarithme du temps de descente et le niveau crête. Le tableau A-4.5 donne les coefficients de cette corrélation.

Régression multiple incrémentielle	R	R ²	R ² _{mod}	P
Log(Energie)	0,87	0,77	0,77	0
Log(Temps de descente)	0,89	0,79	0,02	0,085
Log(Niveau crête)	0,89	0,80	0,01	0,18

Tableau A-4.5 : Coefficients de corrélation obtenus par une régression linéaire multiple incrémentielle entre le logarithme de la sonie et le logarithme de l'énergie (2^{ème} ligne). La troisième ligne donne la corrélation quand le temps de descente est ajouté à l'énergie pour l'établissement de la régression linéaire. Dans la dernière ligne, on donne la corrélation quand le niveau crête est ajouté au temps de descente et à l'énergie dans la régression linéaire. Le coefficient R²_{mod} traduit l'augmentation de R².

Une relation linéaire entre le logarithme de la sonie et le logarithme de l'énergie permet d'expliquer 77 % de la variance de la sonie (R²=0,77). Lorsque le temps de descente est également pris en compte dans la régression linéaire, on augmente le pourcentage de variance expliquée par la régression de 2 % (R²_{mod}=0,02). Le logarithme de la sonie est alors corrélé à 79 % à la relation linéaire de l'énergie et du temps de descente (voir relation A-4.3). Ajouter le niveau crête à l'énergie et au temps de descente dans l'ajustement de la régression linéaire augmente le pourcentage de variance expliquée de 1%.

D'après les résultats de l'analyse incrémentielle de la régression multiple, l'influence du temps de descente sur la sonie ne serait pas significative (p=0,085>0,05). Cependant, il nous semble important de ne pas éliminer ce paramètre. En effet, de nombreuses études (voir Scharf, 1978) ont montré que la durée du signal avait une influence sur la sonie indépendamment de l'énergie. D'après nos données, la probabilité que le temps de descente n'ait pas d'influence sur la sonie est de 0,085 (soit 8,5 %). Ce pourcentage n'étant pas beaucoup plus élevé que 5%, nous préférons ne pas éliminer le temps de descente dès à présent. L'analyse permet aussi de montrer qu'il n'y a pas d'influence du niveau crête ni du temps de montée sur la sonie. En effet, la probabilité que le niveau crête n'ait pas d'influence sur la sonie est de 18 % (p=0,18 >>0,05), ce qui est loin d'être négligeable.

Ainsi, une régression linéaire entre le logarithme de la sonie et les logarithmes de l'énergie et du temps de descente nous permet d'établir un premier estimateur simple de la sonie de sons impulsionnels (Estimateur de sonie d'impulsion noté ESI) :

$$\text{Log}(S)=0,27 \text{Log}(E) + 0,1 \text{Log}(T_d) + C \quad (\text{A-4.3})$$

Ce qui revient à écrire la relation : $S = k \cdot E^{0,27} \cdot T_d^{0,1}$ (A-4.4)

avec S la sonie mesurée, k est une constante d'ajustement, E l'énergie du signal et T_d le temps de descente du signal.

Le coefficient de corrélation entre la sonie calculée par ce nouvel estimateur et la sonie mesurée par les 15 auditeurs sur notre corpus de sons est de 0,89.

Cet estimateur de la sonie de sons impulsionnels a été testé sur une autre série de sons impulsionnels, pour vérifier sa validité. Les résultats sont donnés en annexe 5.

3- Conclusion

Les travaux antérieurs à notre étude montrent que la sonie des sons impulsionnels dépend non seulement de l'énergie ou du temps de descente mais aussi du niveau crête ou du temps de montée. Ces études ont été faites pour la plupart sur des bruits en créneau, c'est-à-dire que les sons présentaient un palier au maximum d'amplitude. Dans notre étude, nous avons voulu étudier des bruits impulsionnels « naturels », par opposition à un son de synthèse. Les sons impulsionnels « naturels » ne présentent pas de palier.

L'influence du temps de montée ou du niveau crête sur la sonie n'a pas été mise en évidence pour les signaux que nous avons étudiés. Il se peut que les temps de montée de nos sons tests soient dans une tranche de durée ($< 1,5$ ms) où effectivement il n'y a pas d'influence sur la sonie (Gustaffson, 1974). De plus, le niveau crête était corrélé à l'énergie. Son influence est donc dissimulée dans la corrélation entre la sonie et l'énergie.

Pour résumer, dans le cas de sons de durées courtes (< 1 s) sur lesquels porte notre étude, la sonie estimée (S) dépend essentiellement de l'énergie ($S=kE^{0,3}$; $R=0,87$) ce qui est en accord avec le modèle de l'oreille comme intégrateur d'énergie. Cependant, on a observé que le temps de descente, proportionnel à la durée du signal, devait être également pris en compte pour une meilleure estimation de la sonie.

ANNEXE 5 : Domaine d'application des modèles de sonie et de l'ESI

1- Lignes isosoniques de sons impulsionnels

Dans l'annexe 3, nous avons présenté les lignes isosoniques à 50, 60 et 75 phones de sons impulsionnels de 22 ms. C'est-à-dire que nous avons déterminé les niveaux crêtes que doivent avoir les sons impulsionnels pour avoir un niveau d'isotonie de 50, 60 ou 75 phones. Les résultats sont rappelés par la figure A-5.1. En ordonnées, nous avons porté les niveaux crêtes pour lesquels les sons sont de sonie égale. La dispersion des données est de l'ordre de 6 dB.

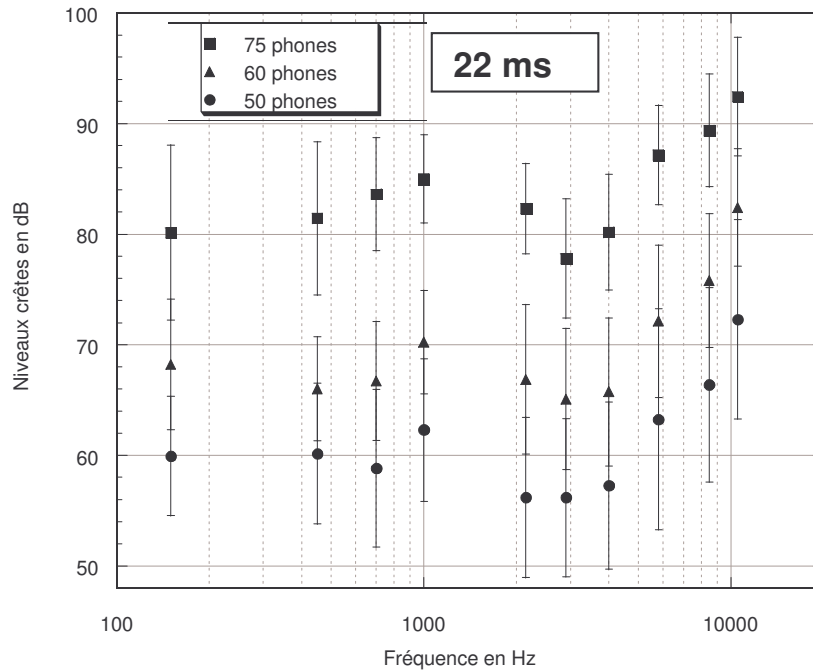


Figure A-5.1 : Lignes isosoniques à 50, 60 et 75 phones de sons impulsionnels de 22 ms, déterminées dans l'annexe 3.

Nous avons calculé le niveau d'isotonie de ces sons en utilisant les différents modèles. Les sons avaient pour niveau crête celui mesuré lors de la détermination des lignes isosoniques (ordonnées de la figure A-5.1).

Les figures A-5.2, A-5.3 et A-5.4 représentent les niveaux d'isotonie donnés par le modèle de Zwicker 1999 (N30ms et Nmax), par notre « estimateur de sonie impulsionnelle » (ESI) et par le modèle de Moore 2002 (STLmax). Les niveaux crêtes des sons sont ceux mesurés pour lesquels les sons sont de niveau d'isotonie égal à 50, 60 et 75 phones. Ainsi, un modèle qui correspondrait parfaitement à nos données fournirait des points alignés sur une droite horizontale passant par 50, 60 et 75 phones respectivement.

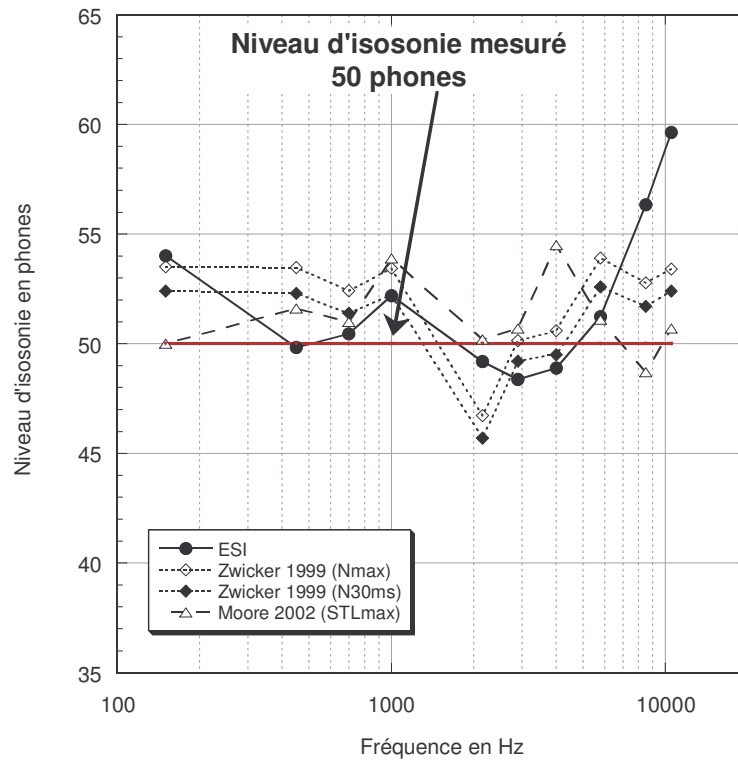


Figure A-5.2 : Niveaux d'isophonie calculés, en fonction de la fréquence de la porteuse, de sons impulsionnels de 22 ms dont le niveau d'isophonie mesuré est de 50 phones.

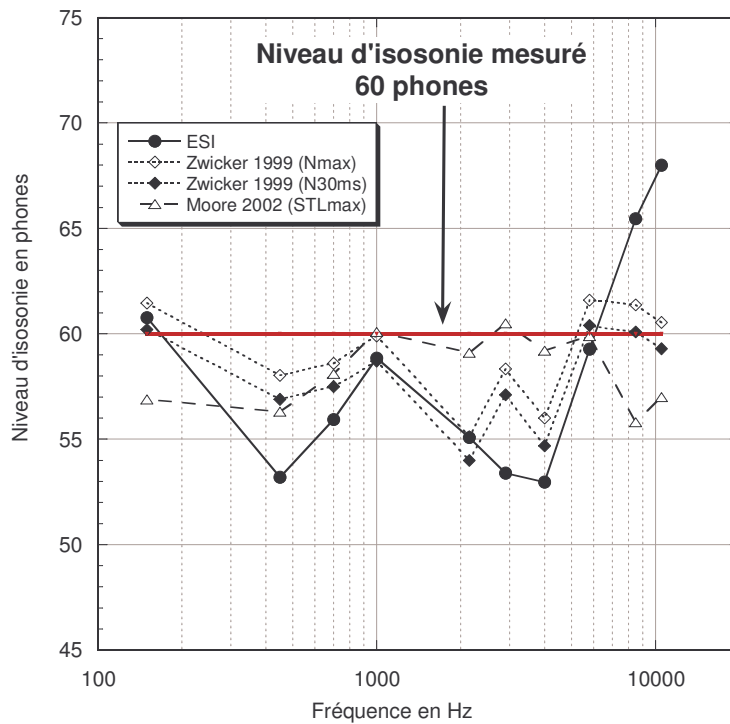


Figure A-5.3 : Niveaux d'isophonie calculés, en fonction de la fréquence de la porteuse, de sons impulsionnels de 22 ms dont le niveau d'isophonie mesuré est de 60 phones.

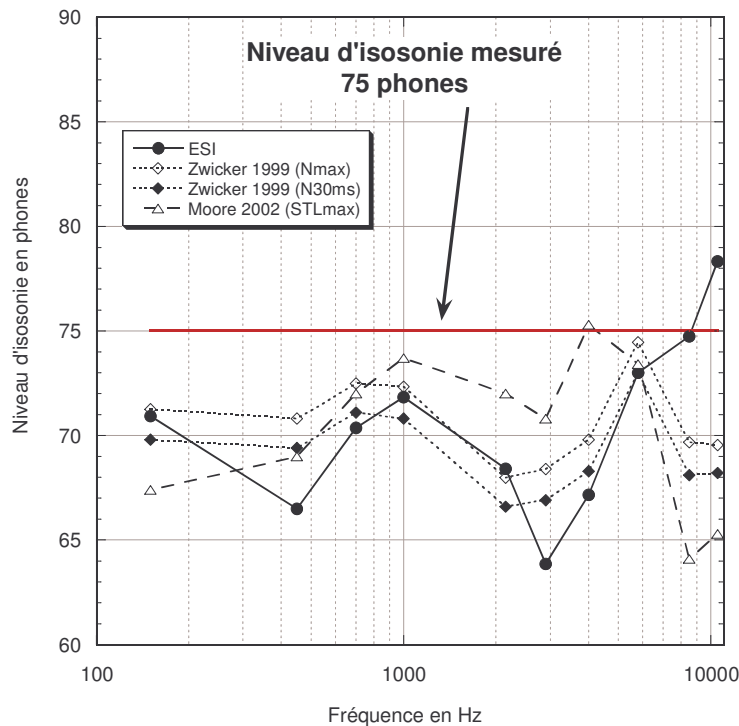


Figure A-5.4 : Niveaux d'isononie calculés, en fonction de la fréquence de la porteuse, de sons impulsionnels de 22 ms dont le niveau d'isononie mesuré est de 75 phones.

De manière générale, et d'après les figures A-5.2, A-5.3 et A-5.4, les modèles donnent le niveau d'isononie à moins de 5 phones près. Pour la ligne à 50 phones, les niveaux d'isononie calculés sont globalement légèrement supérieurs. A 60 phones, au contraire, les modèles donnent des niveaux d'isononie inférieurs à ceux mesurés. En ce qui concerne la dernière ligne isotonique à 75 phones, les modèles donnent un niveau d'isononie au-dessous de celui mesuré. On observe aussi que l'écart entre le niveau d'isononie calculé et mesuré est plus grand qu'à 60 phones. Plus le niveau augmente, plus le niveau d'isononie calculé est inférieur à celui pour lequel les niveaux des sons ont été ajustés.

L'estimateur de sonie d'impulsion (ESI) ne tient pas compte des variations de la sensibilité de l'oreille en fonction de la fréquence, or ce dernier donne des résultats similaires aux modèles. L'ESI dépend de l'énergie et du temps de descente de l'impulsion, soit du niveau du son ajusté. Cela confirme la validité de l'exposant de l'énergie, du moins à 50 et 60 phones.

Le tableau A-5.1 donne le MAR (moyenne des valeurs absolues des écarts), le Rmax (écart maximum), le Rmean (écart moyen) et le Rstd (écart type des écarts moyens) pour les trois lignes isotoniques confondues. En plus des indicateurs calculés à partir du modèle de Zwicker 1999 (Nmax et N30ms), nous avons représenté, dans le tableau A-5.1, le Zwicker 1999 (N5) pour permettre une comparaison plus poussée entre les différents indicateurs.

D'après le tableau A-5.1, trois indicateurs (Nmax, N30ms et STLmax) donnent en moyenne des niveaux d'isononie proches de ceux mesurés (MAR = 3 phones). Par contre l'écart maximum (Rmax) de ces trois indicateurs peut être de 7 phones pour Zwicker 1999 (Nmax), de 8 phones pour Zwicker 1999 (N30ms) et de 11 phones pour Moore 2002 (STLmax). Les niveaux d'isononie calculés avec l'estimateur de sonie d'impulsion (ESI) sont comparables à ceux donnés par les modèles (MAR = 4 phones, Rmax = 11 phones). Cependant, l'écart est plus important (5 à 10 phones) pour des sons dont les structures fines sont des hautes fréquences, en particulier pour des niveaux d'isononie inférieurs à 75 phones.

Lignes isosoniques de sons impulsions	
MAR	
Zwicker 1999 (Nmax)	3
Zwicker 1999 (N30ms)	3
Zwicker 1999 (N5)	6
ESI	4
Moore 2002 (STLmax)	3
Rmax	
Zwicker 1999 (Nmax)	7
Zwicker 1999 (N30ms)	8
Zwicker 1999 (N5)	15
ESI	11
Moore 2002 (STLmax)	11
Rmean	
Zwicker 1999 (Nmax)	-1
Zwicker 1999 (N30ms)	-2
Zwicker 1999 (N5)	-6
ESI	-1
Moore 2002 (STLmax)	-2
Rstd	
Zwicker 1999 (Nmax)	3
Zwicker 1999 (N30ms)	3
Zwicker 1999 (N5)	4
ESI	5
Moore 2002 (STLmax)	3

Tableau A-5.1 : Récapitulatifs des indices proposés (définis au paragraphe I-4 du document principal) afin de pouvoir évaluer l'efficacité des modèles de sonie.

Pour résumer, les modèles ont tendance à donner un niveau d'isotonie inférieur au niveau d'isotonie attendu ($R_{mean} < 0$), en particulier pour des niveaux d'isotonie supérieurs à 50 phones. En effet, à 50 phones, les niveaux d'isotonie calculés sont supérieurs à ceux mesurés. On n'observe pas non plus d'écart constant ($R_{std} > 0$).

Le modèle qui fournit le niveau d'isotonie le plus proche des valeurs attendues est **le Zwicker 1999 (Nmax et N30ms) qui présente un écart moyen de 3 phones** et un écart maximum de l'ordre de 7 ou 8 phones.

2- Intégration temporelle

Nous avons décrit, dans l'annexe 3, l'effet de la durée sur le niveau d'isotonie. Les sons testés étaient soit des créneaux, soit des impulsions dont les niveaux crêtes étaient tous de 72 dB SPL.

Les figures A-5.5 et A-5.6 représentent, en fonction de la durée, respectivement pour les créneaux et les impulsions, les niveaux d'isotonie mesurés et calculés par le modèle de Zwicker et Fastl (N5, N30ms et Nmax), le modèle de Glasberg et Moore (STLmax) et par notre estimateur de sonie d'impulsion (ESI). Les segments verticaux représentent les écarts types à la moyenne des mesures du niveau d'isotonie.

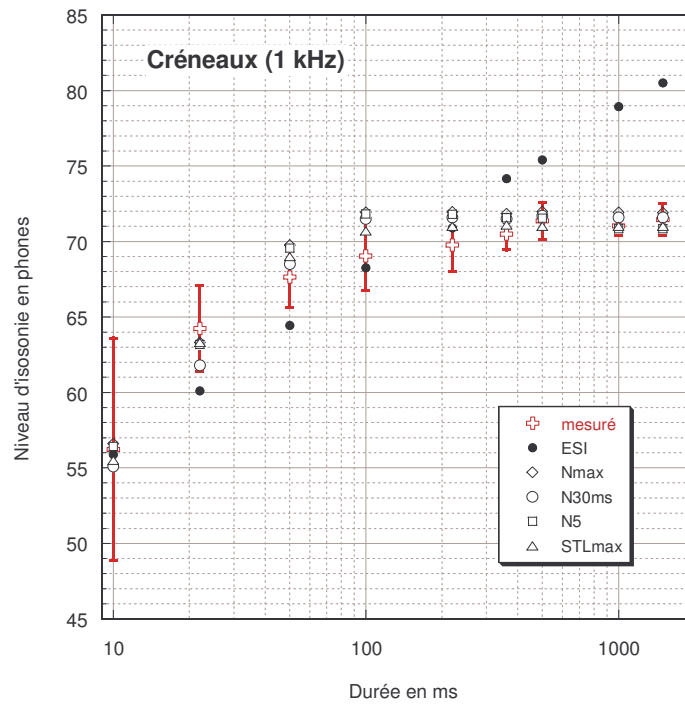


Figure A-5.5 : Niveaux d'isophonie calculés des créneaux de 1kHz en fonction de la durée en millisecondes. Les croix correspondent aux niveaux d'isophonie mesurés. Les barres verticales représentent les valeurs des écarts types des mesures du niveaux d'isophonie.

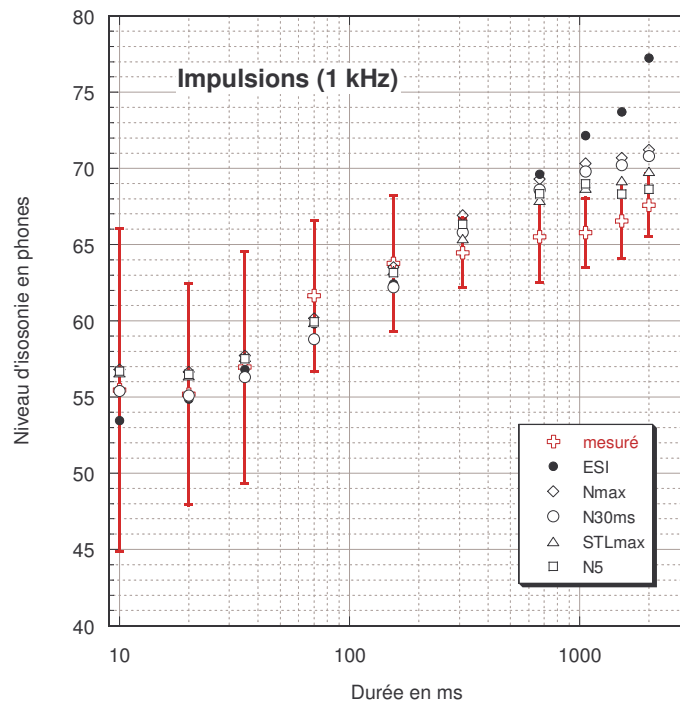


Figure A-5.6 : Niveaux d'isophonie calculés des impulsions de 1kHz en fonction de la durée en secondes. Les croix correspondent aux niveaux d'isophonie mesurés. Les barres verticales représentent les valeurs des écarts types des mesures du niveaux d'isophonie.

On observe sur la figure A-5.5 que, pour les créneaux, les modèles donnent, à moins de 2 phones près, les niveaux d'isophonie mesurés.

En ce qui concerne les impulsions (figure A-5.6), on observe des écarts plus grands entre les niveaux d'isotonie donnés par les modèles et ceux mesurés, en particulier pour des durées supérieures à 300 ms. Pour ces durées, les valeurs du niveau d'isotonie calculé se situent au-delà des valeurs des écarts types. Pour des durées inférieures à 300 ms, les modèles calculent à moins de 2 phones près en moyenne le niveau d'isotonie mesuré, comme pour les créneaux. Pour les très courtes durées, les créneaux et les impulsions sont très proches perceptivement. Au-delà de 300 ms, on entend clairement la décroissance des impulsions. Les sujets semblent alors se baser sur une pondération entre le maximum de sonie et une sonie moyenne alors que les indicateurs (Nmax, N5, N30ms) calculent essentiellement la sonie maximum.

Le tableau A-5.2 fournit les valeurs du MAR, du Rmax, du Rmean et Rstd pour les modèles et l'estimateur de sonie d'impulsion (ESI).

MAR	Créneaux	Impulsions
Zwicker 1999 (Nmax)	1	2
Zwicker 1999 (N30ms)	1	2
Zwicker 1999 (N5)	1	1
ESI	4	6
Moore 2002 (STLmax)	1	2
Rmax		
Zwicker 1999 (Nmax)	3	5
Zwicker 1999 (N30ms)	2	4
Zwicker 1999 (N5)	3	3
ESI	9	14
Moore 2002 (STLmax)	2	3
Rmean		
Zwicker 1999 (Nmax)	1	2
Zwicker 1999 (N30ms)	0	1
Zwicker 1999 (N5)	1	1
ESI	2	5
Moore 2002 (STLmax)	0	1
Rstd		
Zwicker 1999 (Nmax)	1	2
Zwicker 1999 (N30ms)	1	2
Zwicker 1999 (N5)	1	1
ESI	5	6
Moore 2002 (STLmax)	1	1

Tableau A-5.2 : Récapitulatifs des indices proposés (définis au paragraphe I-4 du document principal) afin de pouvoir évaluer l'efficacité des modèles de sonie concernant les données expérimentales sur l'intégration temporelle de créneaux (colonne du milieu) et d'impulsions (colonne de droite) de 1 kHz.

Globalement, les indicateurs tels que le Nmax, le N30ms, le N5 ou le STLmax prédisent bien l'intégration temporelle (MAR de 1 ou 2 phones et $R_{max} = 5$ phones).

Les résultats montrent que **le niveau d'isotonie maximum (Zwicker 1999 (Nmax)) estime assez bien le niveau d'isotonie mesuré des stimuli que nous avons étudiés**. La sonie maximum semble être un des principaux critères, sur lequel s'appuie l'auditeur pour juger la sonie. Cependant, lorsque l'auditeur perçoit une décroissance, ce qui est le cas pour des impulsions de durée supérieure à 300 ms, son jugement ne s'appuie plus uniquement sur la sonie maximum mais semble être pondérée. On atteint ici la limite des indicateurs basés sur le maximum du niveau d'isotonie. Pour des durées plus

longues (> 1 s), ces indicateurs se rapprochent du niveau d'isotonie mesuré.

L'estimateur de sonie d'impulsion que nous avons proposé donne un niveau d'isotonie avec un MAR de 6 phones. Notons que l'estimateur ESI a été construit pour des sons de durée très courte. **Dans le domaine pour lequel il a été créé, il donne des niveaux d'isotonie proches de ceux mesurés (MAR=2 phones).**

Nous allons calculer maintenant les niveaux d'isotonie sur les données expérimentales concernant l'intégration temporelle en fonction de la fréquence. Ainsi, dans ce qui suit, nous avons calculé les niveaux d'isotonie, obtenus par les différents modèles, de sons décrits au paragraphe 3-1.1 de l'annexe 3. Les fréquences des sons étaient de 125 Hz, 350 Hz, 1 kHz, 3 kHz et 8 kHz. La durée des sons était comprise entre 10 et 500 ms pour les créneaux et entre 10 ms et 3 s pour les impulsions. Tous les sons avaient le même niveau crête de 72 dB SPL. Nous avons représenté sur les figures A-5.7 à A-5.16 les niveaux d'isotonie calculés et mesurés pour des créneaux et des impulsions de diverses durées et fréquences.

Pour synthétiser les résultats que nous pouvons observer sur les figures ci-après, nous avons reporté, dans le tableau A-5.3, les valeurs du Mar, du Rmax, du Rmean et Rstd.

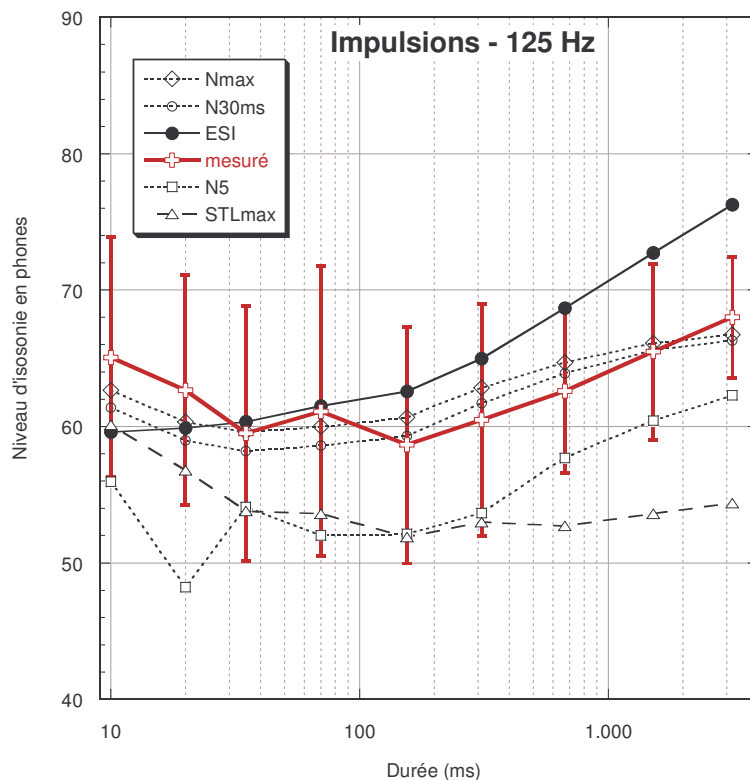


Figure A-5.7 : Niveaux d'isotonie calculés (Nmax, N30ms, ESI, N5 et STLmax) et mesurés (trait épais) d'impulsions de 125 Hz en fonction de leur durée.

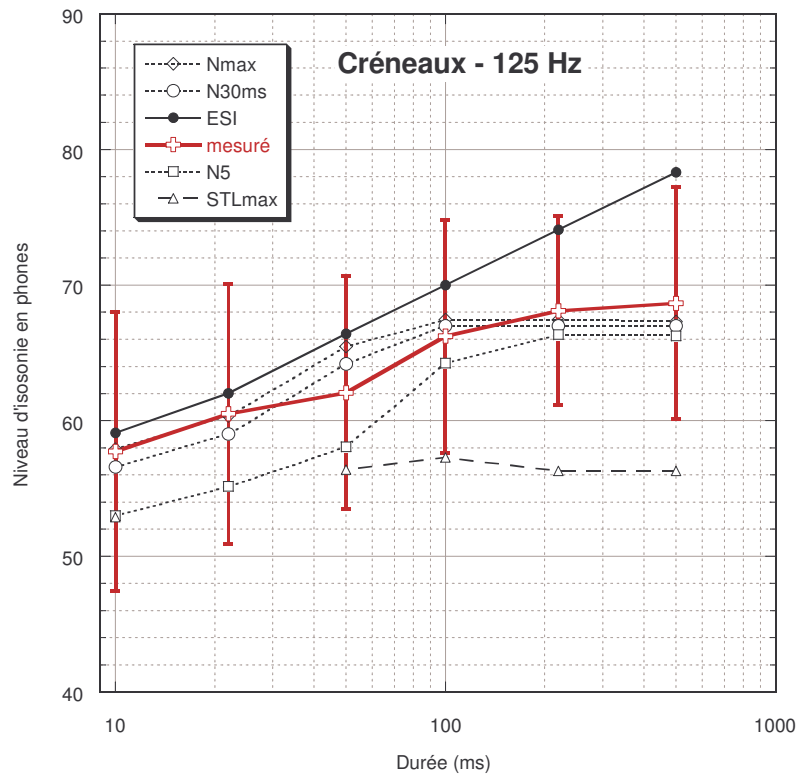


Figure A-5.8 : Niveaux d'isophonie calculés (Nmax, N30ms, ESI, N5 et STLmax) et mesurés (trait épais) de créneaux de 125 Hz en fonction de leur durée.

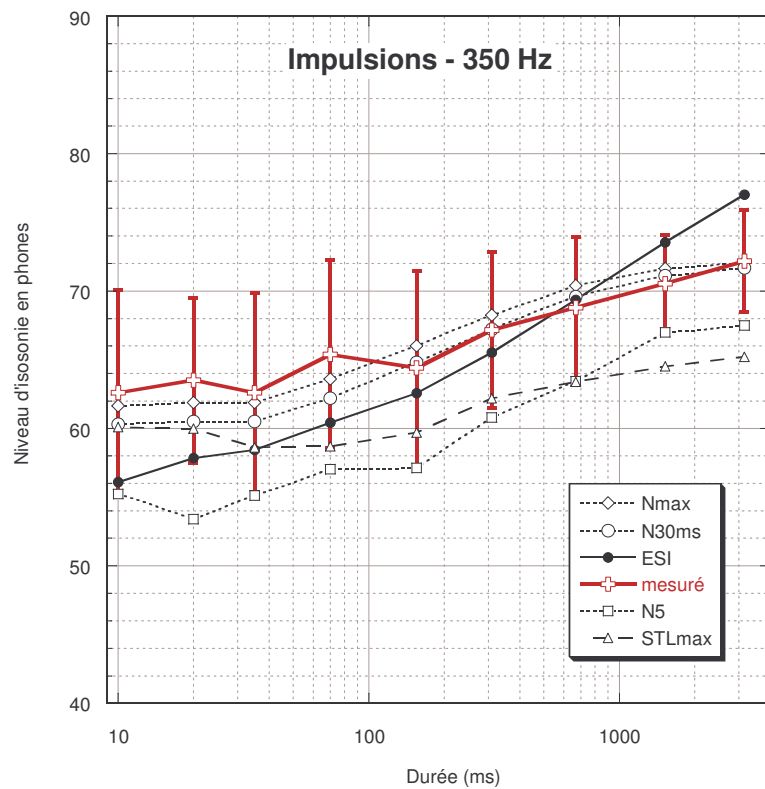


Figure A-5.9 : Niveaux d'isophonie calculés (Nmax, N30ms, ESI, N5 et STLmax) et mesurés (trait épais) d'impulsions de 350 Hz en fonction de leur durée.

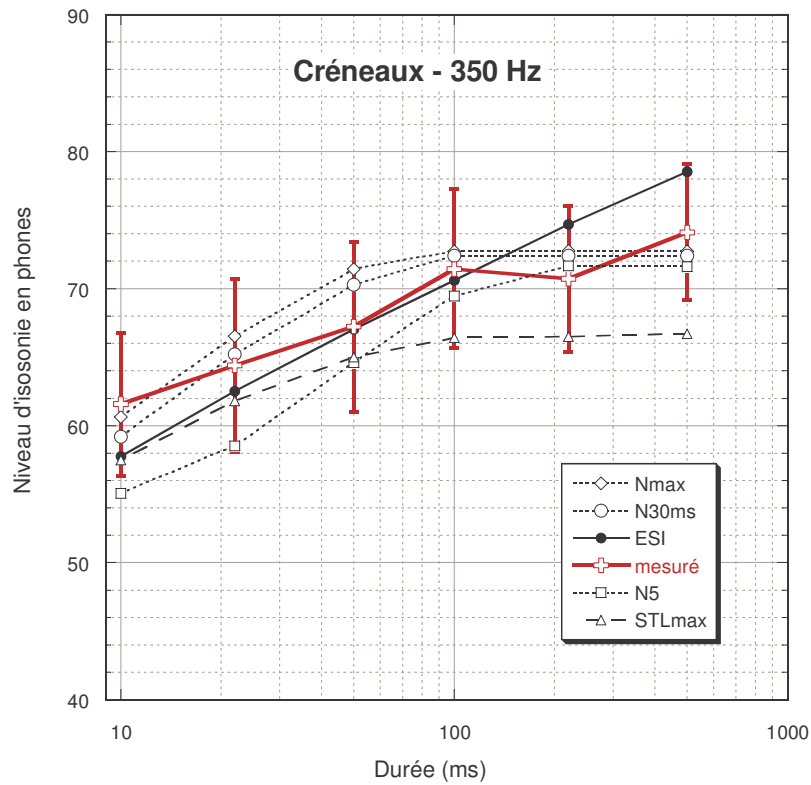


Figure A-5.10 : Niveaux d'isonie calculés (Nmax, N30ms, ESI, N5 et STLmax) et mesurés (trait épais) de créneaux de 350 Hz en fonction de leur durée.

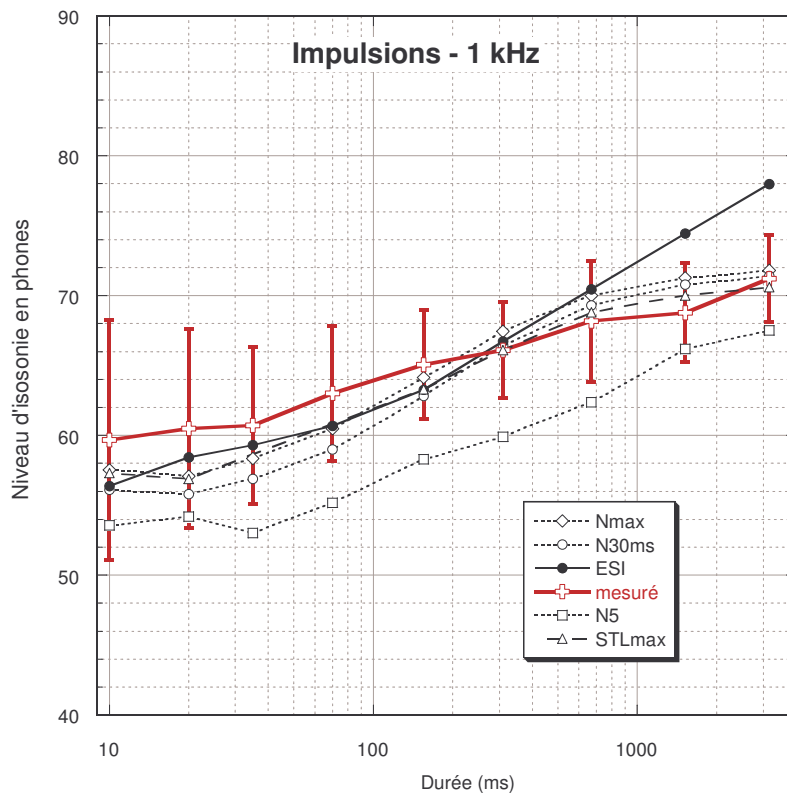


Figure A-5.11 : Niveaux d'isonie calculés (Nmax, N30ms, ESI, N5 et STLmax) et mesurés (trait épais) d'impulsions de 1 kHz en fonction de leur durée.

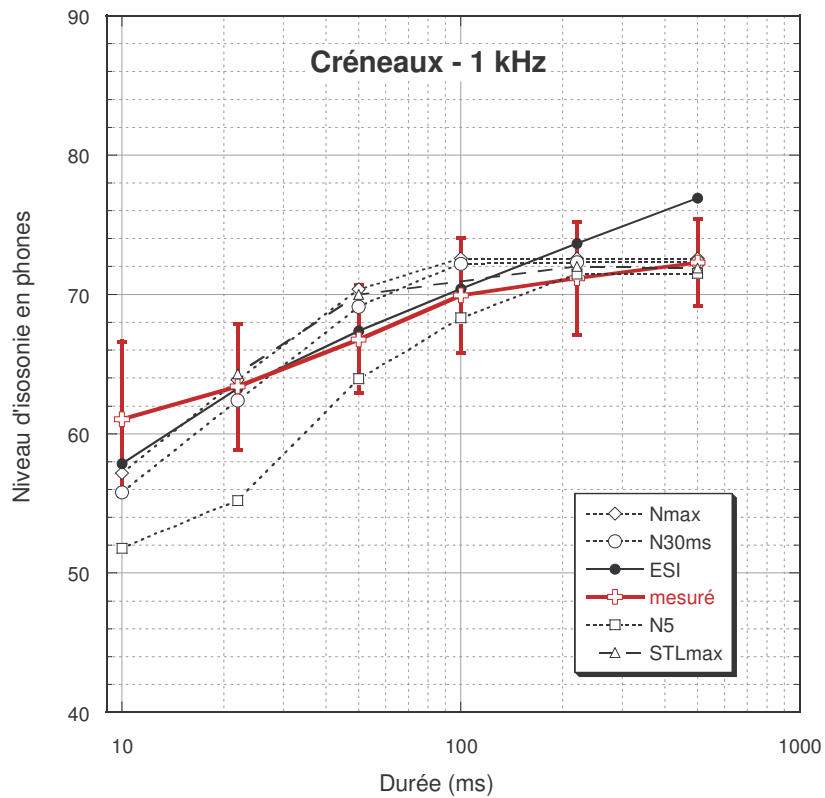


Figure A-5.12 : Niveaux d'isotonie calculés (Nmax, N30ms, ESI, N5 et STLmax) et mesurés (trait épais) de créneaux de 1 kHz en fonction de leur durée.

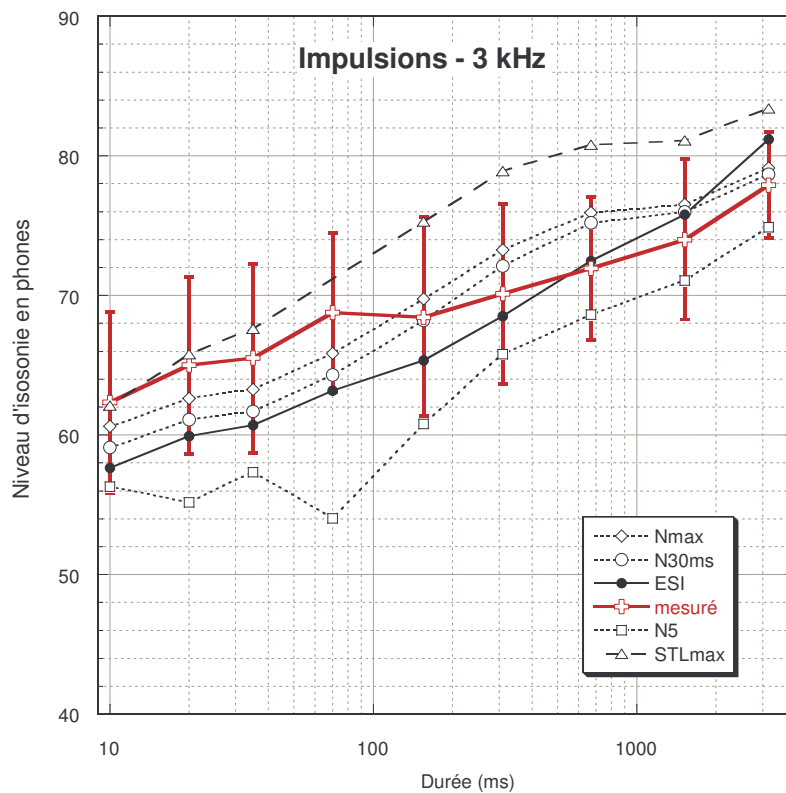


Figure A-5.13 : Niveaux d'isotonie calculés (Nmax, N30ms, ESI, N5 et STLmax) et mesurés (trait épais) d'impulsions de 3 kHz en fonction de leur durée.

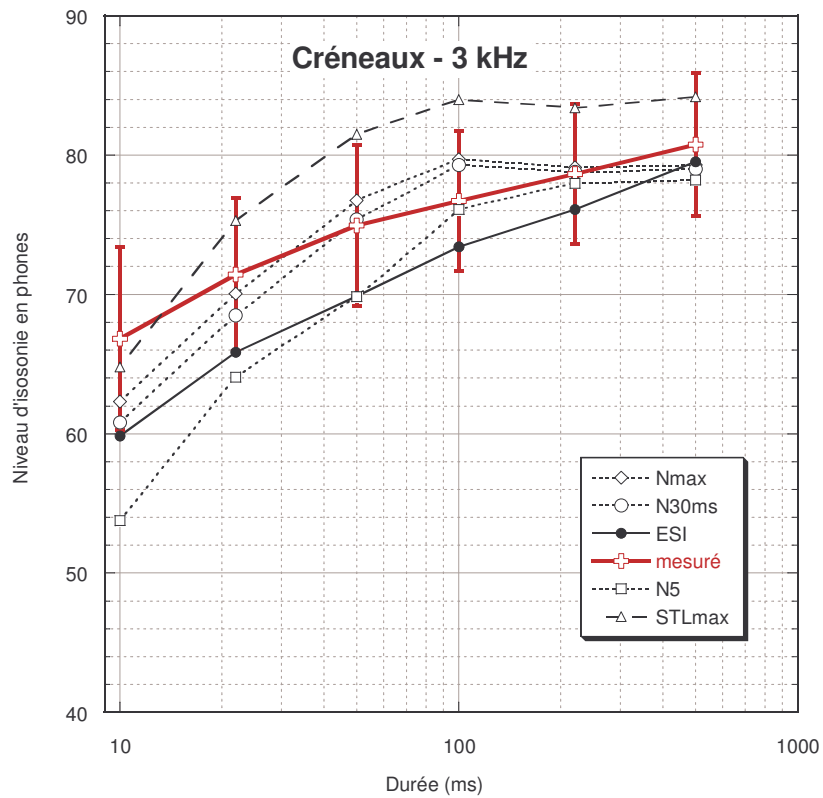


Figure A-5.14 : Niveaux d'isophonie calculés (Nmax, N30ms, ESI, N5 et STLmax) et mesurés (trait épais) de créneaux de 3 kHz en fonction de leur durée.

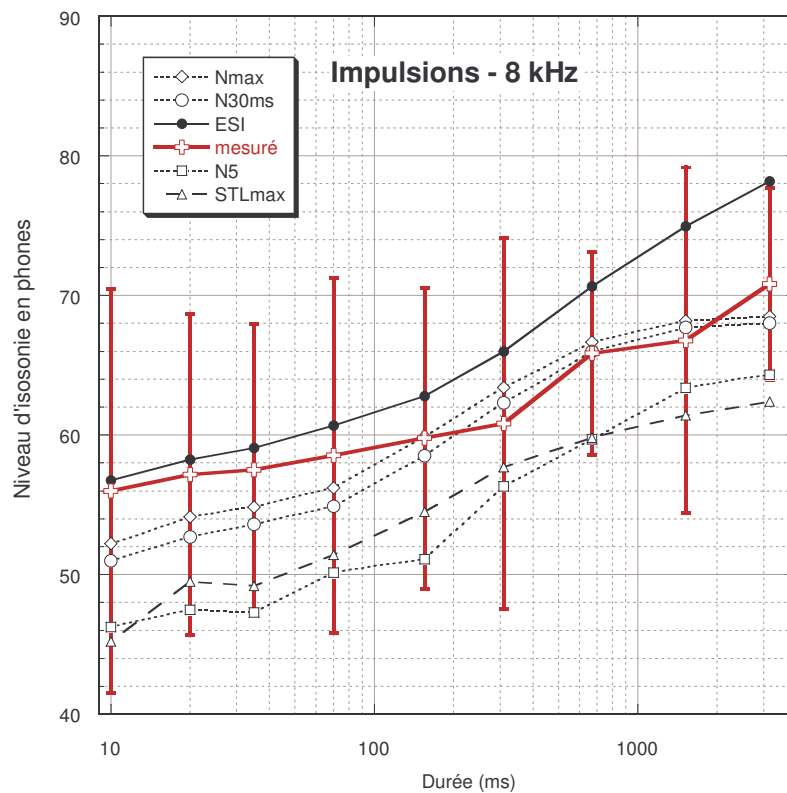


Figure A-5.15 : Niveaux d'isophonie calculés (Nmax, N30ms, ESI, N5 et STLmax) et mesurés (trait épais) d'impulsions de 8 kHz en fonction de leur durée.

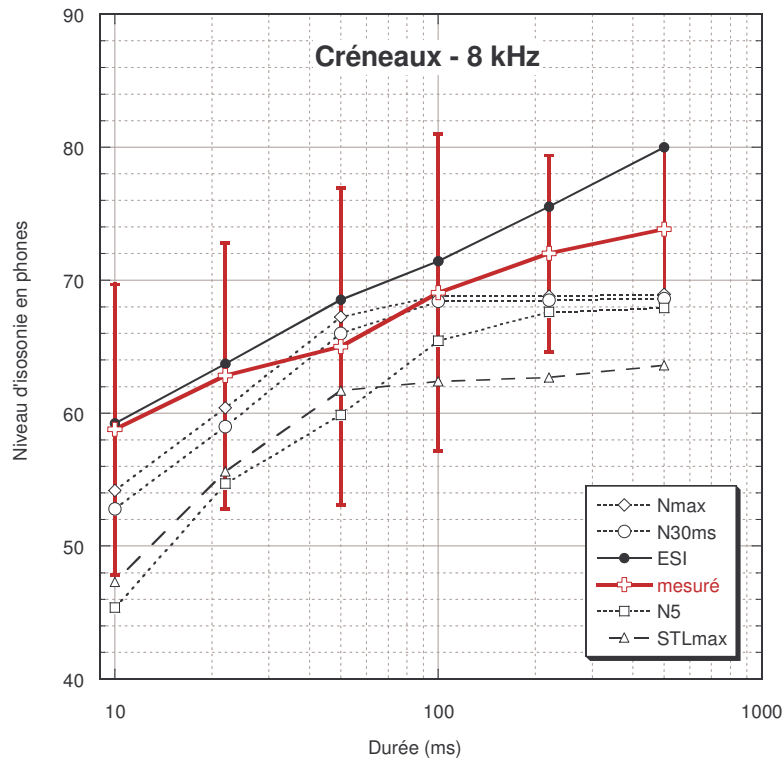


Figure A-5.16 : Niveaux d'isotonie calculés (Nmax, N30ms, ESI, N5 et STLmax) et mesurés (trait épais) de créneaux de 8 kHz en fonction de leur durée.

D'après les valeurs du tableau A-5.3, les écarts entre les niveaux d'isotonie calculée et mesurés, pour chaque fréquence, sont équivalents entre les créneaux et les impulsions.

On observe tout de même que le N5 présente des écarts plus importants pour les impulsions (MAR moyen de 7 phones pour les impulsions et de 4 phones pour les créneaux). De plus, le N5 donne des écarts le plus souvent au-dessous du niveau d'isotonie mesuré ($R_{mean} < 0$) mais pas de manière systématique ($R_{std} \geq 2$).

Il ressort des figures A-5.7 à A-5.16 et du tableau A-5.3, que le Zwicker 1999 (Nmax et le N30ms), quel que soit le signal (créneau ou impulsion) et la fréquence, donnent en moyenne à moins de 2 phones près le niveau d'isotonie mesuré. L'écart maximum pour ces deux indicateurs est de 6 phones, ce qui reste largement acceptable compte tenu de la grande dispersion des données. Le Nmax et le N30ms, donnent en général (sauf pour les créneaux à 8 kHz) des niveaux d'isotonie inférieurs à ceux mesurés pour des sons dont les durées sont inférieures à 20 ms. Pour des sons de durées plus longues, ces deux indicateurs ont tendance à surestimer le niveau d'isotonie mesuré. Globalement, le Nmax donne des écarts répartis uniformément autour du niveau d'isotonie mesuré ($0 < R_{mean} < 1$). Le N30ms, quant à lui, a tendance à donner des niveaux d'isotonie inférieurs à ceux mesurés ($R_{mean} < 0$) mais toujours pas de manière systématique ($R_{std} > 2$).

Le STLmax donne un niveau d'isotonie à moins de 2 phones près, par rapport à celui mesuré, uniquement pour les sons de 1 kHz. Le STLmax donne des niveaux d'isotonie d'autant plus éloignés de ceux mesurés, que la fréquence de la porteuse des sons brefs est plus éloignée de 1 kHz (voir figures A-5.7 à A-5.16 et tableau A-5.3). En effet, pour les créneaux et les impulsions à des fréquences dont le seuil d'audition est supérieur à celui à 1 kHz (125 Hz, 350 Hz et 8 kHz), le STLmax donne un niveau d'isotonie le plus souvent au-dessous de celui mesuré ($R_{mean} < 0$), mais pas avec un biais systématique

($R_{std} \geq 2$). Pour les fréquences dont le seuil d'audition est inférieur à celui de 1 kHz, le STL_{max} surestime le niveau d'isotonie mesuré.

Créneaux					
MAR	Nmax	N30ms	N5	ESI	STLmax
125 Hz	1	1	3	4	9
350 Hz	2	2	3	3	4
1 kHz	2	2	4	2	1
3 kHz	2	2	5	4	5
8 kHz	3	3	7	3	7
Moyenne	2	2	4	3	5
Rmax	Nmax	N30ms	N5	ESI	STLmax
125 Hz	3	2	5	10	12
350 Hz	4	3	6	4	7
1 kHz	4	5	9	5	3
3 kHz	5	6	13	7	7
8 kHz	5	6	13	6	10
Max	5	6	13	10	12
Rmean	Nmax	N30ms	N5	ESI	STLmax
125 Hz	0	0	-3	4	-9
350 Hz	1	0	-3	0	-4
1 kHz	1	0	-4	1	1
3 kHz	0	-1	-5	-4	4
8 kHz	-2	-3	-7	3	-7
Moyenne	0	-1	-4	1	-3
Rstd	Nmax	N30ms	N5	ESI	STLmax
125 Hz	2	2	2	3	3

Impulsions					
MAR	Nmax	N30ms	N5	ESI	STLmax
125 Hz	2	2	7	4	8
350 Hz	1	1	7	4	5
1 kHz	2	2	6	3	2
3 kHz	2	3	7	3	5
8 kHz	2	3	7	4	7
Moyenne	2	2	7	4	5
Rmax	Nmax	N30ms	N5	ESI	STLmax
125 Hz	2	4	14	8	14
350 Hz	2	3	10	6	7
1 kHz	3	5	8	7	4
3 kHz	4	4	15	6	9
8 kHz	4	5	10	8	11
Max	4	5	15	8	14
Rmean	Nmax	N30ms	N5	ESI	STLmax
125 Hz	0	-1	-7	2	-8
350 Hz	0	-1	-7	-2	-5
1 kHz	-1	-2	-6	0	-1
3 kHz	0	-1	-7	-2	5
8 kHz	-1	-2	-7	4	-7
Moyenne	-1	-1	-7	1	-3
Rstd	Nmax	N30ms	N5	ESI	STLmax
125 Hz	2	2	3	5	3

Tableau A-5.3 : Récapitulatifs des indices proposés (définis au paragraphe I-4 du document principal) afin de pouvoir évaluer l'efficacité des modèles de sonie concernant l'intégration temporelle de créneaux et d'impulsions pour diverses fréquences. A gauche les indices concernant les créneaux et à droite les impulsions.

On observe, à partir de ces résultats, que **l'estimateur de sonie d'impulsion (ESI) donne un niveau d'isotonie en moyenne à moins de 4 phones de celui mesuré**. Les écarts les plus grands entre le niveau d'isotonie calculé et mesuré sont obtenus pour des sons à basses et hautes fréquences (10 et 8 phones), de durée supérieure à 1 seconde pour les impulsions et à 200 ms pour les créneaux. Rappelons encore une fois que cet estimateur a été créé pour des sons de très courtes durées. Donc, dans le domaine d'application pour lequel il est prévu, **il est aussi performant que d'autres indicateurs tels que le Nmax et le N30ms sur ce type de sons**. De plus, les valeurs du tableau A-5.3 permettent de montrer que l'estimateur ESI donne des niveaux soit au-dessous, soit au-dessus de ceux mesurés et ne comportent pas de biais systématique ($R_{std} \geq 2$).

Globalement, nous constatons les limites de certains indicateurs, tels le N5 et le STL_{max} , pour calculer le niveau d'isotonie d'un son bref. Notons tout de même que ces indicateurs sont préconisés par les auteurs pour estimer le niveau d'isotonie de sons variant dans le temps. Il ressort aussi que le Nmax ou le N30ms sont de bons indicateurs pour calculer le niveau d'isotonie de créneaux et d'impulsions dont les structures fines sont des sons purs ($MAR = 2$ phones). Les écarts maximaux sont compris entre 4 et 6 phones. En ce qui concerne l'estimateur ESI, celui-ci donne d'aussi bons résultats ($MAR = 3$ et 4 phones) que le Nmax et le N30ms, pour ce type de sons brefs.

Il est intéressant, à présent, de comparer les niveaux d'isononie calculés et mesurés de signaux de synthèse plus complexes et de sons naturels.

3- Domaine d'application des modèles de sonie en fonction du type de sons

Nous présentons ici les résultats des tests des modèles de sonie et de l'estimateur ESI sur des sons plus complexes que ceux étudiés jusqu'alors. Pour cela, nous avons repris les données expérimentales décrites en annexe 2. Les sons testés étaient des sons impulsionnels, des sons stationnaires et des trains d'impulsions, naturels et synthétiques. Nous avons donc choisi de comparer les données des modèles et de l'estimateur ESI aux niveaux d'isononie mesurés par la méthode d'ajustement. On rappelle que quatorze auditeurs ont participé aux tests.

Notons que pour les sons stationnaires, nous avons aussi testés les modèles de sonie de Zwicker (1958) et de Moore et coll. (1997) prévus pour ce type de son. Ils sont notés dans ce qui suit par Zwicker 1958 et Moore 1997.

Les figures A-5.17 et A-5.18 représentent les niveaux d'isononie mesurés et calculés respectivement pour des sons impulsionnels et des trains d'impulsions. Nous avons relié les points expérimentaux par des segments pour une meilleure visibilité des différences entre les modèles. Les segments verticaux représentent les écarts types des mesures.

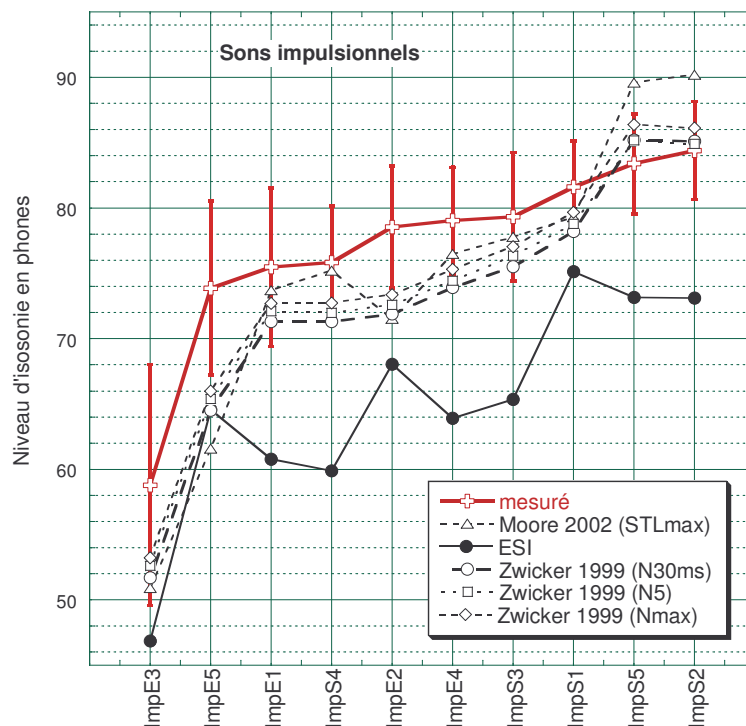


Figure A-5.17 : Niveaux d'isononie en phones de sons impulsionnels mesurés et calculés par le modèle de Moore 2002 (STLmax), l'estimateur ESI et le modèle de Zwicker 1999 (N30ms, N5 et Nmax).

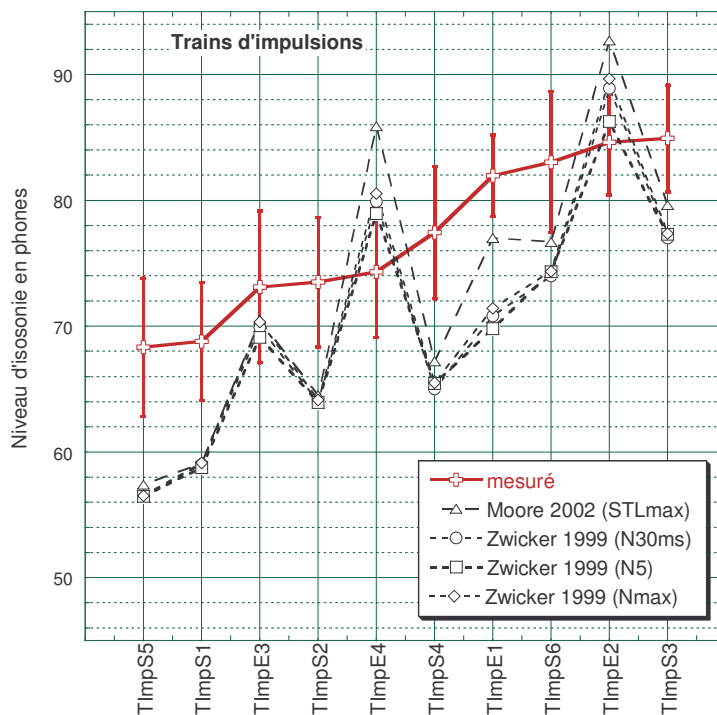


Figure A-5.18 : Niveaux d'isononie en phones des trains d'impulsions mesurés et calculés par le modèle de Moore 2002 (STLmax) et le modèle de Zwicker 1999 (N30ms, N5 et Nmax).

Le tableau A-5.4 récapitule les indices permettant de quantifier les écarts entre les niveaux d'isononie calculés et mesurés en ce qui concerne les sons impulsionnels et les trains d'impulsions.

MAR	Impulsionnel	Train d'impulsions	Moyenne
Zwicker 1999 (Nmax)	4	8	6
Zwicker 1999 (N30ms)	5	8	6
Zwicker 1999 (N5)	4	8	6
ESI	12		10,5
Glasberg et Moore 2002 (STLmax)	7	8	6
Rmax	Impulsionnel	Train d'impulsions	Maximum
Zwicker 1999 (Nmax)	8	12	12
Zwicker 1999 (N30ms)	9	12	12
Zwicker 1999 (N5)	8	12	12
ESI	16		16
Glasberg et Moore 2002 (STLmax)	12	12	14
Rmean	Impulsionnel	Train d'impulsions	Moyenne
Zwicker 1999 (Nmax)	-2	-6	
Zwicker 1999 (N30ms)	-3	-7	
Zwicker 1999 (N5)	-3	-7	-5
ESI	-10		-9,5
Glasberg et Moore 2002 (STLmax)	-1	-4	-3
Rstd	Impulsionnel	Train d'impulsions	Moyenne
Zwicker 1999 (Nmax)	4	7	5
Zwicker 1999 (N30ms)	5	7	3
Zwicker 1999 (N5)	4	6	4
ESI	8		6
Moore 2002(STLmax)	6	8	6

Tableau A-5.4 : Récapitulatifs des indices afin de pouvoir évaluer l'efficacité des modèles de sonie concernant des données sur des sons impulsionnels et des trains d'impulsions.

De manière générale, les modèles ont tendance à donner des niveaux d'isotonie inférieurs à ceux mesurés ($R_{mean} < 0$). On ne constate pas non plus de biais systématique ($R_{std} \geq 4$) sur l'ensemble des résultats.

Pour les sons impulsionnels, c'est le modèle de Zwicker 1999 qui donne les écarts les plus faibles par rapport aux niveaux d'isotonie mesurés ($Mar = 4$ phones et $R_{max} = 8$ phones). Le modèle de Moore 2002 (STLmax) donne des niveaux d'isotonie inférieurs à ceux mesurés pour des niveaux d'isotonie faibles et supérieurs pour des niveaux d'isotonie élevés. L'estimateur ESI donne des niveaux d'isotonie toujours inférieurs à ceux mesurés. Il semble sous-estimer d'avantage les niveaux d'isotonie mesurés pour des niveaux élevés. Ce résultat a aussi été observé lors de la comparaison des modèles sur les résultats des lignes isotoniques de sons impulsionnels. Il est important de souligner ici que l'estimateur ESI a été ajusté sur un nombre d'échantillons sonores limité, obtenus avec une méthode d'estimation directe. Cette méthode avait tendance à donner des niveaux d'isotonie inférieurs à ceux obtenus par la méthode d'ajustement (voir annexe 2). Ceci peut donc expliquer que l'estimateur ESI donne, pour ces données, des niveaux d'isotonie inférieurs à ceux mesurés par la méthode d'ajustement.

Enfin, en ce qui concerne les trains d'impulsion, même si tous les modèles donnent des valeurs équivalentes, aucun ne donne des niveaux d'isotonie proches de ceux mesurés (MAR de l'ordre de 8 phones et R_{max} de 12 phones).

De manière générale, on peut penser qu'il est nécessaire de continuer à vouloir améliorer les modèles existants et en particulier l'ESI pour qu'il tienne compte de l'influence de la fréquence. C'est ce que nous allons présenter dans l'annexe 6.

Annexe 6 : Mesure des fonctions de sonie pour l'ESI_{bc}

Nous présentons dans cette annexe les résultats concernant la détermination des exposants des différentes fonctions de sonie, pour plusieurs fréquences de porteuse et plusieurs durées.

1- Stimuli

Les sons impulsionnels étudiés ont tous été synthétisés. Leur enveloppe temporelle avait une attaque modélisée par une fonction linéaire croissante sur 5 ms, pas de palier, et une décroissance de 20 ms modélisée par une fonction exponentielle décroissante. Le signal était coupé à -60 dB. La structure fine de remplissage des enveloppes était composée de sons purs de 250, 700, 1 370, 1 850, 2 500, 3400, 4800 et 10500 Hz. Chaque son, pour une fréquence donnée, était joué à 11 niveaux crêtes compris entre 50 et 100 dB SPL par pas de 5 dB.

Le tableau A-6.1 récapitule les caractéristiques physiques des signaux que nous avons étudiés.

Nom	Temps de descente (ms)	Durée totale (ms)	Fréquence de la porteuse (Hz)	Niveaux crête (dB SPL)
I250_20ms	20	72	250	50, 55, 60, 65, 70, 75, 80, 85, 90, 95 et 100
I450_20ms			450	
I700_20ms			700	
I1k_20ms			1 000	
I1k37_20ms			1370	
I1k85_20ms			1 850	
I2k5_20ms			2 500	
I3k4_20ms			3 400	
I4k8_20ms			4 800	
I7k_20ms			7 000	
I10k5_20ms			10 500	

Tableau A-6.1 : Tableau récapitulatif des caractéristiques physiques des sons étudiés pour établir les fonctions de sonie.

2- Dispositif expérimental

Le signal numérique est converti en signal analogique grâce à une carte son Oros. Le signal est ensuite atténué par un atténuateur programmable Wilsonics (model PATT) avant d'être restitué par une enceinte amplifiée Génélec (type 1031A) située dans la salle anéchoïque du LMA. La mesure des niveaux est faite par un microphone Bruel & Kjaer (type 2669) situé à la place de la tête de l'auditeur en l'absence de celui-ci.

3- Auditeurs

Les tests ont été réalisés par dix-neuf auditeurs (dix femmes et neuf hommes) âgés de 20 à 45 ans. Dix d'entre eux n'avaient jamais passé de test d'estimation directe. Tous les sujets, soumis à un audiogramme de Békésy sous casque en salle anéchoïque, ne présentaient aucune déficience auditive.

4- Procédure

La mesure de la sonie a été faite par une méthode d'estimation directe sans référence. Dans ce test, les sons étaient répétés trois fois chacun et étaient présentés dans un ordre aléatoire, différent pour chaque

auditeur. L'auditeur estimait la sonie de chaque son en donnant un nombre de son choix proportionnel à la sonie. Pour chaque auditeur et pour chaque son, on calcule la moyenne géométrique des trois estimations. Les données sont ensuite normalisées par rapport à une référence de sonie connue. La sonie d'un son est la moyenne géométrique des estimations de tous les auditeurs.

5- Résultats

Les fonctions de sonie peuvent être ajustées par une fonction puissance. Les valeurs des exposants ont été introduites dans l' ESI_{bc} pour tenir compte de l'influence de la fréquence.

Annexe 7 : Processus de validation

Le but de cette étude est de commencer un processus de validation de l'ESI_{bc} sur un grand nombre de stimuli afin de valider ce dernier. Les données expérimentales ont été obtenues à partir de mesures de sonie de sons impulsionnels de synthèse de différentes durées (231 sons) et de sons impulsionnels « naturels » (26 sons).

1- Stimuli

Les sons impulsionnels de synthèse avaient les caractéristiques suivantes : leur enveloppe temporelle avait une attaque modélisée par une fonction linéaire croissante sur 5 ms, pas de palier, et une décroissance de 5, 40, 60, 80, 100, 200 et 600 ms, modélisée par une fonction exponentielle décroissante. Le signal était coupé à -60 dB. La structure fine de remplissage des enveloppes, pour ces durées, était composée de sons purs de 450 Hz, 1 kHz et 7 kHz. Chaque son, pour une fréquence et une durée données, était joué à onze niveaux crêtes compris entre 50 et 100 dB SPL par pas de 5 dB.

Les tableaux A-7.1 et A-7.2 récapitulent les caractéristiques physiques des signaux que nous avons étudiés respectivement de synthèse et « naturels ».

<i>Nom</i>	<i>Temps de descente (ms)</i>	<i>Durée totale (ms)</i>	<i>Fréquence de la porteuse (Hz)</i>	<i>Niveaux crête (dB SPL)</i>
<i>I_5ms</i>	<i>5</i>	<i>25</i>	<i>450, 1 000 et 10 500</i>	<i>50, 55, 60, 65, 70, 75, 80, 85, 90, 95 et 100</i>
<i>I_40ms</i>	<i>40</i>	<i>135</i>		
<i>I_60ms</i>	<i>60</i>	<i>198</i>		
<i>I_80ms</i>	<i>80</i>	<i>261</i>		
<i>I_100ms</i>	<i>100</i>	<i>324</i>		
<i>I_200ms</i>	<i>200</i>	<i>638</i>		
<i>I_600ms</i>	<i>600</i>	<i>1 896</i>		

Tableau A-7.1 : Tableau récapitulatif des descriptions et caractéristiques physiques des signaux de synthèse étudiés.

Nom des sons	Description	Durée	Niveau crête
Son1	Agrafeuse	50	73,7
Son2	Clignotant	30	65,3
Son3	Clignotant	40	65
Son4	Clé de contact	50	71,2
Son5	Fermeture centralisée	150	76
Son5	Fermeture centralisée	150	66
Son5	Fermeture centralisée	150	56
Son6	Frein à main (enlevé)	40	73
Son7	Frein à main (mis)	130	80,2
Son8	Guitare	1620	69,2
Son9	Interrupteur électrique	25	66,3
Son10	Klaxon	520	92,5
Son11	Coup de maillet sur bois	315	88
Son12	Coup de maillet sur fer	530	78,7
Son13	Coup de marteau sur bois	215	72,5
Son14	Coup de marteau sur fer	450	86,8
Son15	Bruit de pas	125	67,2
Son16	Percussion en bois	170	87,1
Son17	Plaque de fer frappée	900	82,3
Son18	Fermeture de portière	200	86,3
Son19	Ouverture de portière	250	82
Son20	Retrait clé de contact	100	80,1
Son21	Son « erreur » de Windows	390	89
Son22	Stylo bic à poussoir	10	72,4
Son23	Fermeture capuchon bic	5	68,7
Son24	Touche de clavier	80	60,8

Tableau A-7.2 : Tableau récapitulatif des descriptions et caractéristiques physiques des signaux « naturels » étudiés.

2- Dispositif expérimental

Le dispositif était le même que celui décrit en annexe 6.

3- Auditeurs

En ce qui concerne la mesure de la sonie des signaux de synthèse, les tests ont été réalisés par dix-neuf auditeurs (dix femmes et neuf hommes) âgés de 20 à 45 ans.

Les tests concernant les signaux « naturels » ont été réalisés par vingt auditeurs (neuf femmes et onze hommes) âgés de 22 à 43 ans.

Tous les sujets, soumis à un audiogramme de Békésy sous casque en salle anéchoïque, ne présentaient aucune déficience auditive.

4- Procédure

La procédure expérimentale de mesure de la sonie est la même que celle décrite en annexe 6.

5- Résultats

Les tableaux A-7.3, A-7.4 et A-7.5 fournissent les valeurs du MAR, du R_{max} , du R_{mean} et R_{std} pour le N30ms, indicateur calculé à partir du modèle de Zwicker et Fastl (1999), de l'estimateur de sonie d'impulsion (ESI) et de l'estimateur de sonie d'impulsion modifié (ESI_{bc}).

N30ms								
450Hz	5 ms	40 ms	60 ms	80 ms	100 ms	200 ms	600 ms	Moyenne
MAR	5,63	3,69	4,06	3,62	4,20	4,99	4,69	4,21
Rmean	5,63	-1,50	-2,82	-2,19	-2,50	-2,98	-2,15	-1,10
Rmax	8,49	6,79	8,66	7,02	8,86	10,58	9,64	10,58
Rstd	2,06	4,20	4,38	3,81	4,59	5,28	5,30	4,09
1 000 Hz	5 ms	40 ms	60 ms	80 ms	100 ms	200 ms	600 ms	Moyenne
MAR	4,16	4,33	4,95	5,61	5,57	5,37	6,05	4,93
Rmean	4,16	0,73	2,16	3,84	4,08	3,86	3,60	2,77
Rmax	6,86	6,45	8,14	10,18	10,04	9,33	10,18	10,18
Rstd	2,06	4,83	5,19	5,57	5,38	5,22	6,19	4,80
7 000 Hz	5 ms	40 ms	60 ms	80 ms	100 ms	200 ms	600 ms	Moyenne
MAR	4,45	5,61	5,17	6,18	4,22	5,86	3,46	4,90
Rmean	4,45	5,61	5,17	6,18	4,13	5,86	1,79	4,66
Rmax	6,57	7,87	7,77	9,28	7,58	8,35	5,56	9,28
Rstd	1,84	1,70	2,37	3,18	3,17	2,49	3,66	2,59

Tableau A-7.3 : Récapitulatifs des indices afin de pouvoir évaluer de la précision quant à l'efficacité du N30ms pour estimer la sonie mesurée de sons impulsionnels de synthèse.

ESI								
450Hz	5 ms	40 ms	60 ms	80 ms	100 ms	200 ms	600 ms	Moyenne
MAR	7,35	2,99	2,42	2,82	2,30	2,61	3,48	3,42
Rmean	7,35	1,64	-1,48	2,13	-0,80	-1,32	2,69	1,46
Rmax	19,17	11,64	6,32	12,64	5,25	5,61	6,92	19,17
Rstd	5,33	4,45	2,91	4,29	2,88	2,91	3,21	3,71
1 000 Hz	5 ms	40 ms	60 ms	80 ms	100 ms	200 ms	600 ms	
MAR	7,43	3,71	3,88	5,08	5,36	5,42	8,30	5,60
Rmean	7,43	3,37	3,82	5,08	5,29	5,42	8,30	5,53
Rmax	18,88	6,85	7,82	9,61	9,37	9,52	13,45	18,88
Rstd	4,26	3,13	2,98	3,41	3,51	3,63	4,78	3,67
7 000 Hz	5 ms	40 ms	60 ms	80 ms	100 ms	200 ms	600 ms	
MAR	9,29	8,52	8,16	7,78	5,67	7,52	9,04	8,00
Rmean	9,29	8,52	8,16	7,78	5,67	7,52	9,04	9,04
Rmax	16,79	17,86	11,91	11,37	9,97	9,76	10,90	17,86
Rstd	2,69	3,16	1,48	2,26	2,65	1,82	1,62	2,24

Tableau A-7.4 : Récapitulatifs des indices afin de pouvoir évaluer de la précision quant à l'efficacité de l'ESI pour estimer la sonie mesurée de sons impulsionnels de synthèse.

ESI_{bc}								
450Hz	5 ms	40 ms	60 ms	80 ms	100 ms	200 ms	600 ms	Moyenne
MAR	4,85	2,44	3,06	3,04	2,51	2,66	0,90	2,78
Rmean	3,99	-0,59	-2,12	-1,53	-2,08	-2,61	-0,05	-0,71
Rmax	13,60	5,77	4,24	6,40	4,28	5,03	2,54	13,60
Rstd	5,34	3,00	2,55	3,21	1,91	1,41	1,18	2,66
1000 Hz	5 ms	40 ms	60 ms	80 ms	100 ms	200 ms	600 ms	Moyenne
MAR	6,37	1,35	1,61	2,36	2,15	1,68	3,12	2,66
Rmean	6,23	0,91	1,30	2,36	2,15	1,66	3,12	2,53
Rmax	18,97	6,50	4,85	4,24	5,76	3,39	6,16	18,97
Rstd	5,97	2,21	1,57	1,14	1,53	1,27	1,90	2,23
7000 Hz	5 ms	40 ms	60 ms	80 ms	100 ms	200 ms	600 ms	Moyenne
MAR	6,24	4,99	3,51	3,27	2,66	2,91	2,07	3,66
Rmean	3,84	2,48	1,06	1,49	-0,91	0,38	-1,86	0,93
Rmax	13,84	11,78	8,11	8,35	5,95	6,90	5,12	13,84
Rstd	6,60	5,54	4,15	3,80	3,41	3,60	2,10	4,17

Tableau A-7.5 : Récapitulatifs des indices afin de pouvoir évaluer de la précision quant à l'efficacité de l'ESI_{bc} pour estimer la sonie mesurée de sons impulsifs de synthèse.

Le modèle de Zwicker 1999 (N30ms) donne globalement des niveaux d'isophonie supérieurs à ceux mesurés. De plus le MAR est en moyenne supérieur à 4 phons et présente un Rmax de l'ordre de 10 phons.

En ce qui concerne les sons impulsifs de synthèse, et ce quel que soit la fréquence et la durée, c'est l'ESI_{bc} qui donne en moyenne les écarts les plus faibles par rapport aux niveaux d'isophonie mesurés (MAR moyen de l'ordre de 3 phons). Il présente toutefois pour les sons de 5 ms une erreur maximale importante (Rmax de l'ordre de 19 phons). On observe que l'ESI_{bc} améliore considérablement l'ESI qui présente un écart moyen supérieur à 5 phons.

Les résultats donnés par l'ESI_{bc} sont ici comparés aux valeurs du N30ms et de l'ESI sur des sons « naturels ». La figure A-7.1 représente les niveaux d'isophonie mesurés et calculés. Nous avons relié les points expérimentaux par des segments pour une meilleure visibilité des différences entre les modèles.

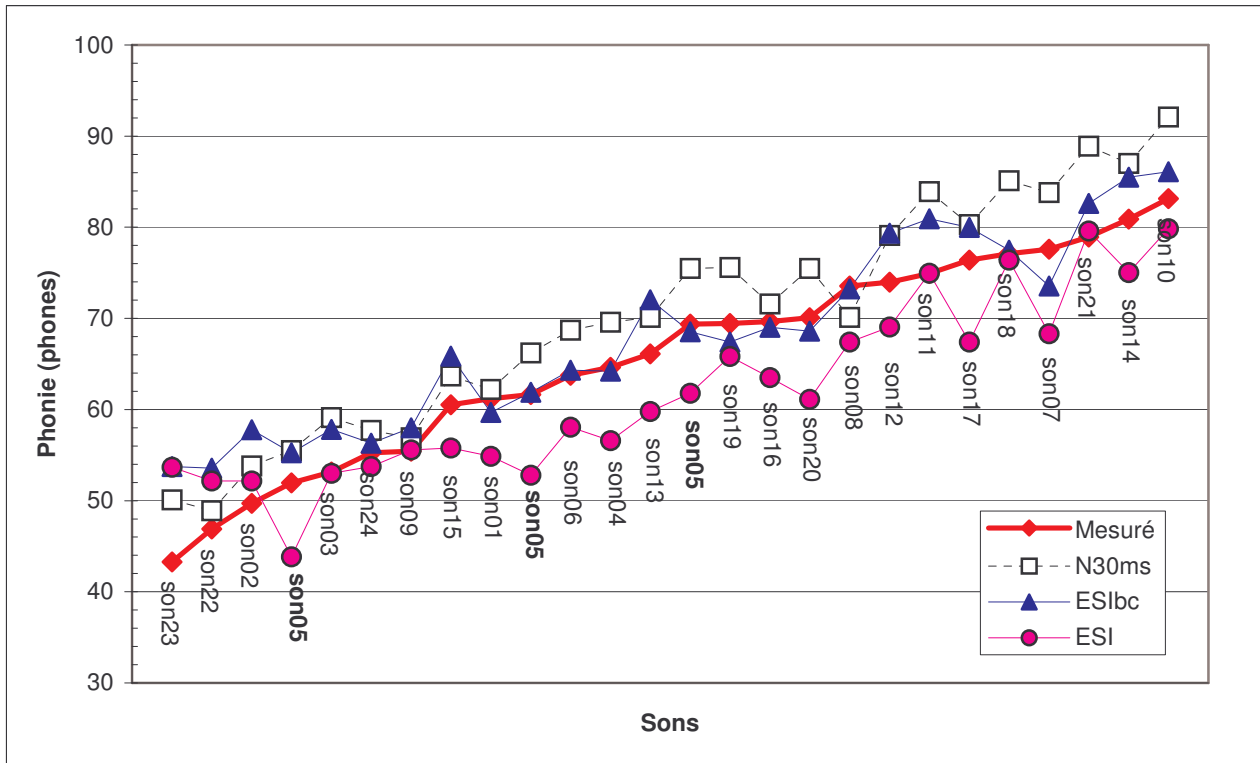


Figure A-7.1 : Niveau d'isononie en phones des impulsions mesurés (losange rouge) et calculés par le modèle de Zwicker 1999 (N30ms), de l'ESIbc et de l'ESI.

Le tableau A-7.6 récapitule les indices permettant de quantifier les écarts entre les niveaux d'isononie calculés et mesurés en ce qui concerne les sons impulsionnels « naturels ».

	N30ms	ESIbc	ESI
MAR	5,25	3,92	6,10
Rmean	4,66	2,96	-3,69
Rmax	9,99	10,49	10,52
Rstd	2,71	2,89	2,74

Tableau A-7.6 : Récapitulatifs des indices afin de pouvoir évaluer de la précision quant à l'efficacité du N30ms, de l'ESI_{bc} et de l'ESI pour estimer la sonie mesurée de sons impulsionnels « naturels ».

De la même manière que pour les sons de synthèse, l'ESI_{bc} donne les écarts les plus faibles (MAR inférieur à 4 phones) et améliore considérablement l'ESI (MAR de l'ordre de 6 phones).

Glossaire

Bande critique : À intensité constante, la sonie augmente avec la largeur spectrale. Mais cette augmentation n'apparaît que lorsque la largeur dépasse une certaine valeur dépendant de la fréquence. Cette largeur spectrale est appelée bande critique. La notion de bande critique modélise le mécanisme de filtrage de l'oreille interne selon Zwicker.

ERB : Equivalent Rectangular Bandwidth. Au lieu d'utiliser le modèle de bandes critiques, Moore et Glasberg (1983) introduisent la notion de filtre rectangulaire équivalent (ERB), dont la largeur dépend aussi de la fréquence.

ESI : Estimateur de sonie d'impulsion

Estimateur : Un estimateur permet de calculer la sonie (ou le niveau d'isotonie) à partir d'un algorithme basé sur des paramètres physiques du signal.

Fonction de sonie : La fonction de sonie est la relation entre l'intensité (ou la pression acoustique) et la sonie. Elle a été établie expérimentalement par Stevens, en faisant estimer la sonie d'un son de 1000 Hz à différents niveaux de pression acoustique. La relation entre la sonie (S) en sones et la pression (P) en Pascals, ou l'intensité (I) ($I = P^2 / \rho c$: avec ρ la densité de l'air et c la célérité du son dans l'air) en watts par mètre carré, est une simple fonction puissance pour des niveaux supraliminaires (relation 1).

L'équation de la fonction de sonie d'un son pur de 1000 Hz, pour des niveaux d'isotonie supérieurs à 30 phones, est :

$$S \text{ (sones)} = k P^{0,6} = 1 I^{0,3} \quad (1)$$

Indicateur : Valeur extraite d'un modèle ou d'un estimateur permettant de calculer la sonie globale de son non stationnaire et impulsionnel.

Intégration temporelle : La sonie augmente régulièrement lorsque la durée d'un son d'amplitude constante augmente, de quelques millisecondes jusqu'à une valeur critique, appelée durée critique, qui varie de 50 à 400 ms selon les auteurs. Ce phénomène est appelé intégration temporelle.

Ligne isotonique : Une ligne isotonique représente le niveau que doit avoir un son pur pour conserver une sonie constante en fonction de la fréquence.

Modèle de sonie : Un modèle de sonie permet de calculer la sonie (ou le niveau d'isotonie) à partir d'un algorithme basé sur la modélisation du système auditif.

Niveau d'isotonie : Le niveau d'isotonie d'un son (son test) est le niveau en dB SPL que doit avoir un son pur de référence de 1 000 Hz pour que les deux sons (son test et son de référence) aient la même sonie. Le niveau d'isotonie s'exprime en phones. L'échelle des phones, établie à partir de méthodes de comparaison (Fletcher et Munson, 1933), coïncide avec l'échelle des dB SPL pour un son pur de 1000 Hz. Par exemple, un son pur à 100 Hz de 50 dB a la même sonie qu'un son pur de 1 kHz à 40 dB, soit un niveau d'isotonie de 40 phones.

Relation sones-phones : La relation se déduit de la fonction de sonie d'un son pur stationnaire à 1 kHz car à 1 kHz le niveau d'isotonie en phone correspond aux dB SPL. Elle est égale à :

$$\text{Phones} = 40 + 10 \log_2(\text{Sones}) \quad (2)$$

Sonie : D'après la norme ISO R532, c'est la désignation numérique de la force d'un son, proportionnelle à sa grandeur subjective estimée par des observateurs normaux. En d'autres termes, c'est l'intensité d'un son telle qu'on la perçoit. Son unité, le sone, est basée sur une échelle sensorielle établie à partir de méthodes directes (Stevens, 1956) qui consistent à demander aux auditeurs de choisir des nombres proportionnels à la sonie des sons présentés. Par convention, la valeur de 1 sone est la sonie d'un son pur de 1000 Hz à 40 dB SPL. Un son de 2 sones est perçu deux fois plus fort qu'un son de 1 sone.

Son impulsionnel : La norme NF (S31-010, décembre 1996) définit un bruit impulsionnel comme un "bruit consistant en une ou plusieurs impulsions d'énergie acoustique, ayant chacune une durée inférieure à environ 1 s et séparée(s) par des intervalles de temps de durée supérieure à 0,2 s". Nous avons précisé cette définition suite à des observations de sons naturels que nous avons pu recenser dans notre environnement sonore. En effet, les sons impulsionnels tels que nous les définissons sont caractérisés par une attaque rapide immédiatement suivie d'une décroissance d'amplitude de forme exponentielle.

Durée de décroissance : Durée nécessaire pour qu'un signal passe de 90 à 10 % de la valeur maximale absolue de la pression acoustique. Par abus de langage nous l'appellerons « temps de descente »

Références

Berglund B., "Relationship between loudness and annoyance for ten community sounds", *Environ. Int.*, 16, 523-531, 1990.

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

Boulet I., "La sonie des sons impulsionnels : Perception, Mesures et Modèles", Thèse de doctorat de l'Université Aix Marseille II, spécialité Mécanique, option Acoustique, 2005.

Boulet I, Marozeau J. et Meunier S., "Temporal summation of loudness as a function of frequency and temporal pattern", 33rd International Congress and Exposition on noise control engineering (Internoise), Prague, République Tchèque, 2004

Boussard P. et Boulet I., "Simulation et perception des environnements sonores : Comparaison des techniques de reproduction et influence de la sonie des bruits de chantier", *Acoustique & Techniques* n°26, p 26-29, 2001.

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.

Cops A. et Myncke H., "Mesures subjectives et objectives du niveau de sonie des signaux sonores impulsifs : un projet de recherche international", *Revue d'Acoustique* n°37, 98-104, 1976.

Fastl H., "The psychoacoustics of sound-quality evaluation", *Acustica* 83, 754-764, 1997.

Fletcher H. et Munson W.A., "Loudness, its definition, measurement and calculation", *J. Acoust. Soc. Am.*, 5, 82-108, 1933.

Gabrielson A. et Sjögren H., "Perceived sound quality of sound-reproducing systems", *J. Acoust. Soc. Am.* 65 (4), 1019-1033, 1979.

Gelfand, S.A., Hearing, "An introduction to psychological and physiological acoustics", Third Edition, Marcel Dekker, Inc., USA, 1998.

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.

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

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.

Gustaffson B., "The loudness of transient sounds as a function of some physical parameters", *J. Sound. Vib.*, 37(3), 389-398, 1974.

- Jestead, W., "An adaptative procedure for subjective judgements", *Percept. Psychophys.* 28, 85-88, 1980.
- Levitt H., "Transformed up-down procedures in psychoacoustics", *J. Acoust. Soc. Am.* 49, 467-477, 1971.
- Marozeau J., Boulet I., Gagneux F., Boussard P., Meunier S., "Evaluation d'estimateurs de sonie", *Acoustique et Techniques n°39*, p 27-34, 2005.**
- Marvit P., Florentine M. et Buus S., "A comparison of psychophysical procedures for level-discrimination thresholds", *J. Acoust. Soc. Am.* 113 (6), 3348-3361, 2003.
- Meunier S., Boulet I. et Rabau G., "Loudness of impulsive sound", actes du 17th International Congress on Acoustics (ICA), Rome, Italie, 2-7 septembre 2001.
- Meunier S. et Marchioni A., "Relations entre les caractéristiques physiques et les attributs subjectifs des signaux acoustiques", rapport final, Programme interdisciplinaire de recherche sur la ville, 1998.
- Meunier S., Santon F., Boussard P., Marchioni A., Rabau G. et Régat X., "Etude perceptive de circulation routière et ferrovière", Contrats n° 98111 du Ministère de l'Aménagement du Territoire et de l'Environnement et n° 226 du Ministère de l'Équipement, du Transport et du Logement, juillet 2000.
- Moore B. C. J., Glasberg B. R. et Baer T., "A model for the prediction of thresholds, loudness and partial loudness", *J. Audio Eng. Soc.*, Vol.45, n°4, 224-240, 1997.
- Moore B. C. J. et Glasberg B. R., "Suggested formulae for calculating auditory-filter bandwidths and excitation patterns", *J. Acoust. Soc. Am.*, 74, 750-753, 1983.
- Ohta K., Kuwano S. et Namba S., "Sound quality of impulsive sounds in relation to their physical properties", *Technoloy Reports Of The Osaka University*, 49, n° 2360, 189-199, 1999.
- Parizet E., Deumier S. et Milland E., "Car road noise annoyance : Significant timbre parameters and inter-individual variability", *Acustica – acta acustica*, 82, Suppl. 1, S 216, 1996.
- Pedersen O. J., Lyregaard P. E. et Poulsen T. E., "The round robin test on evaluation of loudness level of impulsive noise", *Acoustics Laboratory, Technical University of Denmark, Report n°22*, 1977.
- Scharf B., "Loudness", in *Handbook of perception*, edited by E. C. Carterette and M. P. Friedman, Academic Press New York, 187-242, 1978.
- Scharf B. et Buus S., "Stimulus, Physiology, Thresholds", in *Handbook of perception and Human Performance Vol. I*, edited by Boff K. R., Kaufman L. And Thomas J. P., Academic Press New York, 1986.
- Schlauch, R. S. et Wier, C.C., "A method for relating loudness matching and intensity-discrimination data", *J. Speech Hear. Res.* 30, 13-20, 1987.
- Stevens S.S., "The direct estimation of sensory magnitudes", *Loudness, Am. J. Psychol.*, 69, 1-25, 1956.

- Stevens S.S., "Concerning the form of the loudness function", J. Acoust. Soc. Am., 29, 603-606, 1957.
- Zölzer U., "Digital audio signal processing", John Wiley et Son Ltd, 1997.
- Zwicker E., "Übe psychologishe und methodishe grundlagen der lautheit, Acustica", Vol. 8, 237-258, 1958.
- Zwicker E. et Fastl H., "Psychoacoustics: Facts and models", 2nd Edition, Springer-Verlag, Berlin, 1999.
- Zwicker E. et Feldtkeller R., "Psychoacoustique", Masson, Paris, 1981.