

Mémoire de fin d'études
Promotion 2018, spécialité Son

Représentation de protéines par la sonification d'objets 3D

Soutenu le 13 juin 2018 par
Valère RAIGNEAU

Directeur de mémoire interne : Alan BLUM

Directrice de mémoire externe : Tifanie BOUCHARA

Rapporteur : Sylvain LAMBINET

Intervenant extérieur : Matthieu MONTES (pour le projet VIDOCK)

Organisme impliqué : Le CNAM (Conservatoire National des Arts et Métiers)

Remerciements

Je tiens d'abord à remercier tout particulièrement Tifanie Bouchara et Alan Blum pour m'avoir accompagné et conseillé au mieux tout au long de ce travail de recherche. Merci aussi évidemment à Matthieu Montes et toute son équipe pour m'avoir si bien accueilli au sein du projet VIDOCK. Enfin, merci à Thomas Constant, William Rao Fernandes, Guillaume Levieux et Stéphanie Mader pour avoir fait de ce stage au CNAM une expérience de travail plaisante et conviviale.

Je remercie également Théo Serror pour ses nombreux conseils et son optimisme de chaque instant. Pour finir, un immense merci à Mia Valomy pour son soutien sans faille durant toute cette période.

Résumé

Ce mémoire propose une entrée approfondie dans le domaine de la sonification de données, une jeune discipline dont le développement s'est particulièrement accru au cours des vingt dernières années. Notre travail de recherche s'inscrit spécifiquement au sein du projet VIDOCK, qui cherche à proposer aux biologistes plusieurs niveaux de représentation de protéines. À terme, l'objectif est de pouvoir identifier des similarités ou au contraire des singularités locales sur et entre différentes macromolécules. Le projet se consacre donc en particulier à la représentation visuelle de la surface des protéines en trois dimensions. Or, il a été montré que l'utilisation de la sonification permettait notamment de renforcer les représentations visuelles. Nous avons donc vu dans ce projet l'opportunité de mettre à profit les apports de la modalité auditive. Nous détaillons ainsi dans le présent manuscrit la conception et l'intégration de notre représentation sonore de protéines par la sonification d'objets 3D.

Mots-clés : Sonification de données • Interface sonore • Modalité auditive • Spatialisation • Visualisation • Protéine • Objet 3D • Surface • Forme

Abstract

This thesis proposes a deep insight into data sonification - a discipline that has gained significant interest over the last few years. Our work is conducted within the VIDOCK project which intends to propose to biologists several levels of displaying proteins. The long-term objective of this project is to identify local similarities or dissimilarities within and between various proteins. Accordingly, it focuses on visual representation of surface in three dimensions. However, it has been shown that having an auditory modality input would enhance visual display. It is why we see an opportunity in using sonification for this project. Therefore, we present in this manuscript the theoretical and practical conception of our own strategy of sonification of proteins as 3D objects.

Keywords: Sonification • Auditory display • Auditory modality • Spatialization • Visualization • Protein • 3D objet • Surface • Shape

Table des matières

Introduction	8
1. La sonification de données	10
1.1. L'émergence d'une discipline.....	10
1.1.1. Définition(s).....	10
1.1.2. Pourquoi sonifier ?.....	12
1.1.3. Les prémices	14
1.1.4. L'intégration du son dans les interfaces multimodales	15
1.2. Taxinomie des différentes fonctionnalités de la sonification	16
1.2.1. Notification	16
1.2.2. Indication de statut ou de progrès	17
1.2.3. Exploration de données.....	17
1.2.4. Autres fonctionnalités	19
1.3. Différentes techniques et approches en sonification	21
1.3.1. Audification.....	21
1.3.2. Sonification par mapping de paramètres.....	22
1.3.3. Sonification par modélisation	24
1.3.4. Sonification par signaux auditifs	25
1.4. Conclusion.....	26
2. Prérequis à la sonification de protéines	27
2.1. Rappel sur les protéines.....	27
2.1.1. La structure des protéines	27
2.1.2. Représentations et visualisation de la structure des protéines	29
2.1.3. Limites des représentations visuelles et nécessité de la sonification.....	33
2.2. La sonification de molécules	34
2.2.1. Musique et importance du facteur séquentiel	34
2.2.2. Vers une sonification paramétrique.....	37
2.2.3. Le développement de l'interactivité.....	44

2.3. La sonification d'objets 3D	46
2.2.1. Coupe transversale d'un volume.....	47
2.2.2. Le voxel	48
2.2.3. Exploration d'un nuage de points	52
2.4. Le son 3D	54
2.4.1. La localisation auditive et ses performances	54
2.4.2. L'Ambisonic.....	55
2.4.3. La synthèse binaurale.....	57
2.5. Conclusion.....	58
3. Conception d'une nouvelle représentation sonore	59
3.1. Présentation générale du projet	59
3.1.1. Idée principale.....	59
3.1.2. Hypothèse de travail	60
3.1.3. Limites et contraintes.....	60
3.2. Stratégie de mapping des paramètres	63
3.2.1. Définition	63
3.2.2. Échelle de mapping.....	65
3.2.3. Mapping des paramètres sonores	66
3.3. Conclusions provisoires	70
4. Implémentation de notre proposition de sonification	71
4.1. Description des ressources utilisées	71
4.1.1. Le format OBJ.....	71
4.1.2. Max 7	75
4.2. Patch de visualisation	77
4.3. Patch de sonification	79
Conclusion	82
Bibliographie	84
Annexes	96
A. Patch de visualisation	96
B. Patch de sonification	97

Introduction

Ce mémoire s'inscrit dans le cadre du projet européen¹ VIDOCK² (pour *Visualization and DOCKing software*), en cours au sein du laboratoire GBA (Génomique, Bio-informatique et Applications) du CNAM (Conservatoire National des Arts et Métiers). Ce projet a pour principal objectif l'étude de nouvelles méthodes de représentation de la surface des protéines, notamment pour du *docking* moléculaire³ et de la classification.

En effet, il existe un lien fort entre les fonctions d'une protéine et sa structure (cette dernière renseigne par exemple sur les sites de fixation potentiels du ligand⁴). Or, c'est cette même structure qui détermine la (ou les) forme(s) que prendra la protéine. Il paraît donc cohérent de se focaliser spécifiquement sur les différentes représentations possibles de la forme extérieure des protéines, et en particulier leur surface (puisque les molécules interagissent entre elles en surface).

Les modélisations sur lesquelles travaille l'équipe de VIDOCK sont pour l'instant uniquement visuelles et consistent principalement en des projections sur deux ou trois dimensions. Or, dès les années 60, différentes études portant sur des données sismologiques ont commencé à suggérer que la capacité des sujets à établir des classifications en utilisant une présentation aurale était comparable à celle des sujets utilisant une présentation visuelle [Speeth, 1961 ; Frantti et Leverault, 1965]. De plus, il a ensuite été montré que l'utilisation simultanée des deux modalités fournit un avantage certain par rapport à l'utilisation d'une des deux modalités séparément. Il serait donc intéressant de se servir des atouts du système auditif par rapport au système visuel pour améliorer la représentation de ces molécules et profiter des apports de la multimodalité. C'est pourquoi on se propose de sonifier les volumes générés par la surface des protéines.

¹ Le projet est en effet doté d'une bourse délivrée par l'ERC (*European Research Council*).

² Pour plus d'informations, voir le site web : <http://www.vidock.comsource.fr>.

³ Le *docking* moléculaire est une méthode qui consiste à déterminer l'orientation préférée d'une première molécule vers une seconde afin de former un complexe stable. On pourrait traduire le terme en français par « connexion » ou plus littéralement « amarrage », mais la terminologie anglaise reste majoritairement utilisée dans la littérature francophone sur le sujet.

⁴ Un ligand désigne un certain type de molécules qui interagit avec une protéine de manière spécifique et joue donc un rôle dans ses fonctions.

L'idée générale est donc de développer puis proposer aux chercheurs et aux étudiants une interface de sonification qui puisse offrir une nouvelles manière d'observer et d'interagir avec des protéines, pour potentiellement faire apparaître des catégorisations inédites ou au moins trouver un angle d'approche différent qui amènerait à des idées originales. Cela va dans le sens du projet VIDOCK, qui cherche à proposer aux biologistes de nouveaux niveaux d'abstraction et de représentation ; le but étant, à terme, de pouvoir identifier des similarités ou dissemblances locales sur les protéines.

De fait, comment envisager la représentation de protéines par du son, et quelle stratégie mettre en place pour proposer *in fine* une sonification adaptée à leur forme ?

Pour répondre à cela, nous commencerons dans un premier temps par définir les contours de la sonification de données, en passant par l'étude de ses enjeux et des différentes approches existantes. Un état de l'art raisonné de la sonification de molécules et d'objets 3D permet ensuite de justifier de la pertinence de la partie pratique du mémoire, en remettant en perspective son originalité et ses possibles apports dans le domaine. Subséquemment, nous détaillons dans une troisième partie la conception théorique de notre propre stratégie de sonification. Son implémentation pratique sous forme d'interface est enfin détaillée dans le dernier chapitre de ce mémoire.

1. La sonification de données

Dans cette première partie, nous allons tout d'abord définir ce qu'est la sonification de données. Nous étudierons ensuite les différentes techniques les plus couramment mises en œuvre pour transmettre des informations sous forme sonore, avant d'aborder enfin les principales utilisations que l'on peut en faire.

1.1. L'émergence d'une discipline

1.1.1. Définition(s)

Le terme « sonification » est un néologisme provenant du mot anglais éponyme. Il désigne une discipline scientifique en même temps qu'une technique. Il a été introduit en 1992, à l'occasion de la fondation de l'*International Community for Auditory Display* (ICAD) par Gregory Kramer⁵. L'ICAD est depuis devenu un réseau influent dans le domaine de la sonification⁶, doublé d'une rencontre éponyme internationale qui réunit annuellement la majorité des acteurs concernés du milieu académique. Ils y partagent leurs recherches ainsi que leurs productions.

En 2011, un premier ouvrage collectif universitaire de synthèse est publié : *The Sonification Handbook*⁷ [Hermann et al., 2011]. Faisant un état des lieux de la discipline, ce manuel en constitue donc un ouvrage de référence. Il s'appuie en effet sur vingt années de recherches pour s'efforcer de circonscrire le domaine de manière exhaustive. Au début de son introduction, on peut y trouver les deux définitions suivantes :

Auditory Display encompasses all aspects of a human-machine interaction system, including the setup, speakers or headphones, modes of interaction with the display system, and any technical solution for the gathering, processing, and computing necessary to obtain sound in response to the data. In contrast, Sonification is a core component of an auditory display: the technique of rendering sound in response to data and interactions.

Hermann, Hunt, Neuhoff, *The Sonification Handbook* (2011), p. 1

⁵ Gregory Kramer est notamment chercheur, auteur et professeur. Il est connu pour être l'une des figures fondatrices du concept d'*auditory display*. Il a en effet publié le premier livre relatif à ce sujet [Kramer, 1994].

⁶ L'ICAD possède notamment un site web qui propose du contenu sur le sujet en libre accès : <http://www.icad.org/>

⁷ Il est intégralement disponible en ligne à cette adresse : <http://sonification.de/handbook/index.php/chapters/>

Cette citation montre bien l'articulation des deux concepts fondamentaux que représentent l'*auditory display* et la sonification. Le premier, qu'on peut traduire littéralement par « affichage auditif » (on lui préférera la terminologie « interface sonore » dans la suite du manuscrit), décrit l'ensemble du dispositif technique qui permet la captation et/ou la construction des données, leur transformation, et enfin leur diffusion sous forme acoustique (voir la figure 1.1). Il englobe ainsi tous les aspects d'un système de rendu sonore : le système de diffusion du son (haut-parleurs, casques), le type de rendu utilisé (stéréophonie, synthèse binaurale, spatialisation ambisonique, etc.) ainsi que l'ensemble des solutions techniques pour la collecte, le traitement et les calculs nécessaires à l'obtention de son en réponse à un jeu de données. La sonification fait justement partie de ces solutions techniques.

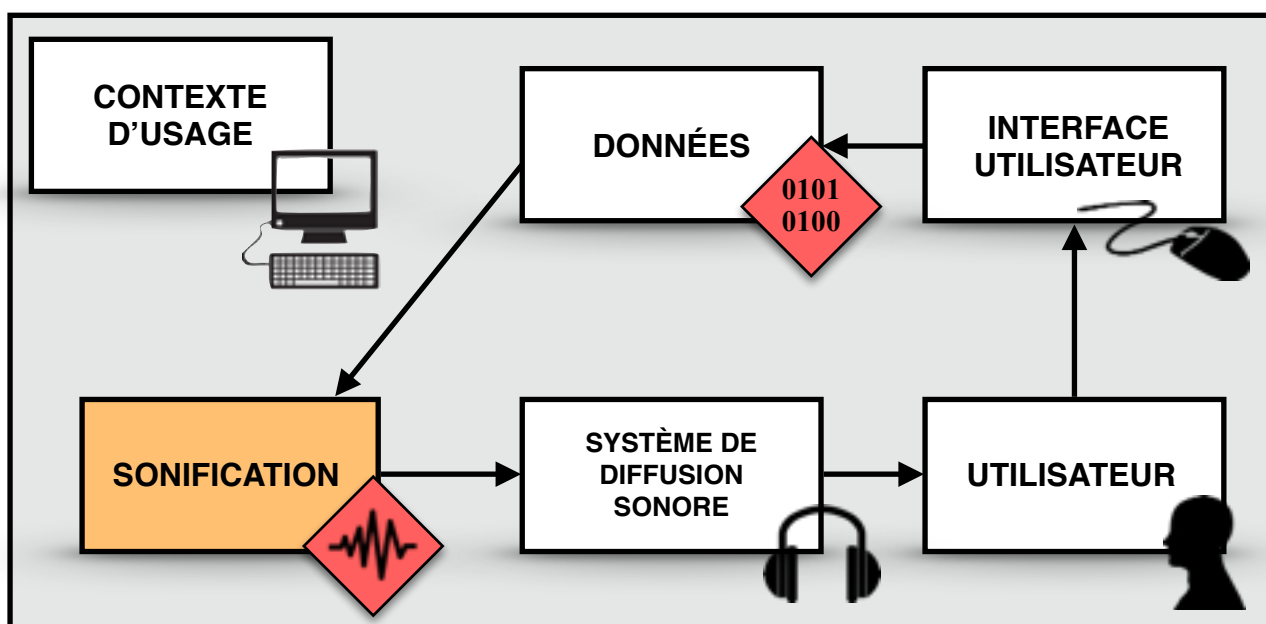


Fig. 1.1 : Les différents composants d'une interface sonore de sonification

En effet, la sonification désigne les divers systèmes construits⁸ qui rendent perceptibles des données grâce à leur transposition dans le domaine sonore. On peut noter que les publications récentes retiennent plus volontiers la définition proposée originellement par [Kramer et al., 1999] :

Sonification is the use of non-speech audio to convey information. More specifically, sonification is the transformation of data relations into perceived relations in an acoustic signal for the purposes of facilitating communication or interpretation.

Kramer et al., *Sonification Report: Status of the Field and Research Agenda* (1999), p. 4

⁸ On peut également désigner par le terme « sonification » un dispositif purement conceptuel.

« La sonification est donc l'utilisation de signaux sonores non-verbaux dans le but de véhiculer de l'information. Plus précisément, elle correspond à la transformation des relations entre des données afin que ces mêmes relations deviennent perceptibles sous forme sonore, dans le but d'en faciliter la communication ou l'interprétation. » En d'autres termes, la sonification consiste à produire des objets sonores qui rendent compte des relations pouvant exister entre des données abstraites, visuelles, cartographiques ou autres. En exploitant la modalité auditive, elle permet notamment de limiter la surcharge d'informations fournies par les interfaces graphiques. On pourrait donc dire que la fonction de la sonification est de produire une forme de « visualisation auditive » de toutes sortes de données grâce à un ensemble de dispositifs techniques et conceptuels de représentation sonore.

Il est important de noter que cette définition de la sonification exclue l'utilisation de messages sonores exclusivement verbaux pour transmettre des données. D'après [Hermann, 2008], la principale raison de cette exclusion est l'utilisation encore majoritaire qu'il était fait de la parole comme premier vecteur d'information dans les interfaces sonores de l'époque. En réaction à cela, Kramer et al. [1999] ont voulu prendre leur distance avec la voix parlée ; surtout qu'il y avait la nécessité pour cette discipline de dessiner des limites claires à un champ de recherche déjà très vaste, sans avoir donc à s'encombrer de considérations linguistiques et sémiotiques.

1.1.2. Pourquoi sonifier ?

Les différentes motivations qui justifient l'utilisation de signaux sonores plutôt que d'autres modalités (visuelle ou haptique⁹ par exemple) pour présenter des informations ont été discutées en détail dans plusieurs études [Sanders et McCormick, 1993 ; Kramer, 1994 ; Kramer et al., 1999 ; Hermann et al., 2011]. On peut en retenir trois raisons principales.

Premièrement, la sonification exploite les capacités de notre système auditif à analyser et identifier avec une grande précision les variations temporelles et/ou fréquentielles d'un signal sonore [Bregman, 1994 ; McAdams et Bigand, 1993 ; Moore, 1997 ; Kramer et al., 1999].

⁹ Qui concerne le sens du toucher.

Deuxièmement, l'utilisation du son permet de grandement soulager la modalité visuelle. On parle alors de « répartition modale ». Introduit par [André, 2000], ce concept décrit l'utilisation spécifique d'une modalité en particulier pour mettre en relief des informations distinctes au sein d'un environnement multimodal. Kristine Jørgensen a notamment montré que l'utilisation du son dans une interface multimodale pouvait aider à augmenter l'attention de l'utilisateur sur une tâche [Jørgensen, 2006]. Il est donc important d'optimiser cette répartition en fonction des contraintes liées à la tâche à effectuer, afin de tirer le meilleur parti possible de chacune des modalités. Dans notre cas, la sonification permet d'éviter une surcharge visuelle lors de la visualisations graphiques de données complexes, et ce dans deux situations principales : lorsque notre vue est déjà focalisée sur une autre tâche [Fitch et Kramer, 1994] et/ou lorsque beaucoup d'informations doivent être affichées en même temps [Brewster, 1997 ; Brown et al., 1989]. On peut par exemple penser aux *smartphones*, tablettes et autres gadgets technologiques mobiles de taille réduite : la sonification permet d'appréhender des informations sans que l'utilisateur n'ait besoin de regarder l'écran et/ou permet d'en compléter la vision lorsque ce dernier est trop petit pour afficher toutes les informations [Brewster et Murray, 2000]. La sonification peut également aider à suppléer l'absence de modalité visuelle, par exemple pour les non-voyants ou des pompiers aveuglés par la fumée d'un incendie [Fitch et Kramer, 1992].

Enfin, Moore [1997] a montré que notre perception auditive nous permet d'écouter, de surveiller et de traiter plusieurs flux sonores différents en même temps, et donc de séparer efficacement l'information que chacun convoie. Il y a deux raisons principales à cela :

- D'abord, notre cerveau est capable d'analyser des informations sonores provenant de toutes les directions, avec plus ou moins de précision selon la position et la définition de la source. En l'occurrence, notre capacité de localisation auditive est bien meilleure dans le plan horizontal qu'en élévation et en distance [Blauert, 1997]. Nous précisons cela dans la sous-section 4.1.3.
- De plus, nous pouvons discerner les composantes spectrales de différents flux sonores superposés. Ainsi, nous sommes capables de diriger notre attention sur un flux sonore en particulier dans une ambiance bruyante et chargée, tout en restant sensible dans une certaine mesure aux autres signaux sonores extérieurs [Bregman, 1994 ; Bronkhorst 2000]. Par exemple, il est aisé de concentrer son attention sur un instrument en particulier dans un

orchestre, ou d'écouter une conversation précise dans une pièce bruyante. C'est d'ailleurs de là que vient le nom de ce phénomène : l'effet « *cocktail party* ».

La modalité auditive paraît donc particulièrement appropriée pour représenter des informations dépendantes du temps, ainsi que des motifs complexes.

1.1.3. Les prémices

On comprend au regard des définitions données précédemment que la sonification est contemporaine de l'expansion des technologies du numérique et de l'informatique. Cette discipline est donc éminemment reliée au contexte technologique actuel. Cependant, bien avant l'émergence de ce contexte, on peut retrouver ce qui s'apparente à des formes de sonification parmi les plus anciennes productions humaines, comme la musique et le langage¹⁰ par exemple. En effet, même si l'écriture textuelle et la partition sont d'abord les modalités d'un enregistrement qui advient grâce à l'inscription sur support, ils entretiennent également tous deux une relation entre des données codifiées et des sons. Ces derniers, qu'ils soient verbaux ou musicaux, rendent audibles et signifiants les codes auxquels ils se réfèrent.

Mais, si l'apparition de la sonification comme champ de recherche est corrélative des développements de l'informatique, on peut faire remonter sa fonction exploratoire aux innombrables dispositifs auditifs expérimentés à partir du XIX^{ème} siècle. Deux de ces appareils en particulier symbolisent parfaitement les deux orientations majeures qu'a pris la sonification depuis son avènement : le stéthoscope (invention de René Laennec traditionnellement datée de 1816) et le compteur Geiger (imaginé vers 1913 par Hans Geiger)¹¹.

On a d'un côté une relation indicielle, et de l'autre une relation symbolique. Le stéthoscope permet en effet la production et l'exploration d'une image auditive du cœur, des poumons ou de la circulation sanguine grâce à l'amplification des sons internes de notre corps. Le compteur Geiger, pour sa part, rend compte de l'intensité d'une énergie ionisante en mettant en correspondance la mesure de ses rayonnements radioactifs avec une impulsion sonore. L'augmentation de la fréquence

¹⁰ Rappelons que la définition de la sonification exclut explicitement le langage de sa définition (cf. 1.1.1).

¹¹ Nous étudierons plus en détail les différents types de sonification qu'ils mettent chacun en place dans la sous-partie 1.3.

de répétition du son exprime alors l'augmentation proportionnelle de la radioactivité d'un environnement : les sons ne rendent pas littéralement audibles le rayonnement mesuré, mais ils signalent symboliquement sa présence et son intensité.

Cette méthode de sonification est bien connue et maintenant intégrée dans la vie quotidienne avec un certain nombre d'applications commerciales. On peut par exemple citer le radar de recul des voitures qui aide le conducteur à éviter les obstacles lorsqu'il se gare. Comme pour le compteur Geiger, plus la voiture se rapproche d'un objet et plus la fréquence de répétition du signal sonore augmente.

1.1.4. L'intégration du son dans les interfaces multimodales

La sonification de données a pris un essor considérable durant les vingt dernières années, essentiellement à cause de la quantité exponentielle de données produites par les calculateurs numériques dont le traitement ne peut plus se satisfaire des seules représentations visuelles et/ou linguistiques. Ces dernières sont en effet devenues insuffisantes ou inopérantes pour traiter cette masse d'informations temporelles. On est dès lors dans ce que la recherche en sonification appelle aujourd'hui communément l'« exploration de données » (*data mining* en anglais) ; où le son peut être utilisé pour révéler certaines particularités¹² au sein de la production massive d'informations qui touche tous les domaines de notre société.

De plus, le développement conjoint des ordinateurs personnels et de nombreux autres dispositifs électroniques a permis aux domaines de la sonification et du son spatialisé de faire leur apparition dans de multiples applications commerciales. Issues de la recherche en acoustique et en interfaçage homme-machine¹³, et premièrement utilisées dans des applications comme les jeux vidéos ou l'informatique, ces technologies commencent à émerger dans les systèmes permettant notamment d'assister un utilisateur dans la réalisation d'une tâche spécifique.

Aujourd'hui, la sonification se manifeste donc dans toutes nos interfaces électroniques (écrans tactiles, *smartphones*, tablettes, lecteurs de musique, jeux, appareils électroménagers,

¹² Comme des saillances ou des anomalies.

¹³ Une interface homme-machine (IHM) désigne l'ensemble des moyens et outils mis en œuvre afin qu'un opérateur humain puisse contrôler et communiquer avec une machine-outil ou une machine logicielle.

transports, médecine...) puisque ces dernières associent désormais un couplage sensoriel interactif visuel, tactile et auditif.

2. Taxinomie des différentes fonctionnalités de la sonification

Étant donné que la modalité auditive a des propriétés inhérentes qui peuvent se révéler bénéfiques pour l’affichage d’informations, nous allons examiner ici quelques types de fonctions que l’affichage auditif et la sonification peuvent remplir.

Buxton [1989], puis Edworthy [1998] ainsi que Walker et Kramer [2004] ont décrit ces différentes fonctions en terme de trois grandes catégories taxinomiques : (1) les alarmes, alertes et avertissements, (2) les messages d’état, de processus et de suivi d’une tâche, et enfin (3) l’exploration de données. Ce sont ces trois taxonomies fonctionnelles principales qui déterminent les applications de la sonification des données. À noter que Walker et Nees [2011] ont plus tard ajouté à cette liste une quatrième catégorie, centrée autour de l’interaction et l’immersion sonore.

1.2.1. Notification

Du simple avertissement à l’alerte, les notifications sonores permettent d’indiquer qu’un événement vient de se produire, est en train de se produire ou va se produire. Le message véhiculé par les alertes est généralement relativement simple et a pour but d’indiquer une information succincte à l’auditeur [Buxton, 1989 ; Sorkin, 1987]. Une sonnette est un bon exemple commun d’alerte sonore, cette dernière indiquant la présence d’une personne derrière la porte d’entrée. C’est le même type de signal qui est aussi utilisé par les micro-ondes pour indiquer la fin du temps de cuisson ou par les téléphones portables pour signifier l’arrivée d’un message texte.

Ces alarmes et avertissements sont des notifications sonores destinées à transmettre l’apparition d’une classe restreinte d’événements, le plus souvent urgent et défavorable. Ils exigent donc une réponse immédiate ou au moins une grande attention [Haas et Edworthy, 2006]. Pour ce type de notification, Spence et Driver [1997] ont montré que la modalité auditive capte plus facilement l’attention du sujet que la modalité visuelle et qu’elle permet d’éviter les problèmes dus à la limitation du champ de vision. Les alarmes doivent ainsi pouvoir véhiculer l’alerte, mais

également d'autres informations comme le niveau d'urgence ou le type de problème rencontré. De nombreuses études ont exploré la relation entre les sons utilisés et le degré d'urgence ressenti par les auditeurs [Edworthy et al., 1991 ; Suied et al., 2008].

1.2.2. Indication de statut ou de progrès

L'interface d'un système doit souvent fournir plusieurs informations simultanées sur l'état d'un processus ou d'une série d'événements en cours. En renseignant sur ces états et leurs variations temporelles, la sonification permet alors de décharger la modalité visuelle de l'utilisateur qui peut ainsi se concentrer sur d'autres tâches. Pour cela, elle tire notamment profit de la capacité de notre oreille à détecter de subtils changements au sein d'une série d'événements sonores [Kramer et al., 1999].

Les systèmes d'exploitation informatiques tels que Windows ou macOS¹⁴ ont depuis longtemps intégrés ce principe à leur propre interface, en créant de véritables « grammaires » sonores. Ces différents répertoires constituent un ensemble de notifications sonores, portant chacune un sens particulier relatif au statut d'une tâche précise. Par exemple, le son qui joue lorsque l'on met sous tension un ordinateur sert à confirmer le bon déroulement de la procédure de chargement du programme initial par le BIOS (*Basic Input Output System*). Ce son n'est pas vecteur d'alerte, mais bien de statut : il indique simplement que la tâche de démarrage est bien lancée.

Les différentes manières de créer un ensemble de notifications sonores ont été étudiées par [Gaver, 1986 ; Blattner et al., 1989 ; Dingler et al., 2008] et sont détaillées dans la sous-section 1.3.4.

1.2.3. Exploration de données

La sonification peut également permettre l'exploration de données complexes, évoluant temporellement sur de multiples dimensions. Cette fonction se distingue des précédentes par son utilisation générale : on s'attache à représenter une vue globale d'un jeu de données plutôt qu'un état momentané et succinct du système. De l'exploration de données scientifiques [Flowers et Hauer, 1993 ; Brown et al., 2003 ; Stockman et al., 2005] à l'écoute de données médicales

¹⁴ Anciennement Mac OS.

[Hermann et al., 2002 ; Baier et al., 2007 ; Pauletto et Hunt, 2009], ce type de sonification permet de suppléer ou même de remplacer la vue lorsque celle-ci se révèle insuffisante pour analyser certains types de données.

On peut par exemple citer les travaux de [Hayward, 1994], qui s'est attaqué à la représentation sonore de données sismiques¹⁵. Pour cela, il a utilisé des vibrations sismiques préalablement enregistrées comme des échantillons sonores, en partant du principe que la propagation d'une onde sismique est relativement similaire à celle d'une onde sonore dans l'air. Comme les fréquences de ces vibrations sont majoritairement subsoniques, il a accéléré la lecture des échantillons¹⁶. Grâce à cette compression temporelle, les composantes des ondes où se trouvait le plus d'énergie se sont transposées dans le spectre audible. Il a ensuite été démontré avec des tests que cette présentation aurale des données sismiques permettait à des scientifiques et même des étudiants d'identifier différents types d'activités sismiques, et ce plus facilement qu'avec une simple représentation visuelle de sismogramme. Les sujets étaient même capable de différencier un tremblement de terre d'une explosion. Cette méthode de sonification est couramment appelée « audification » ; on en détaille le principe dans la sous-section 1.3.1.

C'est ce type de fonctionnalité qui nous intéresse particulièrement pour la conception de notre propre sonification.

À noter que certaines artistes explorent aussi l'utilisation de données physiques ou abstraites normalement imperceptibles ; non pas pour les analyser scientifiquement *a posteriori* mais pour alimenter et/ou concevoir leurs œuvres. On pense surtout à des compositeurs de musique contemporaine et expérimentale, les plus connus dans ce domaine étant sûrement John Cage (1912-1992) et Iannis Xenakis (1922-2001). Attention, la plupart de ces artistes n'ont pas le concept de sonification en tête lorsqu'ils manipulent ces données, mais l'utilisation qu'ils en font rentre dans le cadre de la discipline. Leurs œuvres sont donc sources d'intérêt pour de nombreux musicologues, mais aussi des scientifiques qui retrouvent dans la composition musicale certaines possibilités

¹⁵ On peut retrouver dans la vidéo YouTube suivante un exemple concret de sonification d'un séisme : <https://www.youtube.com/watch?v=3PJxUPvz9Oo>.

¹⁶ Le séisme de la note précédente passe ainsi d'une durée d'activité de 2 jours à un signal sonore de seulement 2 minutes (vitesse de lecture multipliée par 1440).

d'application des modèles de sonification. Par exemple, Xenakis utilise dans sa pièce *Achorripsis*¹⁷ (1956-1957) différentes lois de probabilité pour générer des « événements musicaux » dont certains paramètres (comme le rythme, la hauteur, le timbre ou encore la vitesse de *glissando*) peuvent varier [Childs, 2002].

1.2.4. Autres fonctionnalités

La sonification permet également de développer d'autres fonctionnalités comme, entre autres, l'interaction et l'immersion sonore. On retrouve souvent ce type de problématiques dans les projets de sonification qui portent sur la notion d'exercices (ré)éducatifs, comme pour le sport et les jeux de divertissement ou d'apprentissage par exemple.

Ainsi, si les interfaces sonores ont d'abord été utilisées pour générer des sons environnementaux et accompagner l'image dans les jeux vidéos [Röber et al., 2006 ; Verron et al., 2010], elles ont ensuite permis de créer des jeux uniquement sonores (*audio games* en anglais) [Friberg et Gärdenfors, 2004 ; Gaudy et al., 2006]. La sonification permet donc de faciliter l'accès des jeux aux non-voyants, tout en favorisant l'interaction directe entre des joueurs voyants et non-voyants [Stockman et al., 2007].

Au niveau des activités sportives, les études de [Schaffert et al., 2009] sur l'aviron ou de [Godbout et Boyd, 2010] sur le patinage de vitesse ont montré que la sonification permet aux sportifs d'améliorer potentiellement leurs mouvements. Cette sonification du mouvement peut aussi être utilisée pour informer du bon déroulement d'une tâche, comme par exemple la justesse du geste d'écriture chez des enfants atteints de troubles dysgraphiques [Thoret et al., 2012]. Jean-Luc Velay et Jérémy Dannay (du Laboratoire de neurosciences cognitives du CNRS) présentaient ainsi au public lors du Forum du CNRS 2017 un outil permettant de sonifier de deux manières différentes l'écriture manuscrite à l'aide d'une tablette graphique et d'un morceau de musique que choisit l'utilisateur¹⁸. Dans le cas de la première méthode de sonification, la vitesse de lecture de la musique est asservie à la vitesse de déplacement du stylo. Ainsi, si le geste est trop lent, le morceau

¹⁷ Une version de la pièce jouée par l'Orchestre Philharmonique du Luxembourg dirigé par Arturo Tamayo est disponible sur YouTube à cette adresse : <https://www.youtube.com/watch?v=RHQM2-MMXzY>.

¹⁸ On peut voir une brève démonstration du prototype dans la vidéo YouTube suivante : https://www.youtube.com/watch?v=S3K9_jgVLis.

va être lu plus lentement (sans que sa hauteur ne soit pour autant transposée vers le grave¹⁹) ; et inversement si le geste est trop rapide. La deuxième méthode de sonification utilise des sons de synthèse qui évoquent des frottement quand la vitesse d'écriture est correcte, des grincements quand la vitesse est en deçà d'un seuil établi et des craquements quand le mouvement n'est pas fluide.

Enfin, au-delà du cadre des loisirs, on peut retrouver des propositions de sonification dans de nombreuses œuvres sonores performatives et interactives. Par exemple, dans *Music for Solo Performer*²⁰ (1965), Alvin Lucier (né en 1931) met en musique l'activité cérébrale de son cerveau (voir la figure 1.2). Grâce à une série de transductions et d'amplificateurs électriques, il anime en temps réel des percussions sur lesquelles sont disposés des haut-parleurs grâce aux ondes électromagnétiques produites par son cortex cérébral [Straebel et Thoben, 2014].



Fig. 1.2 : Photo de John Cage en train de réajuster les électrodes sur la tête d'Alvin Lucier avant une représentation de *Music for Solo Performer* à l'université Wesleyenne en 1988

¹⁹ On parle alors de *time stretching* (littéralement « étirement temporel » en français), un effet qui consiste à modifier la vitesse de lecture et donc la durée d'un son sans en affecter pour autant la hauteur.

²⁰ Une captation vidéo de la performance est disponible en ligne sur YouTube : <https://www.youtube.com/watch?v=bIPU2ynqy2Y>.

3. Différentes techniques et approches en sonification

Différentes techniques de sonifications peuvent être utilisées en fonction du type d'information à afficher (i.e. message de notification ou variation d'une donnée continue). Nous allons décrire dans cette sous-partie les principales approches utilisées en sonification. On peut les classer selon deux grandes catégories :

- Les systèmes « haut niveau » convertissent une information en un son signifiant symboliquement. Comme ils font appel à l'expérience préalable des utilisateurs, ces derniers les appréhendent aisément et rapidement.
- Les systèmes « bas niveau » génèrent du son directement à partir d'une donnée. Les sujets ont donc besoin d'une période d'apprentissage pour comprendre et intégrer la relation entre les sons produits et la propriété qui a été sonifiée.

1.3.1. Audification

L'audification (du mot anglais éponyme) est la méthode de sonification de la plus directe. Selon [Dombois et Eckel, 2011], elle consiste à présenter des données physiques évoluant au cours du temps en les transposant dans le domaine audible. Ces données, souvent issues de capteurs ou de simulations, rassemblent essentiellement des mouvements et vibrations inaudibles qui appartiennent pour la majorité au domaine physique des ondes électromagnétiques ou des ondes mécaniques. Cette approche nécessite souvent de transformer le signal à sonifier avec des dilatations temporelles ou fréquentielles afin de le transposer dans le domaine des fréquences audibles.

L'audification d'un jeu de données permet notamment de faciliter leurs classifications, d'y détecter d'éventuelles anomalies et d'explorer l'effet d'un changement de condition sur l'ensemble du jeu de données. Utilisée dans divers domaines tels que la médecine (pour les électro-encéphalogrammes ou l'étude des battements du cœur par exemple), la sismologie²¹ ou encore les statistiques, c'est cette technique qui est par exemple à l'œuvre dans le principe du stéthoscope (cf. 1.1.3), avec lequel on rend directement audible un mouvement matériel grâce à une amplification immédiate.

²¹ Voir la note de bas de page n°14, p. 18.

1.3.2. Sonification par *mapping*²² de paramètres

En anglais *Parameter Mapping Sonification* [Grond et Berger, 2011], la « sonification par mise en correspondance de paramètres » est la plus commune des méthodes de sonification utilisée pour l'analyse de données scientifiques. Elle consiste à associer une ou plusieurs dimensions des données à la variation d'un ou plusieurs paramètres sonores. Le compteur Geiger illustre parfaitement cette méthode : il rend perceptibles des quantités physiques (le niveau de radiation) en les mettant en correspondance avec un paramètre sonore (le nombre de pulsations sonores par minute). C'est un *mapping* droit et proportionnel : plus le niveau de radiation est élevé, plus la fréquence de répétition du son augmente.

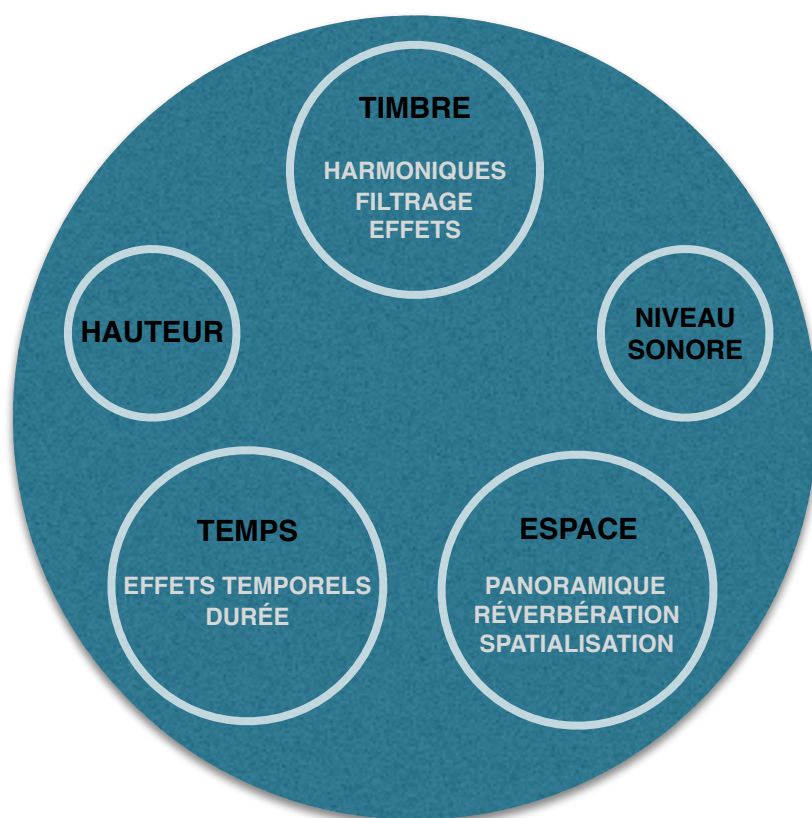


Fig. 1.3 : Les cinq catégories de paramètres sonores selon [Dubus et Bresin, 2013]

Les paramètres sonores habituellement utilisés sont d'une grande diversité. Dans leur revue des différentes stratégies de *mapping* pour la sonification de données physiques, Dubus et Bresin

²² On peut définir le *mapping* comme la mise en cohérence de minimum deux types d'informations distinctes (pour plus de précision, voir la sous-partie 3.2.1). Si le terme est parfois traduit par « mise en correspondance » en français, on lui préférera la terminologie anglaise qui est majoritaire dans la littérature francophone car plus facile à manipuler et plus courte. De plus, on se réserve également le droit d'utiliser le néologisme « mapper » comme verbe désignant cette action de mise en correspondance.

[2013] ont analysé par moins de soixante projets différents. Ils ont ainsi dénombré l'utilisation de trente paramètres sonores distincts. Ces paramètres ont ensuite été classés dans cinq catégories différentes (voir la figure 1.3), selon qu'ils soient relatifs à la hauteur, le timbre (harmoniques, filtrage, effets comme la distorsion, etc.), le niveau sonore, l'espace (panoramique, spatialisation et réverbération) ou le temps (durée, effets temporels comme le *delay*, etc.). Beaucoup de ces paramètres sont évidemment inter-classes, puisque qu'ils mettent justement en relation la variation d'une classe par rapport à une autre. Par exemple, l'enveloppe ADSR (*Attack Decay Sustain Release*) d'un son représente l'évolution de son niveau en fonction du temps. Autre exemple : si la réverbération est évidemment liée à la sensation d'espace, elle touche aussi à la temporalité puisqu'elle altère la durée du son réverbéré.

Grâce à ce grand nombre de paramètres sonores utilisables, la sonification par *mapping* de paramètres permet de nombreuses représentations différentes d'un même phénomène. Le choix de ces paramètres et de leur *mapping* doit cependant prendre en considération un certain nombre de contraintes, comme le nombre de données à sonifier, leur échelle de variation ainsi que les tâches à effectuer avec ces données (surveillance, interprétation, analyse, etc.). Il est également nécessaire de tenir compte des différentes limites psychoacoustiques du système auditif, telles que le seuil de discrimination (*Just-Noticeable Difference* ou JND), l'effet de masquage ou encore les différents seuils de perception de l'oreille humaine en terme d'amplitude, de fréquence et de variations temporelles. Walker et Kramer [1996] ont également montré que le choix des paramètres sonores, aussi bien que la polarité et l'étendue de leur variations, pouvaient avoir une grande influence sur les performances finales. Ils préconisent notamment de se méfier des *mappings* qui peuvent paraître intuitif. Il convient donc de tester plusieurs *mappings* différents avant de déterminer celui qui représentera au mieux les données de départ.

À noter que la sonification par *mapping* de paramètres est généralement appliquée à des sons MIDI²³ ou des sons de synthèse, principalement à cause de la facilité d'accès de leurs différents paramètres. Notre proposition de sonification est elle-même basée en partie sur des sons de synthèse (cf. 3.2).

²³ Le MIDI (*Musical Instrument Digital Interface*) est un protocole standard de communication, ainsi qu'un format de fichier dédié à la musique et utilisé notamment pour la communication entre instruments électroniques, contrôleurs, séquenceurs et logiciels de musique assistée par ordinateur (MAO).

1.3.3. Sonification par modélisation

L'approche de la « sonification par modélisation » (*Model-Based Sonification*) [Hermann, 2011] est originale. Plutôt que de relier les paramètres des données aux paramètres du son, Hermann et Ritter [1999] proposent de construire un modèle virtuel dont la réponse sonore à une entrée de l'utilisateur dépendrait des données à sonifier. Ainsi, l'idée est de : (1) construire un scénario virtuel à partir des données à sonifier, (2) définir un « modèle physique virtuel » qui régirait la réaction vibratoire des données à des excitations externes, (3) permettre à l'utilisateur d'exciter et d'écouter le système de manière interactive.

Cette technique de sonification utilise généralement la synthèse sonore par modèle physique²⁴ [Smith, 1992], ce qui la rend souvent plus agréable à l'écoute que la sonification par *mapping* de paramètres.

Une parfaite illustration de l'usage de cette technique peut être trouvée dans les travaux de [Hermann et Ritter, 1999]. Ces derniers proposent deux stratégies différentes de sonification par modélisation. Nous nous intéressons à la première, qui permet de sonifier des nuages de points. Hermann et Ritter ont imaginé que chaque point du nuage se comporte comme une « planète » dont la position est fixe dans un espace tridimensionnel. Chaque planète va alors contribuer à un potentiel global, qui se traduit par une force gravitationnelle. Le tout constitue donc le modèle physique virtuel.

Pour déclencher la sonification, il suffit d'injecter aléatoirement de nouvelles particules au sein de l'espace. Ces « satellites » vont alors subir la force gravitationnelle du modèle physique, qui va leur imprimer à chacun un mouvement différent que l'on peut traduire comme une oscillation harmonique et, par extension, une sinusoïde dont la fréquence dépendra des caractéristiques du système. Le signal audio résultant de la sonification correspond alors à la somme de l'ensemble des sinusoïdes produites par les mouvements de chacune des sondes spatiales introduites dans le modèle.

²⁴ La synthèse sonore par modélisation physique consiste à produire des sons à partir d'un modèle informatique qui décrit les propriétés physiques d'objets virtuels. Une simulation numérique est ensuite utilisée pour calculer les mouvements de ces objets lorsqu'ils sont excités. Le signal sonore final est alors synthétisé à partir des oscillations harmoniques résultant de ces mouvements.

1.3.4. Sonification par signaux auditifs

On appelle « signaux auditifs » les événements sonores, généralement de courte durée, qui sont utilisés dans les IHM pour donner une information sur un état, une action réalisée, le contenu d'un dossier ou encore le résultat d'une action de l'utilisateur sur le système. Il en existe une multitude de catégories différentes. Dans cette sous-partie, nous allons nous concentrer sur les deux types de signaux auditifs les plus couramment étudiés et utilisés dans le domaine des interfaces sonores : les *auditory icons* et les *earcons*.

On peut cependant noter qu'il existe d'autres types de signaux auditifs, comme les *spearcons* [Walker et al., 2006] (utilisés principalement dans les applications pour personnes aveugles ou malvoyantes) ou encore les *morphocons* [Parseihian et Katz, 2012].

1. Auditory Icons

Gaver [1986] définit les icônes sonores (*auditory icons* dans le texte) comme « des sons du quotidien reliés à des événements informatiques par analogie avec les événements produisant ces sons ». Ce sont généralement des sons brefs non verbaux, qui peuvent être considérés comme l'équivalent sonore des icônes visuelles utilisées dans l'interface graphique de l'ordinateur (voir la figure 1.4).



Fig. 1.4 : Évolution de l'icône macOS de la corbeille

L'idée simple de ces icônes sonores est d'utiliser l'expérience quotidienne d'écoute de l'utilisateur pour représenter différents objets, fonctions ou actions au sein d'une interface. Pour cela, le son doit être relié sémantiquement à la donnée qu'il est censé représenter, avec un lien naturel et intuitif. Par exemple, l'action de vider la corbeille d'un ordinateur va être représentée par le son d'un papier que l'on froisse.

L'avantage de ces icônes sonores réside dans leur très faible temps d'apprentissage. Leur principal inconvénient reste la difficulté à trouver une représentation sonore iconique, ayant une

relation directe et claire avec l'information à sonifier. Ils sont de fait beaucoup moins utiles lorsque l'on doit représenter des concepts plus abstraits.

2. *Earcons*

Les *earcons* représentent des sons abstraits, souvent musicaux ou synthétiques, qui peuvent être combinés pour créer une ou plusieurs « grammaires » sonores [Blattner et al., 1989 ; McGookin et Brewster, 2011]. Principalement utilisés pour l'interfaçage sonore des ordinateurs et appareils mobiles, ce sont des messages symboliques courts, non verbaux et construits autour de blocs simples qu'on appelle « motifs ».

Contrairement aux *auditory icons*, les *earcons* ont l'avantage de pouvoir représenter tout type d'information ou de concept, car aucune relation sémantique n'existe entre le son et l'information représentée. Leur utilisation nécessite en contrepartie un apprentissage parfois non négligeable de la part de l'utilisateur.

Les exemples d'*earcons* sont nombreux. Nous avons déjà mentionné précédemment les sons de démarrage d'ordinateur (cf. 1.2.2), mais il existe également leur équivalent pour la mise hors tension de l'ordinateur : celui créé par Brian Eno pour le système Windows XP de Microsoft reste comme l'un des plus connus²⁵. On peut également citer les signaux d'alarme utilisés en voiture pour signifier que notre ceinture n'est pas attachée. Mais c'est avec nos téléphones portables que nous expérimentons une utilisation quotidienne des *earcons* : tous les sons liés à la réception ou l'envoi de messages relèvent en effet de cette catégorie.

4. Conclusion

Nous avons donc présenté dans cette première partie les fondements théoriques et pratiques de la sonification de données. En passant en revue ses différents domaines d'application ainsi que quelques techniques et approches parmi les plus utilisées, nous avons pu dresser un panorama exhaustif de la discipline. Cela nous permettra d'appréhender plus facilement les enjeux et contraintes des différents projets de sonification que l'on va s'attacher à analyser dans la partie suivante.

²⁵ Voir la vidéo YouTube suivante : <https://www.youtube.com/watch?v=Gb2jGy76v0Y>.

2. Prérequis à la sonification de protéines

Cette deuxième partie dresse un état de l'art raisonné de la sonification de molécules et d'objets tridimensionnels²⁶. À travers ce panorama, on dévoile notamment les différentes approches méthodologiques ayant cours dans chacun de ces sous-domaines, ainsi que le raisonnement qui a mené aux différents choix de conception de notre propre méthode de sonification.

2.1. Rappel sur les protéines

2.1.1. La structure des protéines

Les protéines forment un constituant essentiel de la matière organique et des êtres vivants. Ce sont des macromolécules, composées par une ou plusieurs chaîne(s) d'acides aminés²⁷. Ces différentes chaînes sont aussi parfois appelées « résidus ». Elles possèdent des propriétés physico-chimiques particulières, par exemple de charge (positive, négative ou neutre à pH physiologique) ou de capacité à faire des liaisons avec l'eau (hydrophilie) ou pas (hydrophobicité). En général, on parle de « protéine » lorsque la chaîne contient plus de 100 acides aminés (en deça, on parle de « polypeptides », voir de « peptides » pour moins de 10 acides aminés). La taille des protéines est de l'ordre de quelques dizaines à plusieurs centaines d'ångströms²⁸. Les plus grosses peuvent atteindre jusqu'à plusieurs dizaines de milliers d'atomes.

Une protéine possède trois types de structures (voir la figure 2.1) :

- La structure primaire (aussi appelée « séquence ») décrit l'enchaînement linéaire des acides aminés qui composent sa chaîne principale²⁹.

²⁶ On se réserve le droit d'utiliser par la suite le sigle « 3D » en remplacement afin de faciliter la lecture.

²⁷ Un acide aminé est une petite molécule organique qui possède un squelette carboné, ainsi que deux fonctions : une amine et un acide carboxylique. Il n'en existe que vingt-deux types différents dans les organismes vivants. Tous sont décrits selon une lettre différente.

²⁸ Pour rappel, $1 \text{ \AA} = 10^{-10} \text{ m} = 0,1 \text{ nm}$.

²⁹ Concrètement, la structure primaire d'une protéine est représentée par une succession de lettres correspondant chacune à l'abréviation de l'un des vingt-et-deux acides aminés standards.

- La structure secondaire se rapporte elle à la description du repliement local de cette chaîne principale en sous-structures stables comme les hélices α et les brins β .
- Enfin, la structure tertiaire décrit la position relative des différents éléments de la structure secondaire dans l'espace les uns par rapport aux autres³⁰.

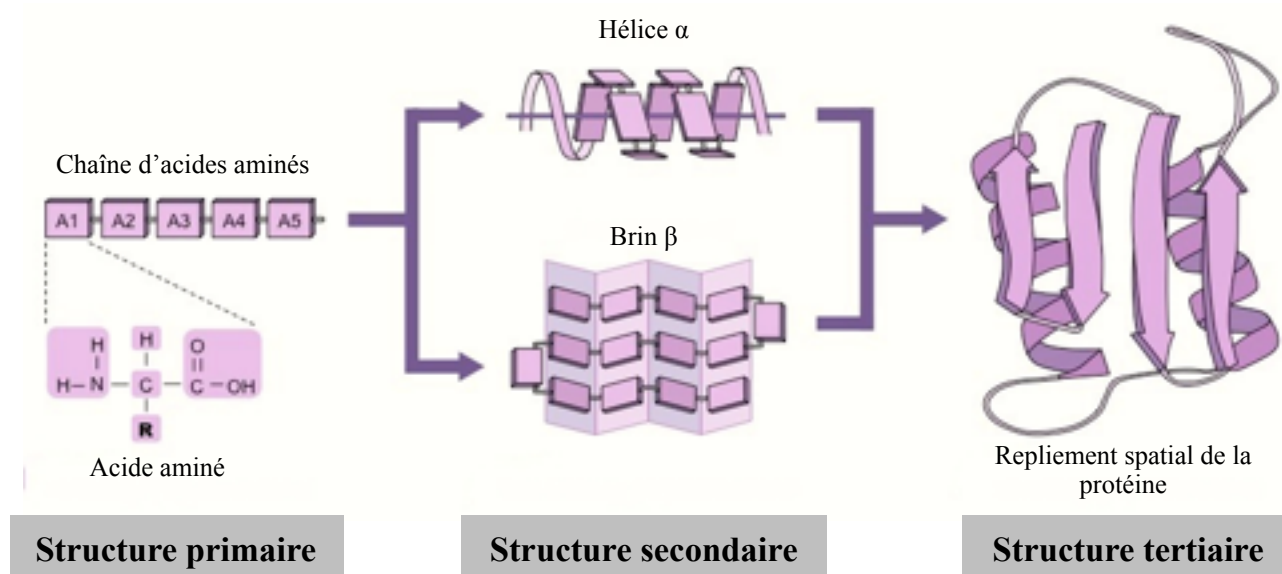


Fig. 2.1 : Les trois types de structures d'une protéine (illustration tirée du site pédagogique BioNinja)

Chaque protéine possède un agencement tridimensionnel qui lui est propre ; on l'appelle plus couramment « structure » ou « repliement spatial ». Comme le rôle de la plupart des protéines est subordonnée à leur interaction en surface avec un ou plusieurs autres partenaires (d'autres protéines, de l'ADN³¹ ou de l'ARN³²), leurs propriétés et leur fonction dépendent concrètement de leur forme. C'est cette conformation spatiale qui va par exemple dicter la spécificité d'une éventuelle activité catalytique en déterminant la position du site actif des enzymes.

Les complexes protéine-protéine formés par ces interactions sont présents dans beaucoup de processus cellulaires, physiologiques et biochimiques. L'étude de ces derniers passe donc nécessairement par la détermination des fonctions des protéines dans leur environnement naturel.

³⁰ Il existe également une structure quaternaire, qui correspond au même agencement que la structure tertiaire mais pour des protéines ayant plusieurs chaînes d'acides aminés.

³¹ L'acide désoxyribonucléique (ADN) constitue le support de l'information génétique.

³² L'acide ribonucléique (ARN) est notamment impliqué dans la synthèse des protéines, ainsi que le transport du matériel génétique.

C'est pourquoi les chercheurs en biologie structurale ont déterminé durant les dernières décennies une énorme quantité de structures spatiales macromoléculaires.

Il existe différentes techniques expérimentales pour déterminer la structure d'une protéine. La plus utilisée est la cristallographie aux rayons X. Cette méthode s'appuie sur le phénomène physique de la diffraction : elle consiste à envoyer un faisceau d'ondes électromagnétiques (comme des rayons X) sur le cristal³³ de la protéine dont on souhaite obtenir la structure. Pour chaque position de la protéine dans le cristal, le faisceau est dévié comme sur miroir sur les différents plans atomiques³⁴ et donne ainsi naissance à des taches de diffraction caractéristiques visibles sur un détecteur. L'étude de la position et de l'intensité de ces taches, couplée à des méthodes de calcul informatisées, permet alors de déterminer la position exacte des atomes et donc de reconstituer l'ensemble de la structure de la protéine. On peut noter que la spectroscopie RMN³⁵ (Résonance Magnétique Nucléaire) ou la microscopie électronique sont également parfois exploitées.

2.1.2. Représentations et visualisation de la structure des protéines

1. Les banques de données moléculaires en ligne

Les structures 3D ainsi obtenues sont intégrées et stockées dans des banques de données moléculaires en ligne. Ces banques proposent ensuite aux utilisateurs de consulter ou télécharger gratuitement les structures, et ce dans différents formats de fichiers.

La banque la plus connue et la plus utilisée est la RCSB PDB (*Research Collaboratory for Structural Bioinformatics Protein Data Bank*, souvent abrégée en PDB). C'est une collection mondiale de macromolécules biologiques constituée essentiellement de protéines d'intérêt pharmaceutique, mais aussi d'acides nucléiques. Les données expérimentales qui sont déposées dans la PDB par des biologistes, chimistes et bio-informaticiens du monde entier appartiennent au domaine public. Leur consultation est gratuite et peut se faire directement depuis le site web de la

³³ Un cristal est un matériau à l'état solide dont les atomes sont disposés selon un schéma ordonné tridimensionnel.

³⁴ En théorie, les faces d'un cristal sont des surfaces planes.

³⁵ Elle se base pour sa part sur les propriétés magnétiques de certains noyaux atomiques pour mesurer leur proximité avec d'autres atomes, exactement sur le même principe que la technique d'imagerie médicale IRM (Imagerie par résonance magnétique).

banque³⁶, qui contient toutes les ressources nécessaires à l’affichage et la manipulation des structures 3D qu’il héberge.

2. Le nombre de structure moléculaires disponibles

Le nombre de structures résolues est en constante évolution. Lors de sa création en 1971, la PDB contenait seulement sept structures. Le nombre de structures déposées a vraiment commencé à augmenter pendant les années 80, parallèlement aux perfectionnements des techniques cristallographiques. Au début des années 1990, la majorité des revues scientifiques de chimie et biologie exigeaient déjà des articles soumis qu’ils indiquent le code PDB³⁷ de toutes les données structurales qu’ils mentionnaient. La croissance du nombre d’entrées de la PDB est alors devenu exponentielle, jusqu’à atteindre entre 6 000 et 7 000 dépôts de structures chaque année. Ainsi, il y avait pratiquement 50 000 structures disponibles en ligne début 2008, puis 100 000 en 2013. On en compte aujourd’hui plus de 130 000.

3. Le format PDB

Le format PDB (extension .pdb) est le format original et principal de la *Protein Data Bank*. Il décrit la localisation exacte de chaque atome de la protéine étudiée, c’est-à-dire les coordonnées cartésiennes de l’atome dans un repère à trois dimensions (x, y, z). À noter qu’il contient également des métadonnées³⁸ sur la structure primaire de la molécule, ses éventuelles structures secondaires, ainsi que des informations sur la méthode expérimentale de résolution de structure utilisée.

4. Les logiciels de visualisation

Les fichiers PDB peuvent être lus par différents logiciels de « visualisation³⁹ » *open-source*⁴⁰ ou commerciaux. Ils permettent de manipuler de différentes manières les structures

³⁶ <https://www.rcsb.org/>

³⁷ Chaque modèle de structure est référencé dans la PDB par un identifiant unique à quatre caractères qu’on appelle « code PDB ».

³⁸ On appelle couramment « métadonnées » un ensemble structuré de données particulières décrivant une ressource ou une donnée générale quelconque.

³⁹ Ce sont en fait plutôt des logiciels de simulation, puisqu’ils n’affichent pas la véritable structure moléculaire mais la recalculent à partir des données encapsulées dans les fichiers.

⁴⁰ On dit d’un logiciel qu’il est « *open-source* » quand son code source est distribué sous une licence permettant à n’importe quel utilisateur de lire, modifier ou redistribuer ce logiciel.

décrites dans les fichiers PDB, en accentuant par exemple certaines caractéristiques ou en choisissant le mode de représentation le plus approprié. Leur interface graphique est également interactive : on peut sélectionner les molécules avec le pointeur de la souris pour les orienter dans n'importe quelle direction et modifier le facteur d'agrandissement afin de zoomer dans leur structure. Enfin, ils proposent un ensemble de fonctions pouvant servir à animer ou simuler l'évolution de ces modèles virtuels. Les trois logiciels les plus connus et démocratisés sont *PyMOL*, *VMD* et *Chimera* - on verra d'ailleurs dans la sous-section 2.2.2. que plusieurs sonifications de protéines sont basées sur des *plugins*⁴¹ développés pour ces mêmes logiciels.

5. Les différentes formes de représentations visuelles

L'exploration visuelle des structures protéiques n'a pas toujours été simple : les premiers modèles structuraux ont été construits comme des maquettes matérielles, à l'aide de tiges reliant des boules représentant les différents atomes (voir la figure 2.2).

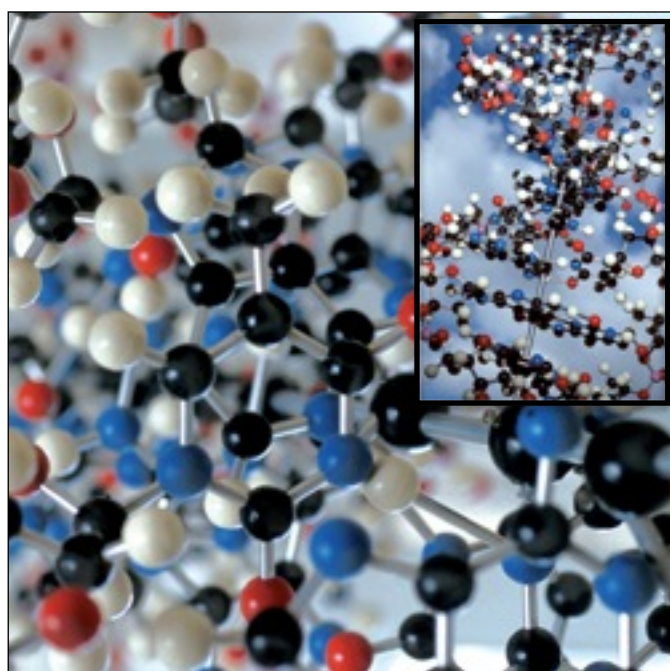


Fig. 2.2 : Maquette réelle d'une molécule d'ADN
(photos de Michel Depardieu pour l'Inserm)

Il faut attendre le milieu des années 1960 et les premiers écrans graphiques - proches de ceux d'un oscilloscope actuel - pour voir la première visualisation d'une molécule sur un écran.

⁴¹ Un *plugin* (aussi parfois orthographié « *plug-in* ») est un paquet de fichiers informatiques qui sert à compléter un logiciel hôte en lui apportant de nouvelles fonctionnalités. Il porte de nombreux autres noms, comme « module d'extension », « *plugiciel* », « *add-on* », etc.

L'arrivée des écrans graphiques à balayage de trame - similaires aux écrans actuels - dans les années 1980, ainsi que la diminution du coût de la puissance de calcul des ordinateurs, a permis un grand pas qualitatif dans le réalisme de la visualisation : il était enfin possible de dessiner des surfaces, de gérer la transparence des objets et de modéliser leur éclairage et les ombres résultantes. Différents modes de représentation de la structure tridimensionnelle d'une molécule sont ainsi apparus [O'Donoghue et al., 2010]. On peut les classer en trois grandes catégories :

- représentation des atomes et de leurs liaisons (voir la figure 2.3) ;

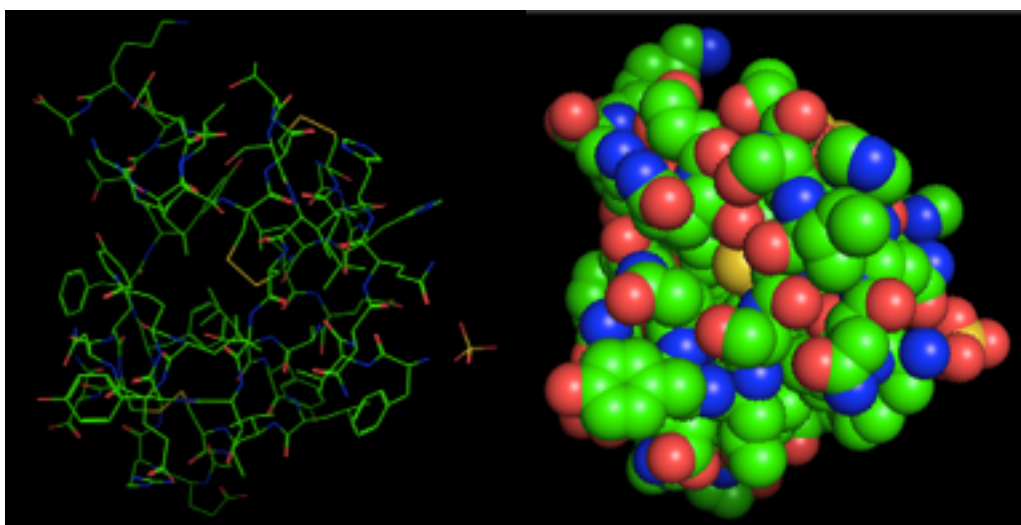


Fig. 2.3 : Représentations des atomes de la protéine 1B18 sous forme de branches (*licorice*) à gauche et de sphères à droite (obtenues avec *PyMOL*)

- représentation des acides aminés (voir la figure 2.4) ;

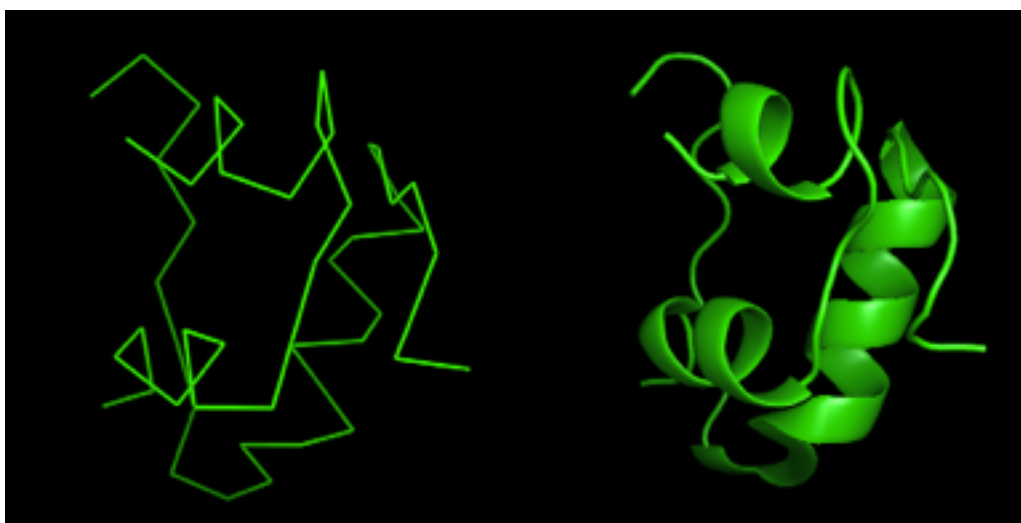


Fig. 2.4 : Représentations de la chaîne principale de la protéine 1B18 sous forme de colonne vertébrale à gauche et de ruban à droite (obtenues avec *PyMOL*)

- représentation de la surface de la protéine (voir la figure 2.5).

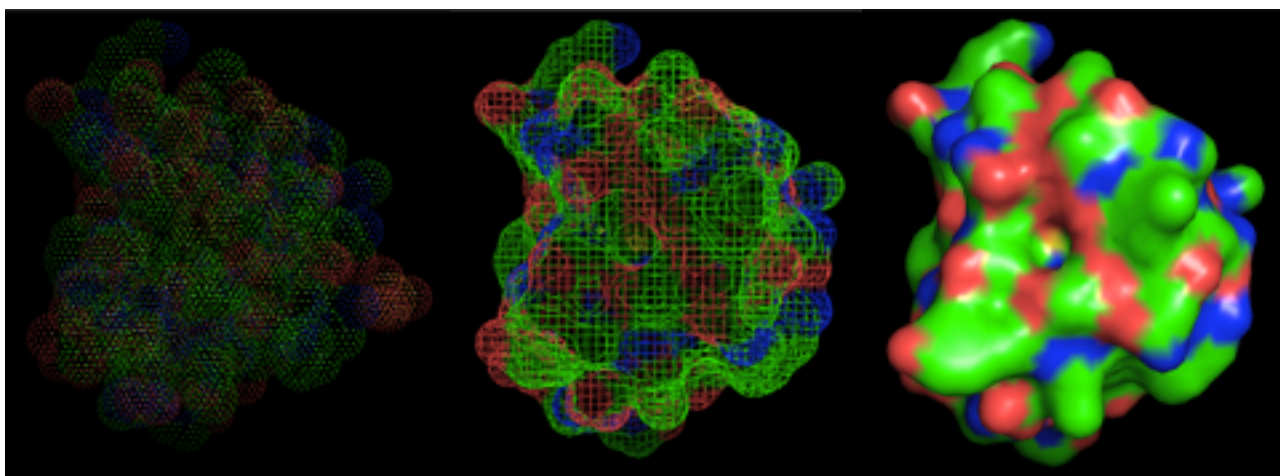


Fig 2.5 : Représentations de la surface de la protéine 1B18 sous forme de nuage de points à gauche, de fil de fer au centre et de solide lissé à droite (obtenues avec *PyMOL*)

2.1.3. Limites des représentations visuelles et nécessité de la sonification

L'analyse d'une représentation moléculaire est un tâche qui nécessite une inspection plus ou moins approfondie de sa structure - cette expérience singulière exige donc de la part de l'utilisateur une certaine expertise. Mais malgré toutes les méthodes d'affichage et de représentation existantes, l'interprétation des fonctions d'une macromolécule ou la détection d'écarts significatifs fonctionnels d'une protéine à une autre restent des procédures très complexes, car elles rencontrent de nombreux obstacles :

- le manque de détail atomique dans une vue d'ensemble de la structure protéique ;
- à l'inverse, la perte du contexte général de la structure lorsque l'on zoome sur la protéine ;
- le phénomène d'obstruction engendrée par la perspective centrée : une partie de la structure est continuellement cachée à l'arrière ;
- la limitation du nombre d'attributs pouvant être visualisés simultanément (l'information passe en effet essentiellement par la couleur et la représentation choisie) ;

- la difficulté à superposer ensemble plusieurs types de représentation (même en jouant sur la transparence, ces superpositions brouillent et complexifient souvent outre mesure la visualisation de la structure moléculaire).

En complément des modalités visuelle et haptique déjà utilisées (pour la représentation de la protéine sur l'écran et l'interaction à la souris avec son orientation et son facteur d'agrandissement), la modalité auditive permet donc de combler certaines lacunes et faiblesses de la représentation graphique - notamment en ce qui concerne les informations visuelles difficiles d'accès. Enfin, ajouter une nouvelle modalité permet également de renforcer par combinaison la perception multimodale.

En s'appuyant sur tous ces avantages, et dans l'espoir de dégager de nouvelles connaissances, de nombreuses méthodes de sonification de molécules ont été proposées. Nous nous proposons de les passer en revue dans la sous-section suivante.

2. La sonification de molécules

Sont ici détaillées différentes propositions de sonification faites à ce jour dans le domaine de la chimie moléculaire. Nous aborderons notamment plusieurs exemples de sonification de protéines.

2.2.1. Musique et importance du facteur séquentiel

Un des premiers projets utilisant la sonification pour de l'analyse de données chimiques consistait à faire correspondre des sons à des séquences d'ADN [Hofstadter, 1980]. Même si cette tentative pouvait aider à l'analyse scientifique des données sonifiées, le but était ici plutôt artistique.

Les premiers travaux traitant véritablement de la sonification de molécules à des fins scientifiques remontent au milieu des années 80. La plupart de la recherche en chimie moléculaire portait alors sur le séquençage de l'ADN. Hayashi et Munakata [1984] ont ainsi été parmi les premiers à publier sur la possibilité d'utiliser de la musique pour représenter puis analyser des molécules d'ADN. L'ADN est formé de deux brins enroulés l'un autour de l'autre en double hélice (voir la figure 2.6). Ces brins sont constitués d'un assemblage de nucléotides, eux-mêmes formés à partir d'une base nucléique. Il en existe quatre sortes différentes : adénine (A), cytosine (C), guanine (G) et thymine (T). C'est la succession de ces bases nucléiques sur l'ADN qui détermine la

suite d'acides aminés constituant les protéines issues de ces gènes. On appelle « code génétique » cette correspondance entre bases nucléiques et acides aminés.

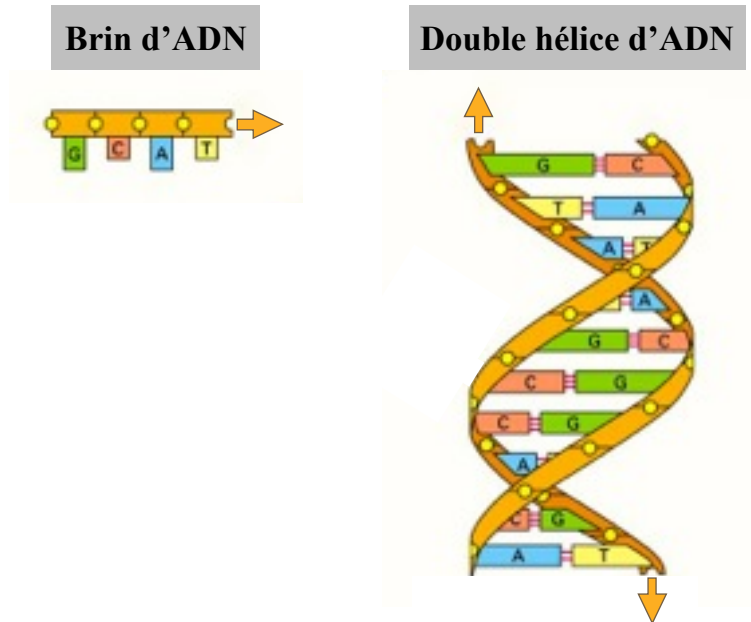


Fig. 2.6 : Illustration d'un brin d'ADN et d'une double hélice d'ADN
(illustration par Marjolaine Faure)

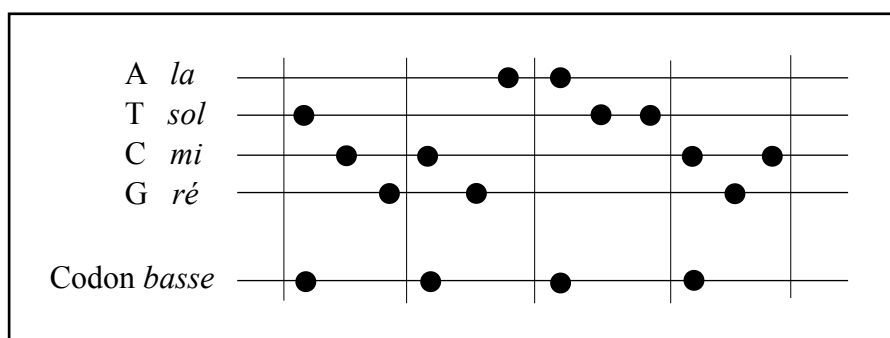
Leur proposition de sonification consiste donc à associer une note de musique différente à chacune de ces quatre bases nucléiques : *la* pour A, *sol* pour T, *mi* pour C et *ré* pour G. Puis, comme les couples AT et CG forment des paires entre les deux brins de la double hélice, il suffit de lire en musique la séquence de bases nucléiques du brin codant⁴² pour sonifier la portion d'ADN choisie. La séquence TCGCGAATTCGC peut ainsi être représentée de la manière suivante, comme une partition :

A	<i>la</i>	_____	•	•	_____		
T	<i>sol</i>	_____	•	_____	•	•	
C	<i>mi</i>	_____	•	•	_____	•	•
G	<i>ré</i>	_____	•	•	_____	•	

⁴² On appelle « brin codant » le brin principal chargé de coder l'ADN. Le second brin n'est qu'une simple transcription du premier par la règle des paires de bases nucléiques.

Si ce *mapping* de paramètres paraît sommaire, il demeure efficace. En sonifiant de cette manière différentes séquences d'ADN, Hayashi et Munakata ont en effet réussi à faciliter leur reconnaissance, tout en mettant en exergue les répétitions de motifs au sein d'une même séquence. De plus, leurs résultats furent concluants quant à la quantité de données pouvant être facilement analysée de cette manière par un utilisateur. Un impact sur la motivation des utilisateurs fut également décelé : leur sonification permettait de diminuer (voire d'éviter) l'ennui causé habituellement par une tâche répétitive mettant en jeu des données exclusivement représentées de manière visuelle.

Le projet *Protein Music* du chercheur Ross King et du musicien Colin Angus constitue une extension intéressante de ce travail [King et Angus, 1996]. Ils ont codé ensemble un programme informatique connecté en MIDI avec un synthétiseur. Comme précédemment, leur *mapping* met en correspondance les quatre bases nucléiques de l'ADN avec quatre notes MIDI différentes. Leur principal apport est l'ajout d'une « basse continue » qui sonifie la séquence protéique codée par l'ADN : à chaque codon⁴³ est assignée une note dont la hauteur dépend de certaines caractéristiques de l'acide-amino qu'il transcrit, comme sa polarité de charge ou sa taille. On a donc une note supplémentaire qui est jouée toutes les trois bases nucléiques :



Plusieurs autres biologistes de cette époque ont pointé les similitudes existant entre la musique et le séquençage biologique. Ainsi, Lubert Stryer détaillait dès 1988 la comparaison qu'il avait faite d'un réseau de molécules d'une fibre de collagène⁴⁴ avec une fugue [Stryer, 1988].

⁴³ Chaque triplet de nucléotides sur l'ADN correspond en effet à un codon chargé de transcrire un acide aminé en particulier.

⁴⁴ Le collagène est la famille de protéines la plus abondante dans un organisme humain, le plus souvent sous forme fibrillaire.

Nombre de chercheurs et de musiciens se sont également associés pour imaginer et composer de nouvelles pièces à partir de ces analogies séquentielles entre musique et données moléculaires [Sansom, 2002]. Par exemple, le compositeur Susan Alexjander a imaginé avec le biologiste David Deamer un *mapping* où la hauteur de chaque son synthétisé associé à une base nucléique est proportionnelle à son niveau d'absorption de la lumière [Alexjander et Deamer, 1999]. On peut enfin citer le projet *Life Music* de l'artiste John Dunn et de la biologiste Mary Ann Clark, qui ont collaboré ensemble sur un ensemble d'algorithmes permettant de convertir la structure primaire de protéines en séquences musicales : à chacun des acides aminés standards était ainsi associée une hauteur de note différente, les amino-acides hydrophobes étant assignés à des notes plus graves que les hydrophiles [Dunn et Clark, 1999].

La fin du vingtième siècle a continué de voir éclore des propositions de sonification musicale liées à des données moléculaires séquentielles [Ohno et Ohno, 1986 ; Ohno, 1988 ; Pickover, 1992]. Tous les travaux présentés dans cette sous-section se sont d'abord attachés à rendre audible l'aspect séquentiel et linéaire qui peut caractériser des données moléculaires (une suite de base nucléiques pour l'ADN et un enchaînement d'acides aminés pour les protéines), mais sans forcément chercher à en dégager une information sur la protéine elle-même ou ses fonctions.

Nombres de ces tentatives fluctuaient donc entre l'analyse scientifique des données sonifiées et le geste artistique. En raisonnant par analogie, on pourrait donc très bien nommer cet écart « musification ». À travers l'audification musicale de séquences de protéines, Picinali [2012] a cependant émis la possibilité de séparer clairement deux sets de protéines en travaillant sur leur aspect séquentiel. Bywater et Middleton [2016] ont confirmé cette intuition en montrant notamment que la sonification « musicale » de séquences d'acide aminées pouvait aider à dégager des patterns de reconnaissance et de classification chez les protéines, par exemple en ce qui concerne le repliement de leur structure.

Si la sonification séquentielle de la structure primaire des protéines concerne une grosse partie des publications du domaine, d'autres travaux se sont penchés sur la sonification des paramètres relatifs à la structure secondaire et tertiaire.

2.2.2. Vers une sonification paramétrique

Il n'est plus question dans cette sous-partie de sonifier les séquences de la structure primaire d'une protéine ; on s'intéresse davantage à leur structure secondaire (qui décrit, on le rappelle, le repliement local de sa chaîne principale) et leur structure tertiaire (la structure tridimensionnelle de la protéine, qui nous intéresse dans notre cas). Ainsi, on passe d'une sonification séquentielle à une sonification de statut et d'états paramétriques. On remarque d'ailleurs que nombre de ces propositions de sonification passent par l'utilisation d'une interface de visualisation ainsi que l'implémentation d'un *plugin* ou d'un logiciel dédié. On se propose donc de détailler plusieurs sonifications différentes, portant sur trois types de paramètres différents.

1. Complémentarité de deux protéines

PROMUSE [Hansen et al., 1999] est un logiciel qui permet de sonifier la superposition de deux protéines dans un espace tridimensionnel⁴⁵ (voir la figure 2.7). Le but principal de ce projet était d'étudier dans quelle mesure l'usage de la sonification peut permettre d'améliorer notre perception et notre reconnaissance des spécificités locales d'une structure tridimensionnelle.

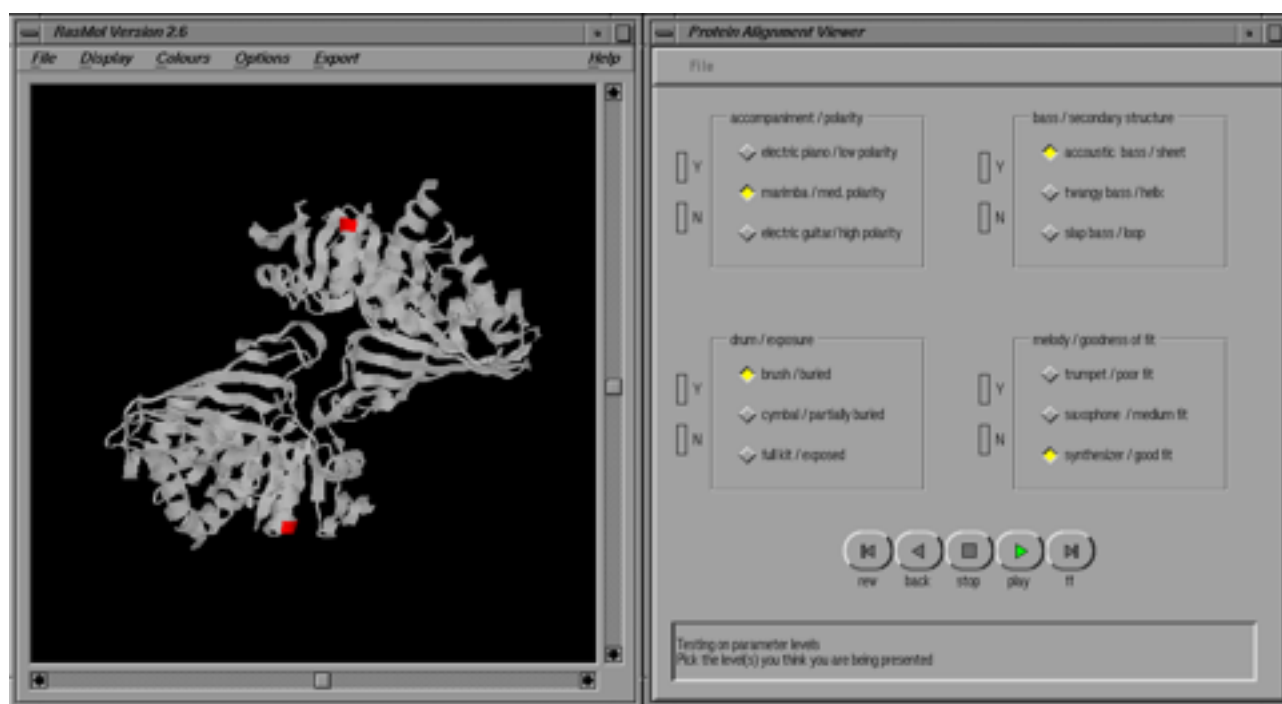


Fig 2.7 : Interface utilisateur de *PROMUSE* (à droite), couplée au logiciel de visualisation *RasMol* (à gauche)

⁴⁵ C'est ce qu'on appelle en chimie un alignement structurel.

Avec leur proposition, Hansen et ses collaborateurs désiraient donc préciser la représentation visuelle de la structure et potentiellement lever l'ambiguïté de certaines de ses caractéristiques.

Jugeant que les sonifications d'ADN ou de protéines déjà existantes sonnaient souvent de manière aléatoire et dissonante, Hansen et le reste de son équipe ont donc choisis pour leur sonification d'utiliser des sons musicaux, non pas distribués comme des notes éparses et indépendantes mais groupés en de courtes phrases mélodiques. De plus, ces phrases étaient jouées par un arrangement musical de type « quarter de jazz », composés de quatre flux sonores distincts : basse, batterie, accompagnement rythmique et instrument mélodique (*lead instrument* dans le texte). Chacun de ces quatre flux représente un type de données en particulier (voir la figure 2.8).

<i>parameter</i>	<i>data value</i>	<i>part</i>	<i>voice</i>	<i>volume^b</i>	<i>pan^c</i>
Secondary Structure	Sheet	Bass	acoustic bass	127	0
	Helix		twangy bass	127	0
	Loop		slap bass	127	0
Polarity	Low	Comp	electric piano	127	127
	Medium		marimba	70	127
	High		electric guitar	70	127
Exposure	Buried	Drums	brush	127	64
	Partially Buried		cymbals	127	64
	Exposed		full kit	127	64
Goodness of Fit	Poor	Melody	trumpet	80	33
	Medium		saxophone	70	33
	Good		synthesizer	70	33

Fig. 2.8 : Tableau de *mapping* de *PROMUSE* [Hansen et al., 1999]

C'est l'instrument mélodique qui a été choisi pour renseigner sur le niveau de complémentarité des deux protéines, car c'est l'instrument sur lequel on va naturellement focaliser notre attention. Les autres flux sonores servent à fournir des informations complémentaires, comme l'accompagnement qui donne par exemple une indication générale sur le degré de polarité du complexe protéique formé. Pour faciliter notre perception de la représentation sonore et affiner notre différenciation de ces différents flux, l'équipe de *PROMUSE* a également intégré des effets de panoramique et de niveau.

Le test qu'ils ont ensuite conduit consistait à juger une expérience d'alignement structurel avec des informations présentées de manière visuelle, puis auditive et enfin audiovisuelle. Ce sont les deux présentations comportant du son qui donnèrent les meilleurs résultats, la présence de l'information visuelle pour l'une d'elle n'améliorant que de très peu la performance. De plus, les chercheurs montrèrent que la sonification aidait à lever l'ambiguïté propre aux visualisations moléculaires très chargées. L'utilisation de la stéréo ainsi que le travail sur les timbres furent des moyens de discriminer les données de manière efficace. On peut trouver dans ce papier une première hypothèse sur l'utilisation du son 3D pour faciliter notre perception de différents flux sonores superposés.

2. Facteur B

SUMO (Sonification Utility For Molecules) est un *plugin* de sonification développé sur *SuperCollider*⁴⁶ [Gond et Dall'Antonia, 2008]. Il communique en OSC⁴⁷ avec *PyMOL* qui sert d'interface de visualisation (voir la figure 2.9).

Ce *plugin* propose à l'utilisateur deux méthodes de sonification de protéines différentes. La première s'attache à la description séquentielle des amino-acides tandis que la seconde, qui nous intéresse ici, propose une sonification du facteur B⁴⁸. Cette représentation sonore est particulièrement intéressante pour l'utilisateur car la méthode de représentation visuelle standard du facteur B rend son interprétation difficile : la valeur du facteur est affichée comme un nombre à côté de chaque atome. Il est certes aussi possible d'utiliser la couleur pour le représenter mais cela revient à bloquer un attribut visuel majeur pour représenter un paramètre mineur de la protéine (on préfère plutôt utiliser la couleur pour indiquer la charge des atomes par exemple).

L'utilisateur peut sélectionner dans *PyMOL* la zone de la molécule dont il souhaite connaître la précision de représentation. La sonification consiste alors en un *mapping* de paramètres : la

⁴⁶ *SuperCollider* est un environnement de programmation pour la synthèse audio en temps réel.

⁴⁷ L'*Open Sound Control* (OSC) est un protocole de transmission de données principalement utilisé pour communiquer entre un ordinateur et un synthétiseur.

⁴⁸ Le facteur B (aussi nommé « facteur de Debye-Waller » ou « facteur de température ») décrit le déplacement des atomes pendant la détermination expérimentale de la structure d'une molécule. Provoqués par l'agitation thermique, ces mouvements font que la position des atomes fixée dans la structure finale est plus ou moins faussée. Ainsi, le facteur B fournit une indication sur la précision de la représentation d'une protéine : plus il est élevé, plus sa structure contient des approximations de position, voire des erreurs. Il peut également donner une appréciation de la flexibilité de la protéine.

2. Prérequis à la sonification de protéines

valeur moyenne du facteur B de chaque résidu est traduite exponentiellement en un son pur d'une fréquence comprise entre 100 et 2000 Hz. Les sons résultants sont alors joués successivement de gauche à droite, avec les plus faibles valeurs de facteur B à gauche et les plus élevées à droite. Le résultat final ressemble donc à une sorte d'arpège. Gond et Dall'Antonia ont montré que l'on pouvait déterminer grâce à cette sonification la précision relative de la cristallisation d'une petite zone de la structure sans avoir à zoomer dans le détail pour lire la valeur numérique affichée à côté de chaque atome.

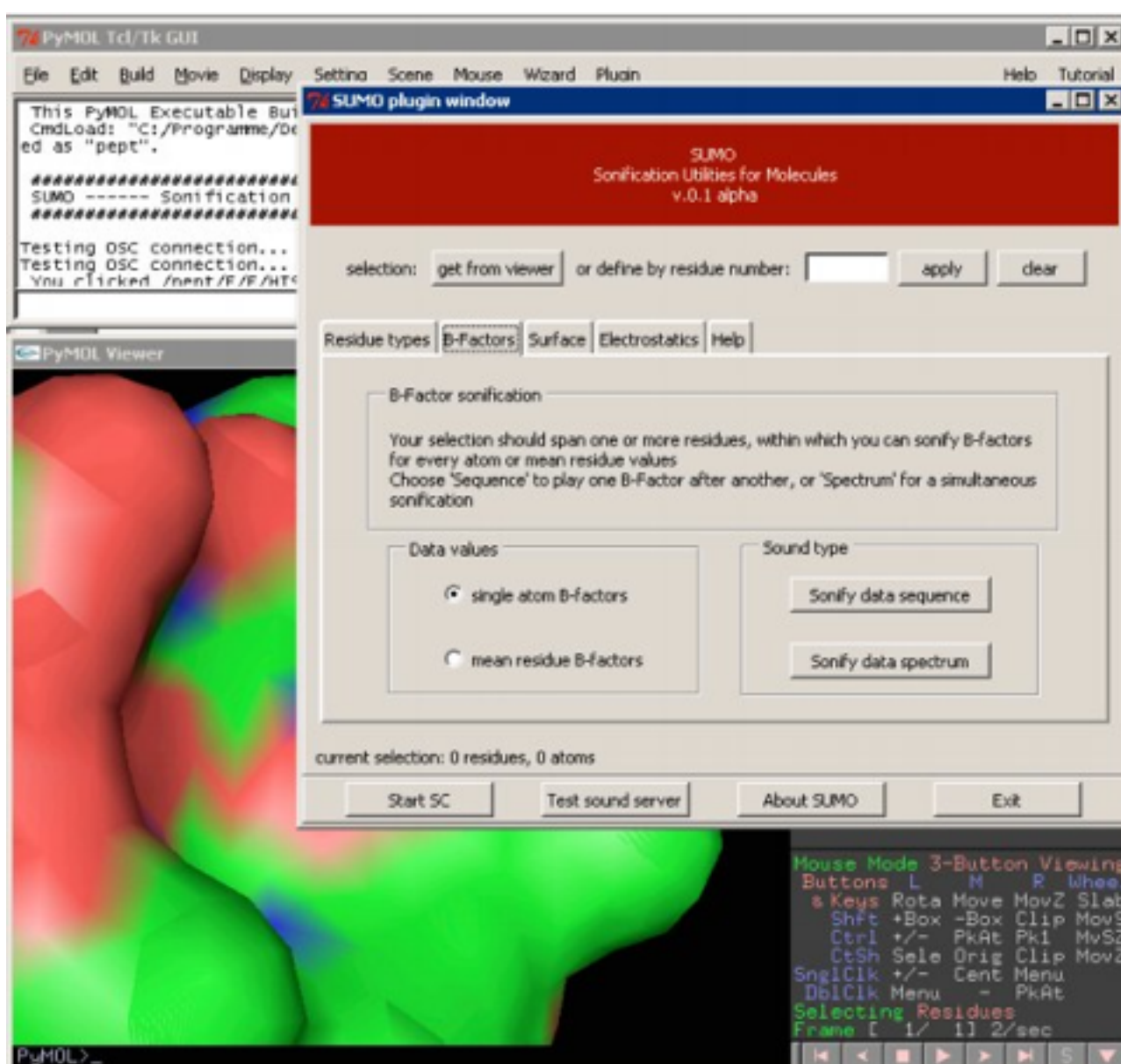


Fig. 2.9 : Interface utilisateur du *plugin SUMO*, avec *PyMOL* en arrière-plan [Gond et Dall'Antonia, 2008]

3. La sonification de données dynamiques

Tous les projets précédents travaillaient sur des jeux de données statiques. Mais différentes propositions de sonification de données dynamiques ont été faites. L'enjeu de ce type de sonification est de rendre audible la variation dans le temps ou l'occurrence d'une donnée, plus que la valeur de la donnée elle-même.

Par exemple, un système d'audification qui rend perceptible les vibrations moléculaires a été développé par [Delatour, 2000]. Une fois la fréquence d'oscillation de la molécule mesurée, il peut choisir de la convertir directement en une forme d'onde qui sera lue ou alors l'utiliser pour régler la hauteur et l'amplitude d'un sinus. À travers le résultat sonore obtenu, Delatour indique la possibilité d'utiliser scientifiquement un tel *mapping* pour identifier et classer différents composés chimiques.

On peut enfin également citer le projet de [Rau et al., 2015]. Il propose parmi d'autres une sonification des liaisons hydrogènes⁴⁹ entre une molécule et un solvant au sein d'une simulation dynamique. La formation et la rupture de ces liaisons hydrogènes sont des événements ponctuels et donc difficilement perceptibles par leur seule représentation visuelle. Pour les mettre en évidence, Rau et ses collaborateurs ont utilisé des icônes sonores spatialisées, combinées à une nappe sonore continue qui correspond au niveau d'activité générale pré-calculé du complexe moléculaire. Cette ambiance sonore en fond permet d'avoir une vision plus large de la simulation, ce qui est utile lorsque que l'on est en train de se concentrer sur une partie précise de la molécule et/ou que l'information n'est pas accessible *via* la représentation visuelle actuellement utilisée.

On retiendra surtout de ce projet la proposition d'utiliser un micro virtuel localisé dans l'espace tridimensionnel de la simulation en cours (voir la figure 2.10). Ce micro possède deux modes de fonctionnement. Le premier correspond simplement à un micro attaché à la caméra virtuelle. C'est son utilisation standard : ce que l'on entend est synchronisé à ce que l'on voit. La seconde option est plus originale et interactive : on propose à l'utilisateur de déplacer le micro à l'intérieur de la simulation 3D. Cela permet de focaliser son attention sonore sur une certaine partie du complexe moléculaire, quand bien même la représentation visuelle n'a pas été modifiée. Ce micro s'accompagne visuellement d'une sonde sphérique qui représente un système de filtrage

⁴⁹ Une liaison ou un pont hydrogène est une force intermoléculaire impliquant un atome d'hydrogène et un atome électronégatif, comme l'oxygène ou l'azote par exemple.

spatial dont l'utilisateur peut contrôler la portée et la force. Ce filtrage fait que les événements occurrents à l'extérieur de la sonde ne sont plus sonifiés. On retiendra pour notre propre conception de sonification l'idée d'une sonde sphérique permettant d'éliminer d'un jeu de données les informations que l'on ne souhaite pas sonifier. L'utilisation que l'on propose d'en faire est cependant assez différente (cf. 3.1.3).

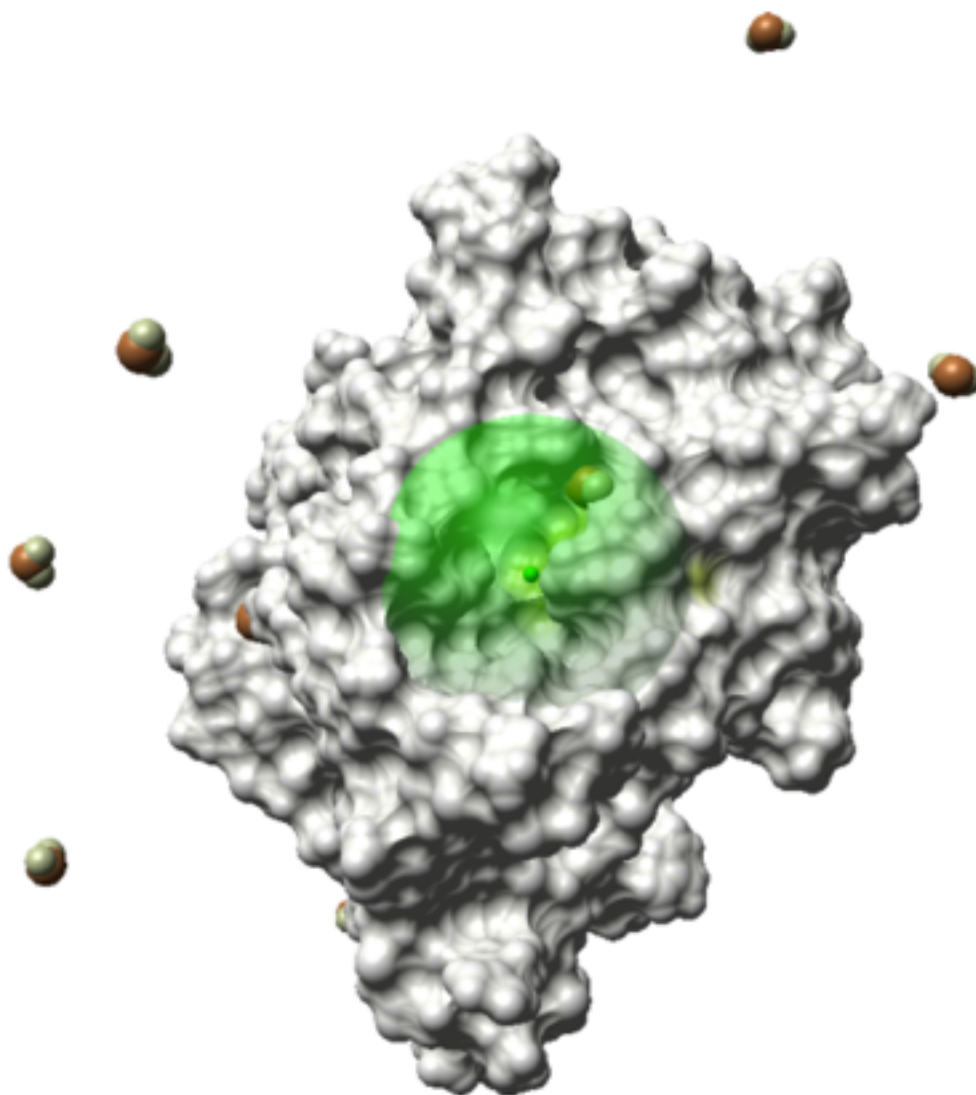


Fig. 2.10 : Le point vert au milieu représente la position du microphone virtuel, tandis que la sphère qui l'entoure illustre sa portée d'action [Rau et al., 2015]

Cette intégration de l'interactivité au sein de la sonification est l'un des principaux enjeux actuels de la discipline. Le domaine de la sonification moléculaire n'y a de fait pas échappé.

2.2.3. Le développement de l'interactivité

Que ce soit sur l'aspect séquentiel, paramétrique ou dynamique des données moléculaires, la majorité des sonifications détaillées auparavant sont figées jusqu'à un certain point. L'utilisateur peut évidemment garder le contrôle du volume, voire du type de sonification si l'interface en propose plusieurs ; mais son contrôle sur l'outil s'arrête là. L'interaction qu'il peut avoir les données sonifiées de la protéine est donc moindre. Un pan de la recherche en sonification de molécules s'est donc intéressé aux différents types d'interactions que peut avoir un sujet avec les données sonifiées, en partant du principe que si l'utilisateur est considéré comme un acteur et non plus un simple observateur, son expérience et ses performances n'en seront que décuplés. On peut par exemple lui donner la possibilité et même la décision de déclencher la sonification, ou encore lui laisser le contrôle de l'échelle de *mapping* ou des données qu'il souhaite sonifier.

On aurait pu citer le projet de *plugin* suivant dans la sous-partie précédente, puisqu'il propose également parmi d'autres une sonification paramétrique du facteur B. Développé cette fois-ci pour *Chimera* par [Ballweg et al., 2016], le *plugin* en question utilise un *mapping* de paramètres plus ou moins similaire à celui de *SUMO*. Les valeurs du facteur B sont encore une fois « mappées » exponentiellement à la fréquence d'un sinus légèrement distordu. Cette distorsion permet de définir un simili timbre harmonique au sinus. Comme les sons joués sont spatialisés en Ambisonic décodé en binaural au casque, cela rend le son plus facile à localiser que pour un son pur. La principale différence avec *SUMO* réside dans le simple fait que tous les paramètres relatifs au *mapping* de la fréquence du sinus et à sa distorsion sont entièrement modifiables par l'utilisateur (voir la figure 2.11).

Pour activer la sonification, il lui suffit de cliquer à la souris sur un résidu ou un atome pour jouer le son correspondant à son facteur B⁵⁰. La partie de la molécule sélectionnée va alors changer de couleur temporairement. Ce léger *feedback* visuel permet notamment à l'utilisateur de visualiser plus facilement le point sélectionné lorsque la valeur du zoom est moindre et que la représentation de la protéine est particulièrement chargée. Il est également possible de paramétrer le déclenchement de la sonification pour qu'un clic sur un atome actionne aussi la sonification de toute la zone périphérique au point sélectionné, avec une propagation de proche en proche sur les

⁵⁰ Une vidéo de démonstration de la sonification est disponible sur YouTube : <https://youtu.be/IjQrL1T3h0s?t=1m56s>.

atomes voisins. La vitesse de propagation et le radius de cette onde sphérique progressive sont aussi contrôlables par l'utilisateur. Cela permet de déterminer la vitesse à laquelle les points alentours vont être sonifiés, ainsi que l'atténuation progressive de l'énergie de l'onde et donc la baisse consécutive du niveau des sons à mesure que l'on s'éloigne de l'épicentre. À noter que le changement de couleur temporaire des atomes est aussi couplé à la propagation de cette onde, l'intensité de la couleur diminuant avec la propagation de l'onde.

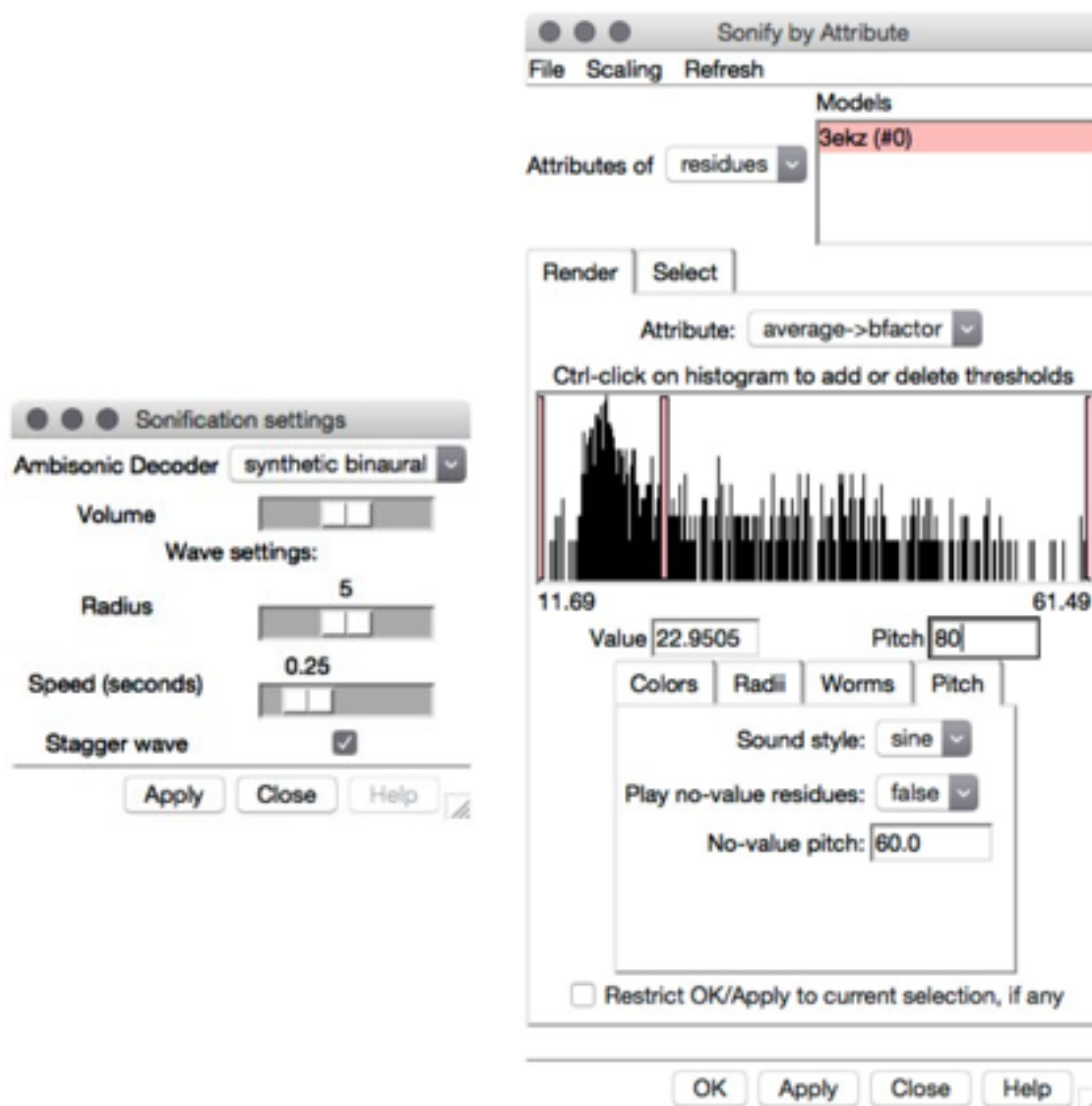


Fig. 2.11 : Différentes fenêtres de l'interface utilisateur du *plugin* développé par [Ballweg et al., 2016] pour Chimera

D'autres travaux ont mis en avant l'interactivité et le contrôle laissé à l'utilisateur dans leurs sonifications, que ce soit pour décrire la forme de structures d'ARN [Grond et al., 2010] ou

proposer une expérience de *docking* moléculaire en réalité virtuelle [Delalande, 2009]. Ce dernier projet, nommé CoRSAIRe (Combinaisons de Rendus Sensori-moteurs pour l'Analyse Immersive de Résultats), donne en effet la possibilité à un sujet qui est placé devant un écran vidéoprojeté de tester différentes positions de *docking* avec un contrôleur haptique qui fait office de capteur de position et d'orientation. Les collisions entre les deux protéines sont alors sonifiées en temps réel, avec notamment une spatialisation ambisonique des sources sonores ; le tout afin de lui permettre de juger au mieux de la justesse du *docking*. La principale force de l'interface de CoRSAIRe réside donc dans l'utilisation simultanée de trois modalités différentes (visuelle, haptique et sonore).

On peut donc dégager de cet état de l'art de la sonification de molécules deux grandes tendances : le travail sur les données séquentielles et le travail sur les données paramétriques (fixées ou dynamiques). On aura aussi mis en avant le développement de l'interactivité ainsi que les apports de la spatialisation 3D.

Mais notre propre proposition ne porte pas sur la sonification des données séquentielles ou paramétriques que l'on peut lier aux structures des protéines. L'aspect novateur de notre hypothèse de travail porte plutôt sur la sonification de la forme et la surface même des molécules, pour potentiellement renseigner sur leurs propriétés plutôt que de justement sonifier ces mêmes propriétés. En cela, nous nous rapprochons davantage de la sonification de formes et d'objets 3D, domaine d'études que nous allons donc aborder dans la sous-partie suivante.

3. La sonification d'objets 3D

Diverses techniques ont été développées pour proposer une représentation sonore de volumes et d'objets 3D. En effet, une utilisation réfléchie de stimulus acoustiques combinés à une représentation visuelle aide grandement à l'immersion, en particulier dans des environnements interactifs tridimensionnels [Röber et al., 2005]. De plus, l'exploration visuelle et sonore de ces environnements permet une meilleure préhension des anomalies locales ou globales de l'objet représenté.

Sont donc détaillés dans cette sous-partie quelques exemples résumant les principales approches et orientations du domaine. À travers ce court panorama, on tentera d'extraire de chaque projet des idées probantes vis-à-vis de la conception de notre partie pratique.

2.2.1. Coupe transversale d'un volume

Le projet SATIN [Alonso-Arevalo et al., 2008, 2012] consiste en une « discrétisation » 2D de la forme d'un objet 3D, *via* la représentation de ce dernier par une coupe transversale (voir la figure 2.12).

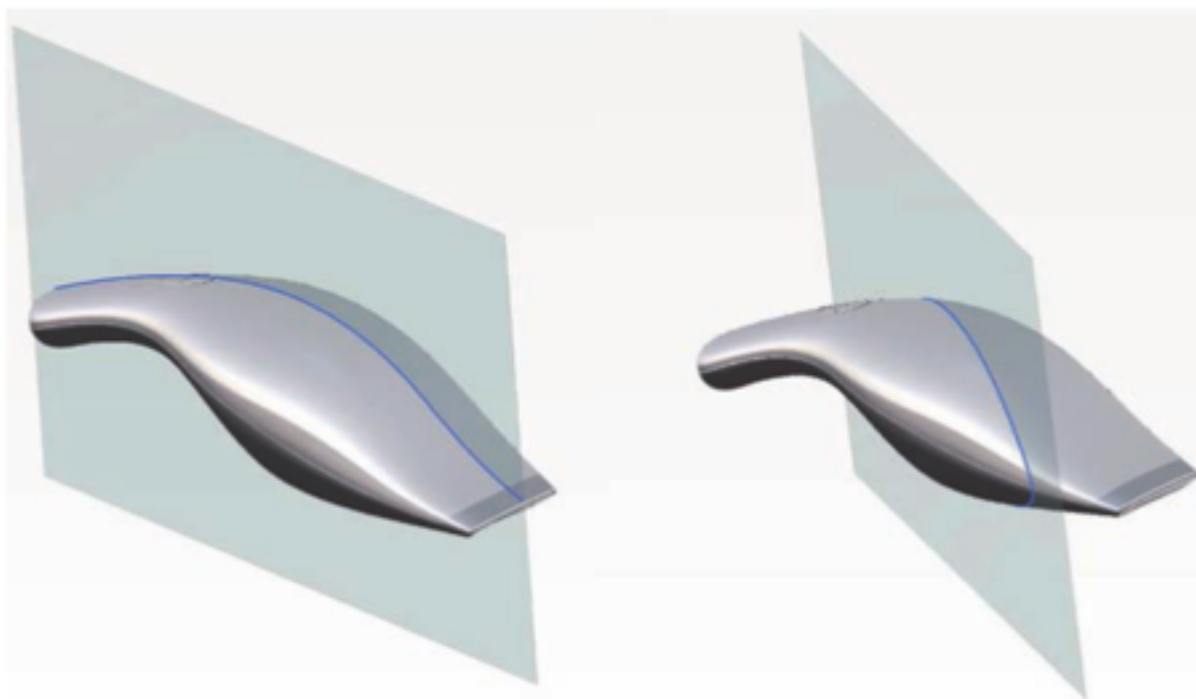


Fig. 2.12 : Exemple de coupes transversales d'un objet 3D [Alonso-Arevalo et al., 2012]

La position et l'angle de cette coupe sont choisis par l'utilisateur en déplaçant un doigt sur un contrôleur haptique tactile. On peut alors représenter sa silhouette de deux façons différentes : en sonifiant au choix la courbe ou bien la courbure⁵¹ associées au contour de la coupe obtenue (voir la figure 2.13). Ainsi, la fréquence du son produit lorsque que l'on touche le contrôleur peut être « mappée » directement sur la courbe obtenue ou fonction de la courbure de la coupe. Avec ces deux propositions de *mapping*, on cherche d'abord à renforcer la compréhension de l'utilisateur de la forme globale d'un prototype virtuel puis, dans un second temps, à informer l'utilisateur de possibles attributs géométriques qui seraient autrement restés invisibles de par les limitations du logiciel de visualisation et/ou de la perception humaine.

⁵¹ On appelle « courbure » d'un objet géométrique la mesure quantitative du caractère plus ou moins courbé de cet objet. La valeur de ce caractère est définie par le rayon de courbure, qui correspond au rayon du cercle permettant d'approximer au mieux la courbe en un point donné.



Fig. 2.13 : Illustration de la courbe et du rayon de courbure d'une coupe transversale d'un objet 3D [Alonso-Arevalo et al., 2012]

Cette nécessité de passer par une discrétisation 2D de l'objet que l'on souhaite représenter est une problématique récurrente dans les travaux relatifs à la sonification de formes et/ou de surfaces. Quelques études ont cependant pris le contre-pied de cette tendance, notamment en cherchant à favoriser l'immersion spatiale au cœur de l'objet que l'on souhaite représenter. Pour cela, elles se sont basées sur un découpage spatial du volume tridimensionnel, que ce soit en voxel ou en nuage de points.

2.2.2. Le voxel

Analogue tridimensionnel du pixel⁵², le voxel (contraction de *volume* et *element*) permet de représenter la structure interne ou externe de n'importe quel volume (voir la figure 2.14). Chaque voxel contient différentes informations comme sa colorimétrie⁵³ et surtout, sa principale différence avec le pixel, des coordonnées spatiales de position sur trois dimensions (x, y, z) au lieu de deux. Le découpage en voxel est notamment utilisé en sonification d'objets 3D car il permet de circonscrire facilement un volume en différentes zones identiques en taille et dont le principal attribut différentiel est leurs coordonnées. Cette discrétisation de la forme de l'objet offre la possibilité d'attribuer à chaque voxel une source sonore localisée ou des attributs relatifs à une source sonore externe.

⁵² Le « pixel » (souvent abrégé en px) est l'unité de base permettant de mesurer la définition d'une image numérique matricielle. Son nom provient de la locution anglaise *picture element*, qui signifie « élément d'image ». C'est le plus petit élément constitutif d'une image numérique. Généralement de forme rectangulaire ou carrée, une couleur ou un niveau de gris ainsi qu'une intensité lui sont associés en plus de ses coordonnées en (x, y) dans l'image.

⁵³ La colorimétrie d'un pixel ou d'un voxel se décompose en trois valeurs : celles-ci peuvent correspondre à un codage trichrome RVB (Rouge Vert Bleu) ou bien un codage perceptuel TLS (Teinte Saturation Luminosité).

2. Prérequis à la sonification de protéines

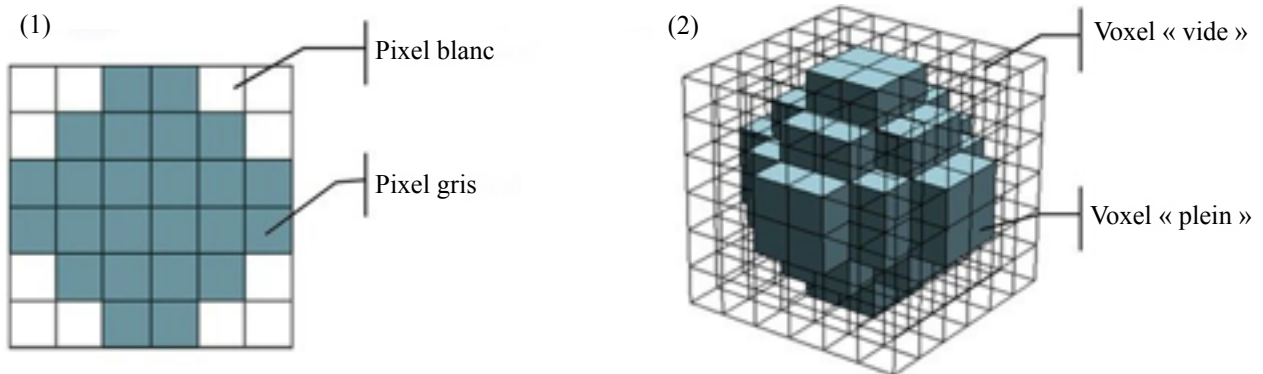


Fig. 2.14 : (1) Image basse résolution d'un cercle (2) Modèle basse résolution d'une sphère [Aremu et al., 2016]

Par exemple, Rossiter et Ng [1996] ont fait le choix d'indexer tous les voxels d'un volume (voir la figure 2.15). Puis, ils ont fait correspondre leurs différentes valeurs à certains paramètres tels que l'amplitude, le timbre ou la fréquence fondamentale d'instruments virtuels, de sons synthétisés (typiquement à partir d'un sinus) mais aussi de fichiers audio.

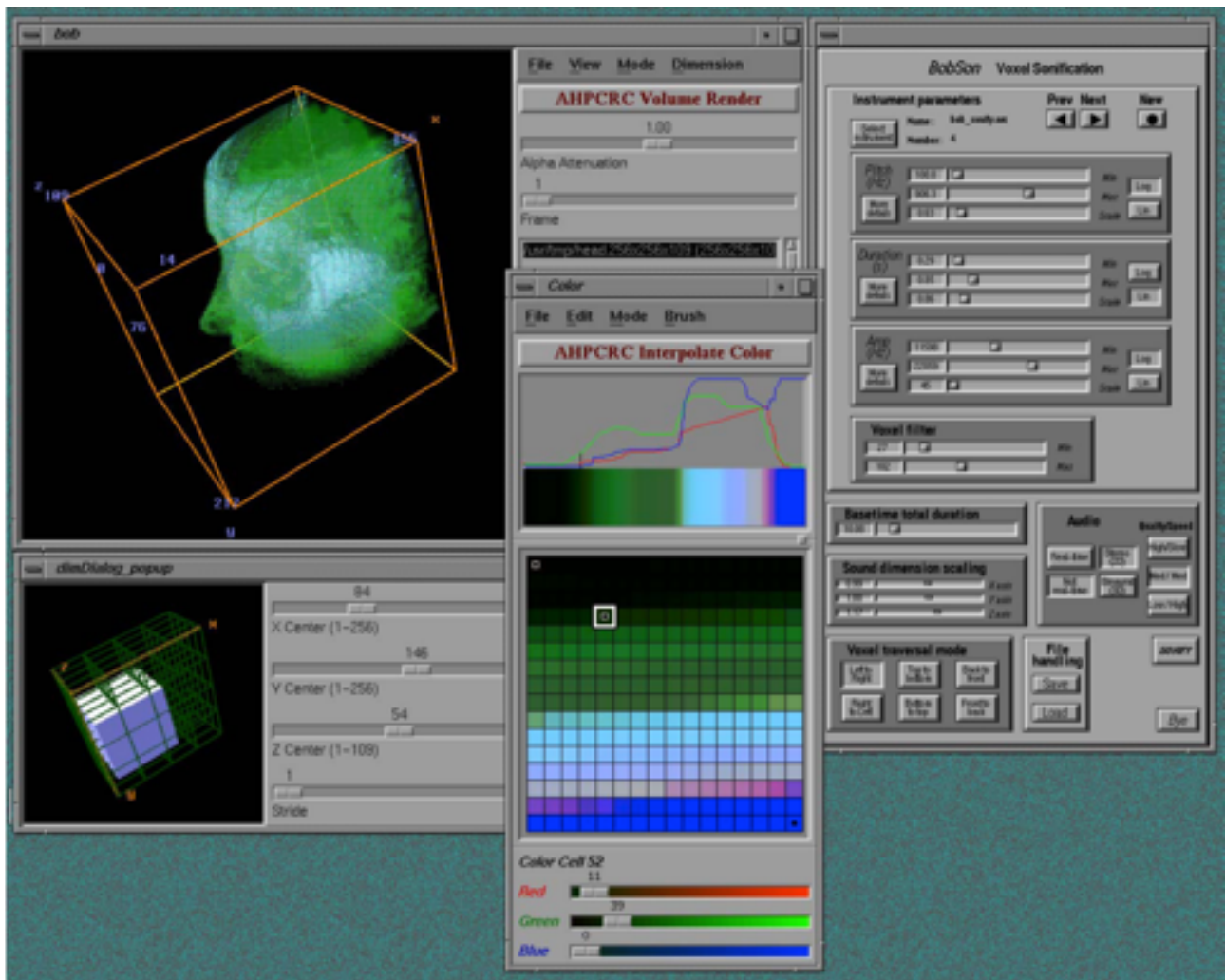


Fig. 2.15 : Capture d'écran de l'interface utilisateur proposée par [Rossiter et Ng, 1996]

Cette méthode de sonification consiste à associer à chaque voxel composant l'objet 3D un évènement sonore spatialisé selon la position du voxel qu'il représente. La restitution se faisant au casque, cette spatialisation peut être stéréo (avec des effets de panoramique pour simuler la position horizontale et des effets de niveau pour simuler la profondeur) ou bien binaurale. Pour donner un exemple imagé du fonctionnement de cette interface, le sujet peut ainsi décider de choisir un *mapping* où la luminosité des voxels à la fois proportionnelle à la fréquence fondamentale d'un sinus et à sa durée. Ainsi, traverser de part en part un volume fait de quelques voxels sombres dispersés en périphérie d'une masse de voxels très clairs revient alors à parcourir un champ sonore où l'on va d'abord entendre quelques sons graves éparses et très brefs avant d'atteindre une masse sonore très dense faite de sons aigus quasi-continus.

Gomez et Ramirez [2011] ont pour leur part construit une sonification de la cartographie tridimensionnelle de l'activité d'un cerveau mesurée lors d'un test cognitif (voir la figure 2.16). L'une des deux sonifications par *mapping* de paramètres qu'ils proposent consiste à synthétiser une nappe sonore à partir des voxels correspondant à une zone du cerveau en activité.

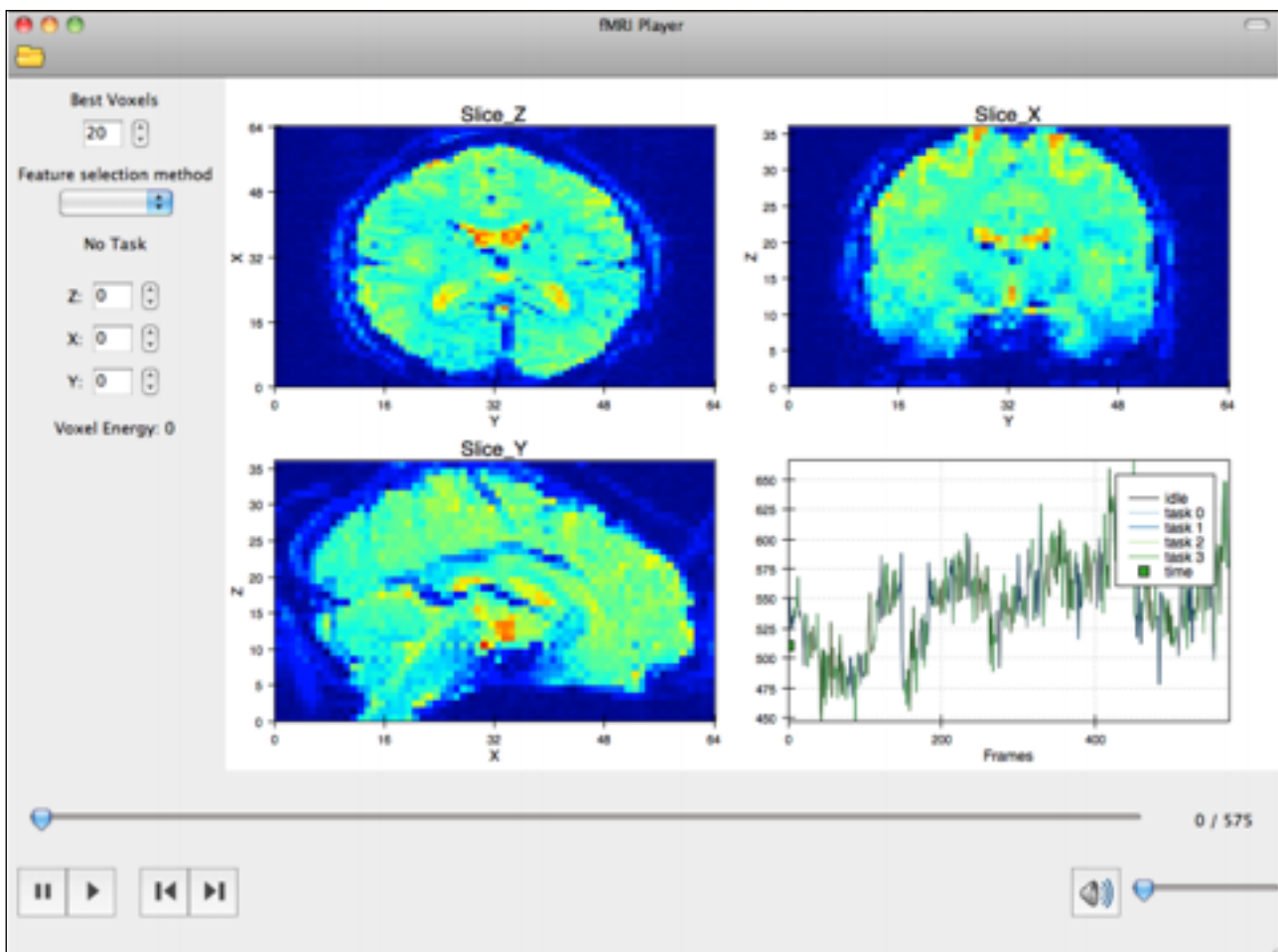


Fig. 2.16 : Capture d'écran de l'interface utilisateur du projet de [Gomez et Ramirez, 2011]

Ainsi, un son pur est associé à chacun de ces voxels, en « mappant » leur énergie (et donc le niveau d'activité, représenté par la couleur dans la figure 2.16) à la fréquence du sinus. Si le résultat final peut parfois se rapprocher d'un bruit blanc, cette méthode donne malgré tout une bonne représentation de l'évolution dans le temps du niveau d'activité de la zone du cerveau préalablement sélectionnée.

Enfin, le projet SSound [Minghim et Forrest, 1995] propose parmi plusieurs sonifications différentes une méthode qui permet d'obtenir une représentation sonore en voxels à partir d'un volume tridimensionnel. Elle consiste à projeter une grille tridimensionnelle sur l'objet 3D désiré, afin de le « découper » en voxels (voir la figure 2.17). L'utilisateur va alors pouvoir déplacer sa souris au sein du volume. Chaque voxel qu'il traverse va changer de couleur et déclencher la lecture d'un son produit par un synthétiseur connecté en MIDI et dont la hauteur de jeu est mise en correspondance avec la valeur du paramètre du voxel que l'on a sélectionné pour le *mapping* (plus la valeur choisie dans le *mapping* est élevée, plus la fréquence est haute). Le son est également spatialisé en stéréo selon la position du voxel joué.

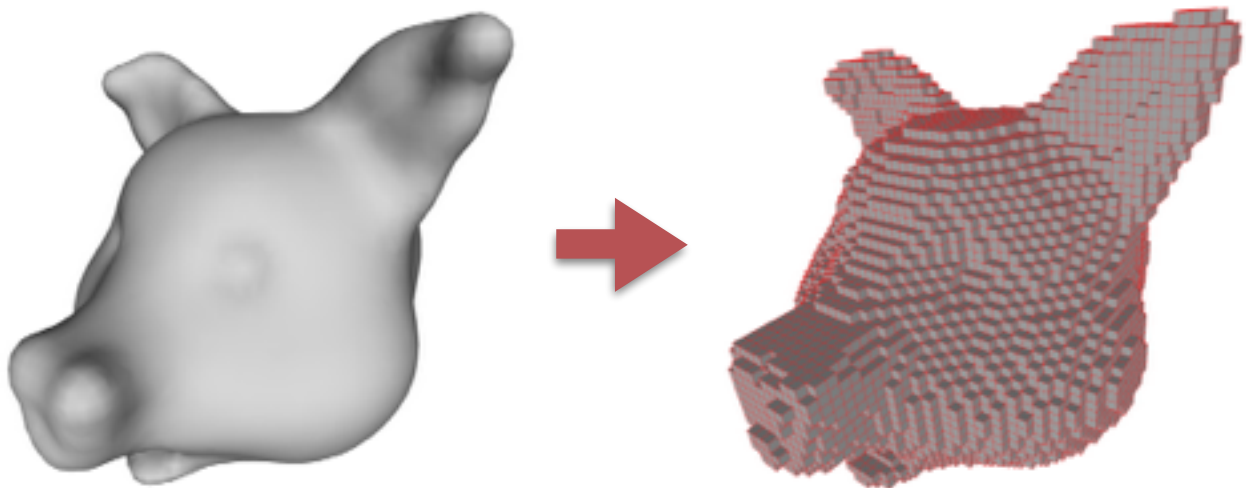


Fig. 2.17 : Discretisation d'un volume 3D en voxel (figures par Andreas Bærentzen)

Découper spatialement un objet 3D en de multiples voxels est donc très avantageux pour affiner la précision de la sonification et potentiellement y introduire une composante spatiale. La représentation en nuage de points représente l'aboutissement de cette réflexion.

2.2.3. Exploration d'un nuage de points

En plus de sa sonification de voxels, le projet SSound que l'on vient de mentionner propose également une sonification de l'isosurface d'un volume 3D. L'isosurface est un rendu graphique proche du nuage de points puisqu'elle consiste en un lissage polygonal de la représentation en voxel (voir les figures 2.18 et 2.19), où l'approximation de la surface par des triangles donne de fait un nuage de points (chaque sommet de triangle constituant un point).



Fig. 2.18 : Représentation 2D du lissage des voxels par l'isosurface

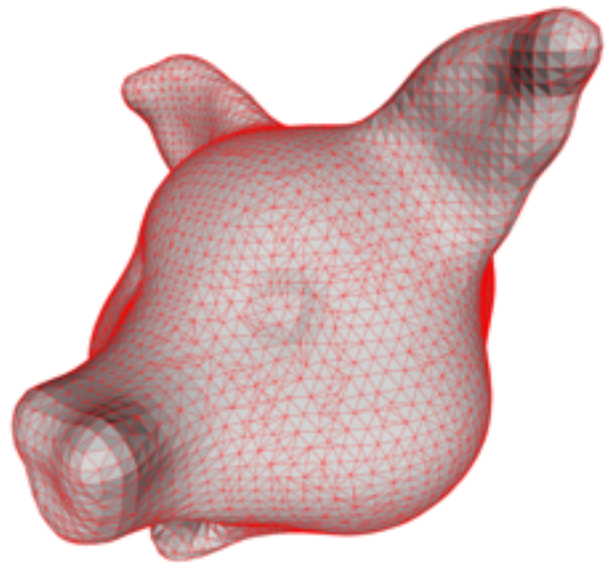


Fig. 2.19 : Représentation de l'isosurface d'un volume 3D (figure par Andreas Bærentzen)

L'objet que l'on souhaite sonifier est ensuite placé dans un monde virtuel tridimensionnel borné par une « boîte de délimitation » (*bounding box* dans le texte, mais on trouve plus souvent la terminologie *world box* dans la littérature concernée par le sujet). À l'intérieure de cette boîte évolue une sonde volumétrique cubique (*volume probe*), que l'utilisateur peut déplacer et agrandir ou rétrécir (voir la figure 2.20). À noter que l'interface peut sauvegarder la position et la taille de la sonde afin de pouvoir comparer de la même manière deux volumes différents.

Ensuite, on peut choisir de sonifier l'intégralité du volume ou bien ce qui est à l'intérieur de la sonde : la densité de triangles dont est composée la surface est alors « mappée » en fréquence (les hautes densités en basses fréquences et inversement). Une option débrayable permet également de jouer un second son qui indique lui la portion du volume total de l'objet que la sonde contient. Les effets de panoramique sont cette fois-ci utilisés pour indiquer l'orientation de la sonde à l'intérieur

du volume : les coordonnées « basses » sont spatialisées vers la gauche et les « hautes » vers la droite. Le volume est quant à lui inversement proportionnel à la fréquence, afin de compenser la non-linéarité en fréquence de notre perception auditive.

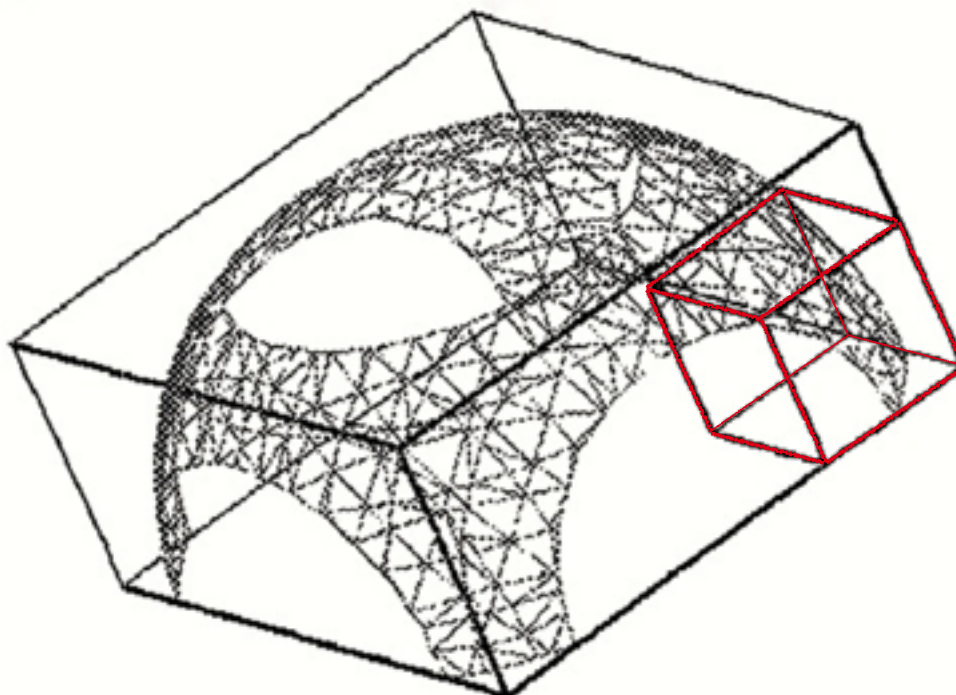


Fig. 2.20 : L'objet 3D ici représenté est une sphère dont certaines parties sont coupées par la *bounding box*. On aperçoit dans le coin à droite la sonde volumétrique cubique en rouge [Minghim et Forrest, 1995]

Toujours dans cet esprit de sonde volumétrique, Stockmann et al. [2008] ont proposé une interface qui permet à un utilisateur de déplacer virtuellement une baguette de carillon (*chime rod* dans le texte) au milieu d'un nuage de point généré par le volume 3D que l'on souhaite représenter et explorer (voir la figure 2.21).

Chaque collision du faisceau vertical avec un point du nuage déclenche un son : la hauteur de ce dernier dépend de la position du contact sur le tube (plus le contact est haut, plus la hauteur est élevée), tandis que son amplitude est proportionnelle à la densité de points contenus dans la zone de l'objet 3D activée. Par ailleurs, les sons résultant des collisions sont également spatialisés afin d'augmenter l'immersion de l'expérience. Même si ce projet reste avant tout pensé comme un instrument de musique et de création plus qu'un outils pour la compréhension de données, il met en

avant des mécanismes d'interactivité et d'immersion que l'on souhaite pouvoir retrouver dans notre propre interface de sonification.

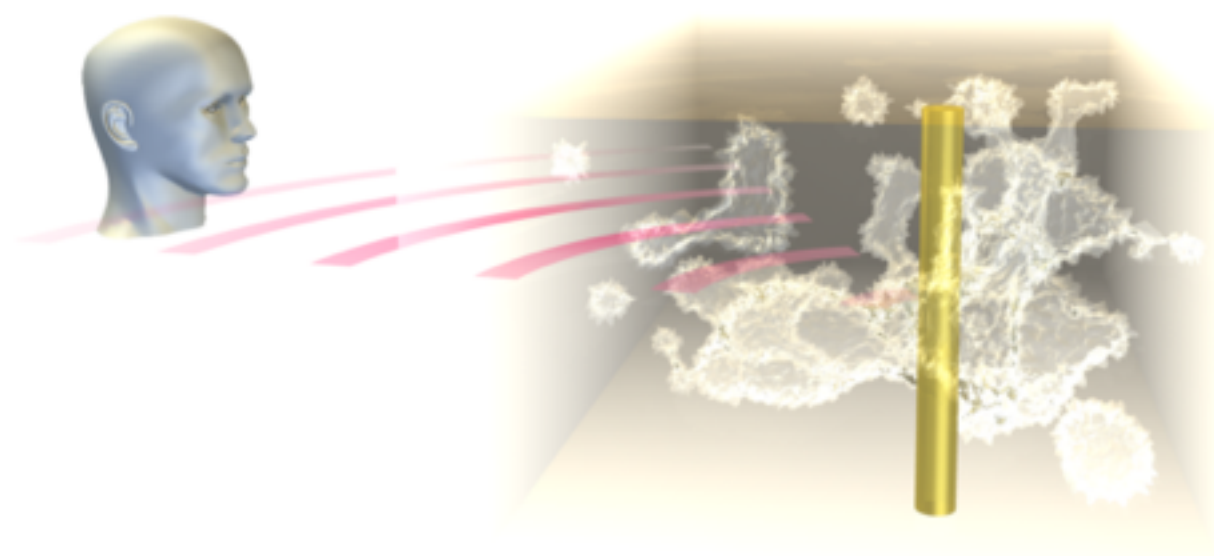


Fig. 2.21 : Illustration du scan d'un volume à l'aide d'un contrôleur haptique faisant office de capteur de position [Stockmann et al., 2008]

On retiendra l'idée commune à ces deux dernières sonifications d'approximer la forme d'un objet tridimensionnel en passant par sa représentation en nuage de points, comme autant de données à sonifier et de sources sonores potentiellement spatialisables en trois dimensions.

4. Le son 3D

2.4.1. La localisation auditive et ses performances

La localisation auditive correspond à notre capacité à identifier la position d'une source sonore dans l'espace. On sait que cette capacité est beaucoup plus précise dans le plan horizontal qu'en élévation et en distance [Blauert, 1997]. Elle est en effet basée sur deux indices acoustiques différents, qui permettent à notre système auditif de localiser les sources sonores principalement dans le plan horizontal.

Le premier de ces indices est la différence interaurale de temps (ITD, pour *Interaural Time Difference*), liée à la différence des trajets acoustiques de la source aux deux oreilles. Cette

différence temporelle est directement occasionnée par un retard entre les instants d'arrivée de l'onde à chaque oreille. Lorsque la source est complètement latéralisée, cette différence de temps est d'environ 0,6 ms. Une différence de phase permet ainsi la localisation, sauf quand la longueur d'onde est plus petite que la distance entre les deux oreilles ; puisque le déphasage devient alors supérieur à 180° . La différence de temps permet seulement la localisation des fréquences inférieures à environ 1500 Hz.

Le second indice est une différence interaurale d'intensité (IID/ILD, pour *Interaural Intensity/Level Difference*), liée à la diffraction engendrée par notre tête. Cette différence d'intensité est due au fait que la tête peut être un obstacle à la propagation du son, spécifiquement pour les hautes fréquences, auxquelles sont associées des longueurs d'onde petites par rapport aux dimensions de notre boîte crânienne. Cette dernière crée ainsi une « ombre acoustique » pour les fréquences supérieures à 1500 Hz environ.

La localisation peut être utilisée en sonification pour fournir des informations de position et/ou de direction en corrélation avec l'image, mais elle peut également servir à discrétiser des flux sonores (cf. 1.3.3). Néanmoins, il faut bien garder à l'esprit que la localisation n'a pas une précision homogène partout dans l'espace. Par exemple, si l'on est capable de discriminer deux sources espacées de 1° si elles sont émises en face de nous, ces deux mêmes sources doivent être espacées de 15° lorsque d'elles sont placées à un azimut de 90° [Strybel et al., 1992].

De nombreuses sonifications se sont malgré tout emparées du son 3D et du binaural (avec notamment pour but de donner l'impression que le son issu d'écouteurs ou d'un casque provient de l'extérieur de la tête), car la localisation auditive permet autant de guider l'attention visuelle que d'éviter la surcharge d'informations visuelles en les remplaçant par du son spacialisé.

2.4.2. L'Ambisonic

L'Ambisonic (en français « ambisonie ») est un procédé développé par Michael Gerzon. Il consiste en une méthode de captation, d'encodage à l'enregistrement puis de reproduction du son en trois dimensions [Gerzon, 1980]. Pour cela, deux étapes sont nécessaires :

- Encodage : on reproduit le champ sonore complet en décomposant ses caractéristiques directionnelles comme les composantes harmoniques d'une sphère. À l'ordre 1, ce modèle en comprend quatre (voir la figure 2.22) : W (composante globale omnidirectionnelle qui correspond à la pression de l'air en un point fixé) et X, Y et Z (composantes directionnelles orthogonales qui correspondent au gradient de pression au point fixé).

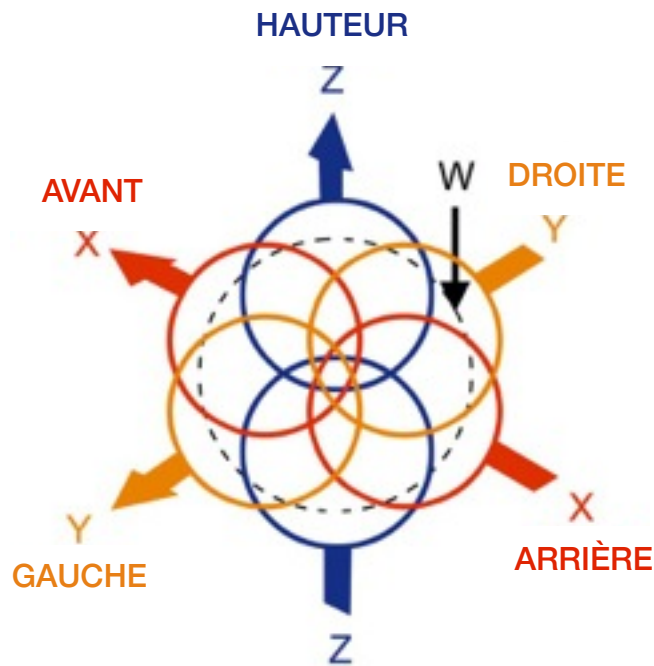


Fig. 2.22 : Représentation en trois dimensions du champ acoustique pour de l'Ambisonic d'ordre 1

- Décodage : le décodage consiste en la restitution du champ sonore par un réseau de haut-parleurs. Chaque haut-parleur émet un son résultant d'une combinaison particulière des composantes W, X, Y, Z. L'intérêt de l'Ambisonic est de pouvoir reproduire le champ sonore avec n'importe quel nombre de haut-parleurs (à partir de deux) et dans n'importe quelle configuration (la répartition des haut-parleurs selon une sphère donne toutefois de meilleurs résultats).

Keating [1996] a décrit dans un compte-rendu les avantages de l'Ambisonic dans des contextes de réalité virtuelle et d'environnements immersifs. Cependant, un système de restitution Ambisonic est coûteux et lourd à mettre en place. Or, nous avons besoin que notre interface puisse fonctionner au casque et de manière nomade. On décide donc de se tourner vers d'autres méthodes

de spatialisation tridimensionnelle comme le binaural, qui utilise des *Head Related Transfer Function* (HRTF).

2.4.3. La synthèse binaurale

La synthèse binaurale est une méthode qui permet de construire virtuellement un espace sonore à partir d'informations morphologiques propres à l'utilisateur qu'on appelle *Head Related Transfer Function* (HRTF). Il s'agit en fait de reproduire le champ sonore tel qu'il serait perçu au niveau des oreilles d'un utilisateur. Pour cela, le rendu sonore de chaque source monophonique que l'on souhaite spatialiser est basé sur un filtrage linéaire du signal par un couple de HRTF propre à chacun et contenant tous les indices physiques de localisation qui sont utilisés par notre système auditif.

Comme les HRTF dépendent de la morphologie de chaque individu, elles doivent donc être individualisées afin de conserver des indices de localisation probants. Cependant, obtenir cette individualisation des HRTF est un processus compliqué et fastidieux. En effet, les systèmes permettant actuellement de mesurer efficacement des HRTF sont difficiles d'accès au grand public.

De nombreux travaux se sont donc penchés sur ces problèmes d'individualisation des HRTF, afin de rendre possible leur utilisation par n'importe quel individu [Parseihian et Katz, 2012]. Nous ne nous penchons pas d'avantage sur ces questions d'individualisation car nous jugeons que les sensations d'espace et de séparation des sources avec un rendu binaural non individualisé sont suffisantes pour l'utilisation que nous avons à en faire.

Le principal défaut du binaural réside dans les ressources nécessaires au calcul en temps réel du filtrage de chaque source ; calculs qui augmentent donc avec le nombre de sources. Pour pallier ce problème, nous avons décidé d'utiliser une méthode d'Ambisonic virtuel. La séparation entre l'encodage et le décodage dans la synthèse ambisonique permet en fait de remplacer la phase de décodage ambisonique (pour un rendu sur haut-parleurs) par un décodage binaural pour une restitution au casque. On parle alors d'Ambisonic virtuel [Noisternig et al. 2003]. Cela revient à spatialiser en binaural ce qu'aurait émis chaque haut-parleur dans une restitution ambisonique, au lieu des sources sonores elles-mêmes. On peut ainsi combiner les avantages de la synthèse ambisonique et du binaural : le temps de calcul reste relativement faible (encodage ambisonique),

même pour un très grand nombre de sources, et la restitution se fait au casque (décodage binaural) et limite ainsi l'équipement nécessaire à l'utilisateur pour utiliser l'interface. De plus, on peut mentionner le fait que le rendu binaural résout partiellement le problème d'externalisation [Begault, 1994 ; Nicol, 2010].

5. Conclusion

Nous avons donc vu dans cette seconde partie que la forme des protéines renseigne potentiellement sur leurs fonctions et les possibles interactions qu'elles peuvent avoir entre elles. C'est pourquoi les chercheurs se sont intéressés à leur modélisation, notamment grâce à des logiciels de visualisation. Cependant, ces derniers possèdent certaines limites que la sonification semble pouvoir aider à dépasser. À travers un examen non-exhaustif des différents types de sonification de molécules, nous avons conclu que la majeure partie des méthodes et techniques déjà développées se focalisaient principalement sur des données séquentielles, paramétriques ou événementielles que l'on peut lier à la structure et donc la forme d'une protéine, mais sans jamais s'intéresser directement à cette même forme. En partant de là, un bref panorama de différentes méthodes de sonification de forme tridimensionnelle nous a permis d'apprécier certains types de représentation, dont celle en nuage de points. Cette représentation est de plus particulièrement adaptée à la restitution d'un champ sonore en trois dimensions. On comprend donc l'intérêt d'utiliser la spatialisation dans notre sonification.

3. Conception d'une nouvelle représentation sonore

La partie pratique de ce mémoire consiste en la conception puis l'implémentation d'une interface de sonification d'objets 3D, en particulier des petites protéines. On détaillera dans ce chapitre les aspects théoriques de notre proposition de sonification, en justifiant notamment les choix que nous avons fait pour notre *mapping*.

3.1. Présentation générale du projet

On souhaite que notre sonification puisse renseigner sur la forme d'un objet 3D représentant une protéine. On s'intéresse en particulier à tout ce qui pourrait caractériser la forme d'une protéine par rapport à une autre : le relief général de sa surface, ses possibles excroissances ainsi que ses spécificités de convexité et de concavité.

3.1.1. Idée principale

La méthode qui nous a semblé la plus directe et la plus efficace pour renseigner sur la forme d'un objet tridimensionnel consiste à associer une source sonore à chaque point composant la surface de l'objet 3D en question. Les paramètres sonores de cette source sont alors liés à la position du point dans l'espace ; c'est donc une sonification par *mapping* de paramètres (cf. 1.3.2). Par exemple, chaque source sonore sera spatialisée selon les coordonnées du point qu'elle représente.

Notre panorama des différentes méthodes de sonification d'objets 3D nous a permis de confirmer la pertinence et la cohérence d'une telle approche. De plus, la spatialisation apparaît comme une solution naturelle pour représenter de manière immersive des données d'origine spatiale. En effet, placer virtuellement une source sonore au sein d'un espace fictif permet de profiter des capacités naturelles de notre système auditif à localiser une source dans l'espace [Wenzel, 1992, 1994 ; Begault, 1994]. Certains travaux suggèrent même qu'une simple différence de position spatiale entre différentes sources similaires suffit à les ségréguer [David et al., 2014, 2015, 2017].

En se basant sur notre capacité à localiser des sources sonores, on décide également de placer le point d'écoute de l'utilisateur à l'intérieur de la protéine. Ce choix permet surtout de répartir le nombre total de sources sonores autour de lui et donc de mieux séparer spatialement les points les uns des autres. On utilise pour cela le son 3D (voir la sous-section 2.4).

3.1.2. Hypothèse de travail

Notre hypothèse de travail consiste donc à penser que l'utilisateur arrivera à dégager de la projection sonore de l'objet 3D des informations générales sur sa forme et le relief de sa surface.

Attention, l'idée n'est pas de suppléer à la représentation visuelle mais de la soutenir, l'affiner. De plus, aux vues des nombreuses imperfections du son spatialisé, il ne faut pas s'attendre à ce que la sonification proposée permette d'identifier précisément la forme de l'objet sans visualisation. Il s'agit encore moins de localiser chaque son individuellement. Notre proposition a avant tout un rôle de support et d'aide. On cherche seulement à explorer une piste qui n'a pas été poussée dans le domaine de la sonification de protéine. Et quand bien même nous n'obtiendrions pas de nouveaux résultats, nous pourrions au moins dégager de ce mémoire des recommandations utiles à d'éventuels futurs travaux de recherche.

3.1.3. Limites et contraintes

Nous n'envisageons pas encore de sonifier la totalité des points servant à composer la surface d'un objet 3D. En effet, le nombre résultant de sources sonores à spatialiser peut être conséquent et il existe un risque non négligeable que cela conduise à une surcharge cognitive et perceptive pour l'utilisateur. Ainsi, bien qu'il ne soit pas nécessaire que celui-ci distingue précisément chaque source séparément, il semble pertinent de limiter la scène sonore perçue (et donc le nombre de sources). Cela permet de faciliter sa compréhension du lien « surface 3D-rendu sonore ».

Nous avons donc cherché un moyen de limiter le nombre de sources sonores à restituer. Une première solution consiste à réduire la résolution du maillage polygonal qui sert à représenter la surface qui nous sert de base pour la sonification. On parle alors de raffinement par décimation de

vertex (voir la figure 3.1). Il s'agit simplement de diminuer le nombre de triangles dont est composée la surface de l'objet restitué. Cela revient en quelque sorte à diminuer sa résolution.



Fig. 3.1 : Raffinement du maillage d'une sphère par décimation de vertex

Nous introduisons également dans notre sonification deux solutions complémentaires pour diminuer encore d'avantage le nombre de sources sonores à diffuser simultanément :

- Tout d'abord, on permet à l'utilisateur de choisir la quantité de points qu'il souhaite sonifier. Pour cela, nous mettons à sa disposition une sphère de filtrage basée sur le modèle de [Rau et al., 2015]⁵⁴ (voir la figure 3.2).

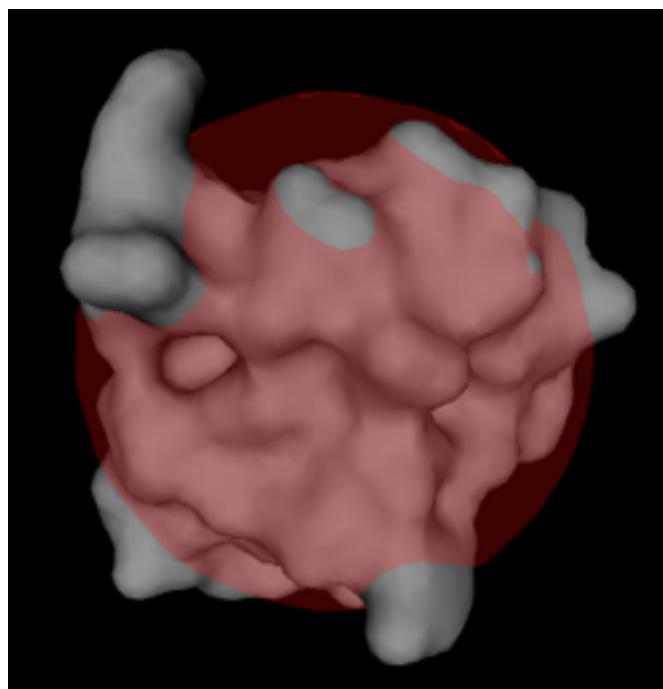


Fig. 3.2 : Illustration du principe de la sphère de filtrage. Ce procédé permet de représenter les appendices de la protéine tout en limitant le nombre de sources à sonifier.

⁵⁴ Ce modèle est décrit dans la sous-partie 2.2.2.

Centrée sur l'objet 3D, cette sphère agit à la manière d'un filtre : aucun des points compris à l'intérieur ou sur la sphère n'est lu. Seuls les points qui lui sont strictement extérieurs sont sonifiés. Pour effectuer cette sélection, on laisse à l'utilisateur le contrôle de la taille de la sphère. Par exemple, il lui suffit de fixer son rayon à zéro pour sonifier entièrement la protéine.

- Ensuite, on propose de jouer sur le plan temporel en sonifiant les points de l'intérieur vers l'extérieur. Ainsi, plus une source sonore est loin de la sphère de filtrage, plus elle est jouée tard. La sonification se propage donc telle une onde de choc, dans un mouvement centrifuge qui part du centre de la protéine vers l'extérieur (voir la figure 3.3).

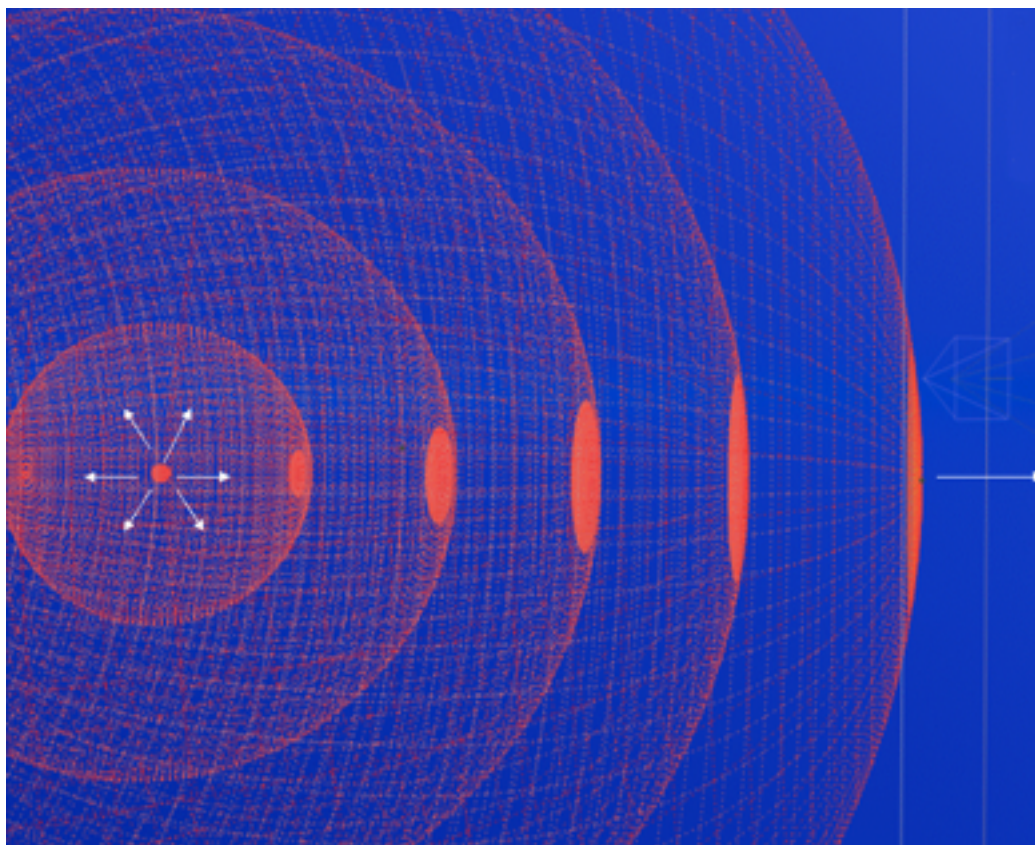


Fig. 3.3 : Illustration de la propagation de la sonification depuis la sphère de filtrage vers l'extérieur.

La vitesse idéale à laquelle les sources sonores doivent être lues reste à déterminer ; mais on peut encore imaginer laisser à l'utilisateur le loisir de décider de la valeur de ce paramètre. Ainsi, la protéine pourrait par exemple être sonifiée quasi-instantanément ou au contraire sur quelques secondes.

À noter qu'il serait aussi intéressant de lier le rayon de la sonde sphérique à la vitesse de lecture afin qu'elle suive visuellement la propagation de l'onde : la sphère grandirait donc à mesure que les points lus sont éloignés du centre, avant de revenir à sa taille initiale fixée par l'utilisateur dès que la lecture de l'objet est terminée. Cela permettrait de renforcer la localisation par l'utilisateur des points qui sont en train d'être sonifiés. Le principe de notre sonde sphérique serait alors similaire à celui de l'écholocation : au lieu d'un signal ultrasonore qui se réfléchit sur les obstacles, c'est l'agrandissement de la sphère qui provoque sa collision avec chacun des points composant l'ensemble de la surface de la protéine, et donc le déclenchement des sources sonores associées.

2. Stratégie de mapping des paramètres

Comme nous venons de le voir, notre proposition constitue une sonification par *mapping* de paramètres. Nous allons donc détailler brièvement dans la présente sous-section les différents types de *mapping* existants, avant de justifier dans la suivante les choix que nous avons effectués pour concevoir le notre.

3.2.1. Définition

On peut définir le *mapping* comme la mise en cohérence de minimum deux types d'informations distinctes. Dans le cadre spécifique de la sonification, le *mapping* consiste à mettre en rapport des données (ou certains attributs de données) avec des paramètres sonores (de synthèse le plus souvent). D'après [Rovan et al., 1997], il existe trois types de *mapping* différents :

- Le *mapping* unique (parfois appelé « un-à-un » dans la littérature concernée) : chaque type de donnée est associé à un seul et même paramètre sonore ; c'est le schéma de *mapping* le plus simple. Par exemple, si l'on choisit de sonifier un objet tridimensionnel, on peut faire varier la fréquence d'un flux sonore qui correspond à la taille de l'objet selon les correspondances audio-graphiques démontrées. Plus l'objet est gros, plus le son est grave ; et inversement.
- Le *mapping* divergent (parfois nommé « redondant » dans la littérature sur le sujet) : plusieurs paramètres sonores différents sont utilisés pour représenter une même donnée. Avec le même exemple, on ferait cette fois varier la fréquence et le niveau du flux sonore associé à la taille de

l'objet. Toujours selon les correspondances audio-graphiques, plus l'objet est gros, plus le son sera grave et fort ; et inversement : plus l'objet est petit, plus le son est aigu et faible.

- Le *mapping* convergent : un seul paramètre sonore est associé à plusieurs données différentes. Ce schéma de *mapping* est le plus complexe ; il nécessite notamment une expérimentation préalable de son utilisation pour arriver à des résultats pertinents. Toujours en poursuivant l'exemple précédant, on utiliserait pour ce type de *mapping* la variation du niveau à la fois pour signifier la taille d'un objet et sa proximité relative avec l'utilisateur. Ainsi, le flux sonore associé à un objet proche et gros serait plus fort que celui d'un objet loin et gros, lui même équivalent à un objet proche et petit et plus fort qu'un objet loin et petit.

Notre méthode de sonification repose donc sur un *mapping* divergent. En effet, pour augmenter la précision de notre localisation de la position de chacun des points, nous avons conçu notre sonification de sorte à ce que cette position soit mise en correspondance avec un maximum de paramètres sonores. Nous détaillons le *mapping* de ces paramètres dans la sous-section suivante.

De manière générale, le *mapping* a une importance déterminante sur l'expérience et les capacités de l'utilisateur. Par exemple, Hunt et al. [2002] ont montré qu'avec la même interface sonore mais des *mappings* différents, les sujets ont des sensations très différentes et obtiennent des résultats plus ou moins corrects. De plus, de par leur facilité, les *mappings* simples lassent les utilisateurs plus rapidement dans la longueur que les *mappings* complexes. Ces derniers nécessitent en effet une plus grande concentration auditive et visuelle ainsi qu'un certain temps d'adaptation. Il s'agira donc de trouver un *mapping* à cheval entre ces deux extrêmes, qui puisse à la fois rendre l'expérience immersive et stimulante sans nécessiter une période d'apprentissage trop longue et possiblement décourageante pour l'utilisateur. Tout l'enjeu de ce travail de sonification était aussi d'obtenir un rendu sonore qui soit le plus esthétique possible, afin de favoriser l'écoute, la concentration et la motivation du sujet lorsqu'il utilise l'interface mise à sa disposition.

[Yeo et Berger, 2005] ont théorisé dans leurs travaux deux concepts différents de *time mapping*, une variante de *mapping* qui concerne des données non-changeantes dans le temps⁵⁵ :

⁵⁵ C'est le cas de notre jeu de données puisque les formes des protéines sont fixées au moment de l'export en .obj via PyMOL.

- Le *scanning* (« balayage ») : C'est un *mapping* qui fige la sonification des données selon un ordre fixe et non-modifiable par l'utilisateur.
- Le *probing* (« sondage » ou « investigation ») : C'est une méthode de *mapping* temporel qui laisse à l'opérateur la possibilité de choisir quelle donnée sera sonifiée et à quel moment.

Notre proposition de *mapping* divergent rentre dans cette deuxième catégorie puisque le sujet discrimine avec la sonde sphérique les parties de la surface de la protéine qui vont être sonifiées ou non. De plus, en lui laissant la possibilité de modifier la durée du balayage sphérique et donc la vitesse de lecture, l'utilisateur peut manipuler le jeu de données au gré de ses envies. À noter qu'il aura également le choix entre actionner manuellement la sonification quand il le désire et un déclenchement périodique donc il pourra fixer la fréquence.

3.2.2. Échelle de *mapping*

Il s'agira dans cette sous-section d'étudier les différents critères relatifs à la construction d'un *mapping* efficient. Il convient ainsi d'interroger le type d'information que l'on va sonifier afin de déterminer l'état des variables exploitables pour un attribut donné. Cela revient par exemple à se demander s'il y a une valeur seuil à représenter. Il est également nécessaire de prendre en compte notre résolution auditive, afin de déterminer de combien doit être l'incrément de la valeur d'un attribut acoustique pour être perçu par notre oreille. De plus, dans une sonification par *mapping* de paramètres, une donnée variable est associée à chaque attribut sonore. Il convient donc d'effectuer un changement d'échelle pour faire correspondre l'intervalle de valeurs de la variable à sonifier avec l'intervalle de valeurs de l'attribut sonore associé.

Enfin, pour que l'utilisateur reste attentif et ne soit pas submergé par un nombre de variations de paramètres trop important, il est préférable de respecter la règle suivante : la fréquence de variation d'un attribut sonore, et la capacité de l'utilisateur à la percevoir, doit suivre la fréquence de rafraîchissement que peut suivre l'utilisateur (à bien différencier de la résolution temporelle de l'oreille). Ces limites de perception ne sont pas quantifiables *a priori*, il convient donc de les étudier en faisant par exemple varier les valeurs des paramètres lors d'expérimentations et de tests utilisateurs.

Nous avons expliqué précédemment que notre *mapping* était considéré comme unidimensionnel et divergent. Il ne s'attache en effet à sonifier qu'un seul type de données à l'aide de plusieurs paramètres sonores : la position d'un point par rapport au centre du repère où est placé le point d'écoute de l'utilisateur. Or, comment caractériser cette position ? Comme nous cherchons à maximiser notre perception de la localisation et donc de la distance de chaque point par rapport à l'origine du monde virtuel, le son devra être mappé sur les coordonnées du point qu'il sonifie. Nous proposons donc de travailler en particulier avec la distance d'un point (x, y, z) à l'origine $(0, 0, 0)$ d'un repère cartésien en trois dimensions. Cette distance r (qu'on appellera « rayon ») est donnée par la formule :

$$r = \sqrt{x^2 + y^2 + z^2}$$

De plus, notre monde virtuel est borné dans une *bounding box* cubique de côté 2. L'objet 3D est donc « normalisé⁵⁶ » pour que tous ses points rentrent dans la *world box*. Le point le plus éloigné possible du centre correspond alors à l'un des sommets du cube, de coordonnées $(\pm 1, \pm 1, \pm 1)$. De là, on en déduit que les points à sonifier auront donc tous un rayon compris dans l'intervalle $[0 ; \text{env. } 1,732]$.

Mis à part la spatialisation qui utilise elle les coordonnées cartésiennes (x, y, z) du point, tous les autres paramètres sonores du son associé au point seront mappés sur cette valeur r .

3.2.3. Mapping des paramètres sonores

1. Fréquence

À chaque point est associé un son pur succinct, joué simultanément à un son de bâton de pluie très large bande. La sinusoïde nous permet d'utiliser dans notre *mapping* la fréquence, un paramètre sonore particulièrement adéquat pour sonifier des variations fines de valeurs [Stockman

⁵⁶ De l'anglais *normalize*, cela veut dire que les coordonnées (x, y, z) de chaque point de l'objet sont « mises à l'échelle » entre $[-1 ; 1]$.

et al. 2005] car notre perception en fréquences est très précise [Wier, 1977]. Le son à large bande permet lui d'améliorer la localisation spatiale des différentes sources. Nous avons choisi un son de bâton de pluie pour tenir ce rôle car c'est un instrument percussif avec un spectre fréquentiel large, qui permet d'enrichir le contenu fréquentiel global de la source sonore et donc de « casser » l'aspect synthétique de la sinusoïde. De plus, le bâton de pluie a cette double connivence avec à la fois la musique et l'effet sonore ou le bruitage. On obtient donc au final un son complexe.

On décide donc que la fréquence du sinus est proportionnelle au rayon du point. Nous fixons ensuite comme intervalle de fréquences l'équivalent de deux octaves [130 Hz ; 1046 Hz]. Utiliser une bande passante aussi réduite permet tout d'abord de s'assurer que notre sonification soit similaire en rendu peu importe la qualité de restitution fréquentielle du casque utilisé. De plus, notre spatialisation est décodée en binaural, un système de localisation basé sur les différences interaurales de temps et d'intensité. Seulement, plusieurs chercheurs ont pointé du doigt le fait que notre perception de la latéralisation d'un signal pur par ITD et IID était seulement effective à des fréquences inférieures à environ 1500 Hz [Hershkowitz et Durlach, 1969 ; Domnitz et Colburn, 1977 ; Yost, 1981].

Ensuite, la polarité de notre *mapping* est fixée de sorte à ce que le rayon minimum corresponde à la fréquence la plus basse, et inversement. Ainsi, plus le point est loin du centre, plus le son qui le représente est aigu. En effet, même si notre expérience auditive nous apprend que ce sont d'abord les fréquences aiguës qui disparaissent avec la distance, des études sur les correspondances audio-graphiques montrent que l'on a plutôt tendance à associer un son aigu à un stimulus visuel de petite taille (qui est donc loin). Inversement, plus l'on fait grossir le stimulus visuel (et donc plus il se rapproche de nous), plus on lui associe un son grave [Marks et al., 1987 ; Kitagawa et Ichihara, 2002 ; Mondloch et Maurer, 2004 ; Gallace et Spence, 2006].

Enfin, comme notre perception de la hauteur n'est pas linéaire, on décide que le *mapping* des deux valeurs suit une loi logarithmique.

La variation de la fréquence sera notre indice principale pour la détection de la distance. Cependant, d'autres ajustements de *mapping* peuvent encore être faits ; notamment pour améliorer la perception de la position et de la distance des points par l'auditeur. Pour ce qui est de la distance, on pense naturellement à un effet d'atténuation en niveau ainsi qu'une réverbération.

2. Atténuation du niveau avec la distance

Sur le modèle de propagation d'une source ponctuelle⁵⁷ en champ libre, nous appliquons à notre son complexe une atténuation du niveau en fonction de la distance. Cette distance correspond donc ici au rayon du point sonifié. On fixe le niveau d'atténuation de sorte à ce que le point le plus éloigné du centre ait un gain de -15 dB. À noter que ce choix est empirique et ne repose que sur des essais. Il mérite donc d'être évalué par des tests approfondis.

Notre *mapping* devrait logiquement être logarithmique, pour respecter encore une fois le fonctionnement de notre oreille. Mais comme l'on ne se rend compte d'un changement franc de niveau sonore que si sa variation est supérieure à 3 dB, on obtient avec ce *mapping* 6 différences de niveaux perceptibles. On utilise donc cette fois-ci une loi de *mapping* par paliers : cela nous donne 5 sphères concentriques virtuelles, séparées d'une distance $d = 0,3$ et correspondant chacune à une atténuation de 3 dB (voir la figure 3.4).

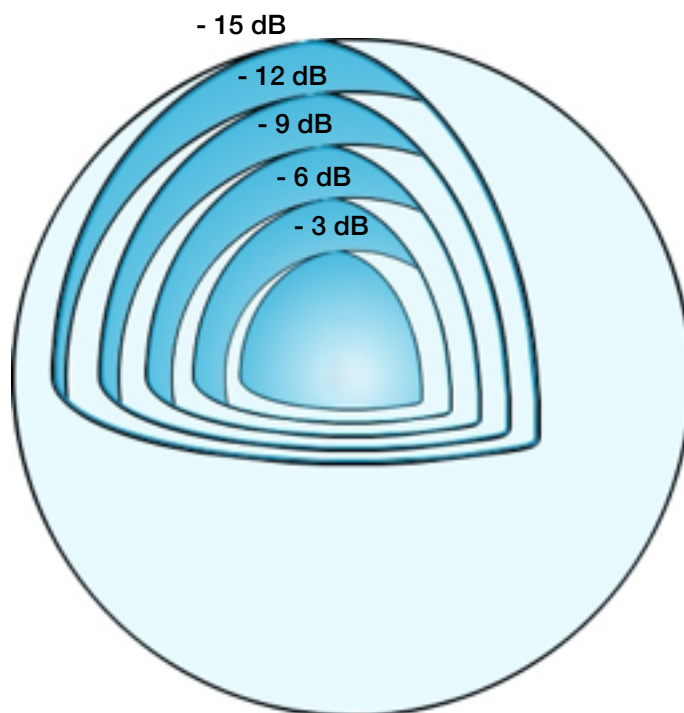


Fig. 3.4 : Illustration des paliers sphériques concentriques. Chaque couche vers l'extérieur correspond à 3 dB d'atténuation supplémentaires.

⁵⁷ Une source est dite « ponctuelle » si l'émission peut être considérée localisée en un seul point.

3. Réverbération

De nombreuses études ont mis en évidence l'amélioration de la perception de la distance avec des indices de réverbération [Mershon et King, 1975 ; Nielsen, 1992]. Begault [1992] a quant à lui montré les avantages liés à l'utilisation d'une réverbération artificielle dans un environnement audio virtuel : cet ajout conduit à une meilleure perception de la distance et augmente l'effet d'externalisation du son⁵⁸. Il a cependant mentionné l'effet néfaste que cela pouvait occasionner sur la précision de la latéralisation des sources, étant donné qu'une réverbération a tendance à augmenter la largeur apparente de la source.

Nous proposons donc d'utiliser un effet de *delay*, pour obtenir un écho à la manière d'un sonar. On décide de fixer le *delay time* (temps entre chaque répétition) à 500 ms ainsi que le réglage de *mix* (qui détermine le mélange entre signal « sec » et signal affecté par le *delay*, soit l'intensité de l'effet en quelque sorte) à 100 %. Comme pour les précédents paramètres, le radius du point sera mappé avec le réglage de *feedback* (qui détermine le nombre de répétition), sur un intervalle allant de 0 à 50 % (0 % correspondant au radius minimum et 50 % au radius maximum) et avec une évolution linéaire.

4. Spatialisation

On a donc à ce stade un son complexe qui, selon le radius du point qu'il sonifie, est plus ou moins aigu, plus ou moins fort et avec plus ou moins de *delay*. Il ne reste plus qu'à le spatialiser selon les coordonnées du point qu'il représente en Ambisonic virtuel décodé en binaural au casque.

Cependant, le binaural a de nombreux défauts, notamment concernant la localisation avant/arrière et en élévation. De plus, le fait d'utiliser des HRTF non individualisées pour la restitution n'aide pas. En effet, plusieurs études ont évoqué les possibles dégradations liées à leur utilisation, comme une moins bonne externalisation du son dans le plan médian et surtout une augmentation des confusions avant/arrière et haut/bas [Wenzel et al., 1993, Wightman et Kistler, 1993].

⁵⁸ Externaliser un son issu d'écouteurs ou d'un casque correspond à avoir l'impression que ce son provient malgré tout de l'extérieur de la tête.

Mais il est possible d'améliorer significativement ces deux types de localisation, notamment par des effets de filtrage et de niveau [Ortega-González et al., 2010a, Ortega-González et al., 2010b].

En suivant les résultats de ces travaux, on se propose donc d'ajouter un filtre passe-bas (LPF, pour *Low-Pass Filter*) ainsi qu'une légère atténuation en niveau (de l'ordre de -3 dB) pour les sources situées à l'arrière. Cela permet de réduire les confusions avant/arrière. On choisit également d'implémenter un filtre passe-haut (HPF, pour *High-Pass Filter*) associé à un léger gain (de l'ordre de +3 dB) pour les sources positives en élévation, afin de mieux les discriminer.

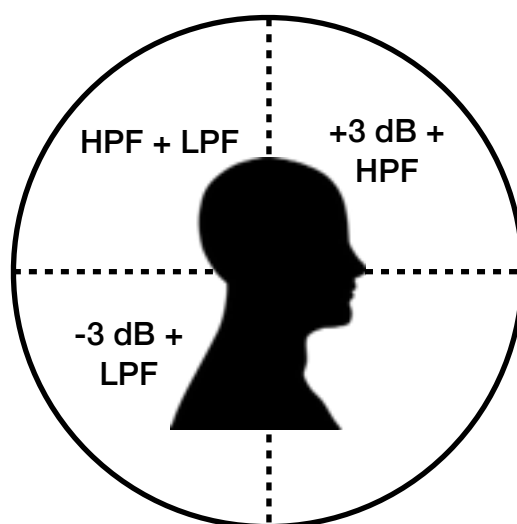


Fig. 3.5 : Illustration des filtres et des gains appliqués aux sources sonores selon leur localisation

3. Conclusions provisoires

La validation de notre *mapping* doit évidemment passer par des tests utilisateur avec des chimistes et des bio-informaticiens, afin que soient évalués le choix des paramètres sonores ainsi que leurs échelles de *mapping*.

4. Implémentation de notre proposition de sonification

Après avoir décrit les différentes ressources convoquées, nous détaillerons dans cette dernière partie quelques briques logicielles des patches de sonification et de visualisation composant notre interface utilisateur développée dans *Max 7*. Il ne s'agit pas d'être exhaustif ou d'expliquer en détail chaque élément des patches, mais plutôt de donner à voir le résultat concret du cheminement théorique qui nous a mené jusqu'à cette implémentation pratique.

4.1. Description des ressources utilisées

4.1.1. Le format OBJ

Le format OBJ est un format de fichier contenant la description d'une géométrie 3D. Défini par la société Wavefront Technologies, ce format de fichier est ouvert et a été adopté par la majorité des logiciels de 3D (Maya, Blender, MeshLab, 3D Studio Max). Il existe plusieurs façons de représenter et de décrire une forme dans un espace en trois dimensions. On a choisit d'utiliser dans notre sonification la représentation polygonale car c'est la plus fréquente et également la plus simple à manipuler.

1. La représentation polygonale

La représentation polygonale d'une surface est composée de trois types de données :

- Des polygones (*faces*), leur assemblage forme ce qu'on appelle un maillage (*mesh*). Ce sont le plus souvent des triangles.
- Des points (*vertices*, *vertex* au singulier), qui sont en fait les sommets des polygones.
- Des arêtes (*edges*), qui correspondent aux côtés des polygones.

On peut alors prendre l'exemple d'un cube (voir la figure 4.1) : il se décompose en 8 sommets, 12 polygones (2 par face du cube) et 18 arêtes.

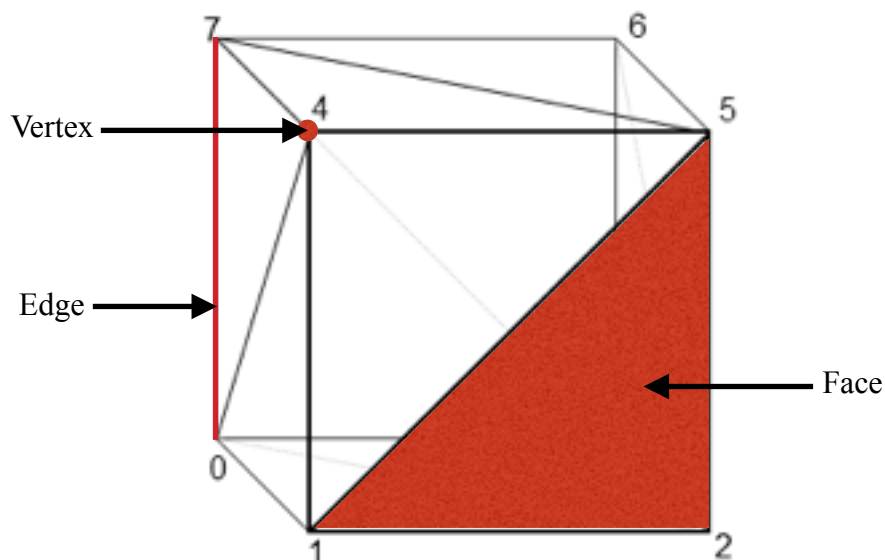


Fig. 4.1 : Représentation polygonale d'un cube

2. Le fichier .obj

Le format OBJ stocke ses informations dans deux fichiers différents :

- un fichier .obj, qui contient les sommets, les polygones et les normales⁵⁹ ;
- un fichier .mtl, qui contient toutes les informations relatives à la texture.

À noter que si le fichier .obj n'utilise pas les informations d'arêtes, c'est que les données « polygones » suffisent à elles seules pour décrire quels sont les trois points à relier ensemble pour former un triangle.

Nous ne nous intéressons qu'aux fichiers .obj pour notre sonification. On peut expliciter leur structure en reprenant l'exemple précédent du cube dans les figures 4.2 et 4.3.

⁵⁹ Les normales servent notamment à décrire la manière dont la texture et la lumière doivent s'appliquer sur la surface à un endroit donné.

4. Implémentation de notre proposition de sonification

Coordonnées des points en (x, y, z) . →
Ce sont ces informations que l'on utilise dans notre *mapping*.

```
# cube.obj
#
g cube
v 0.0 0.0 0.0
v 0.0 0.0 1.0
v 0.0 1.0 0.0
v 0.0 1.0 1.0
v 1.0 0.0 0.0
v 1.0 0.0 1.0
v 1.0 1.0 0.0
v 1.0 1.0 1.0
```

Coordonnées des normales →

```
vn 0.0 0.0 1.0
vn 0.0 0.0 -1.0
vn 0.0 1.0 0.0
vn 0.0 -1.0 0.0
vn 1.0 0.0 0.0
vn -1.0 0.0 0.0
```

Chaque face est définie par un ensemble d'indices faisant référence aux coordonnées des points et des normales définies précédemment. Par exemple, la première face est composée des points v_1 , v_7 et v_5 . Ils sont tous les trois associés à la normale vn_2 .

```
f 1//2 7//2 5//2
f 1//2 3//2 7//2
f 1//6 4//6 3//6
f 1//6 2//6 4//6
f 3//3 8//3 7//3
f 3//3 4//3 8//3
f 5//5 7//5 8//5
f 5//5 8//5 6//5
f 1//4 5//4 6//4
f 1//4 6//4 2//4
f 2//1 6//1 8//1
f 2//1 8//1 4//1
```

Fig. 4.2 : Fichier .obj d'un cube

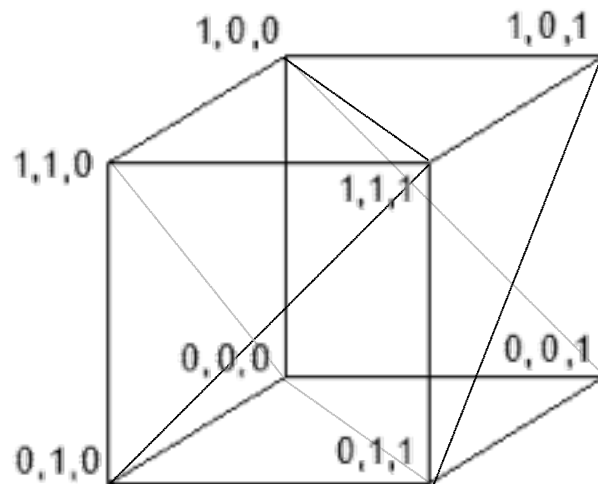


Fig. 4.3 : Représentation du cube du fichier .obj de la figure 4.2

3. Du .pdb au .obj

Pour obtenir un fichier .obj de la structure de la protéine que l'on souhaite sonifier, on télécharge tout d'abord le fichier .pdb qui lui est associé sur le site de la PDB. Ensuite, on lit ce fichier dans PyMOL. Nous avons alors la possibilité de l'exporter en .obj. On passe enfin par MeshLab pour appliquer un filtre de décimation de vertex et ainsi réduire la résolution polygonale de notre objet 3D (voir la figure 4.4). Cela nous permet de limiter le nombre de points à sonifier.

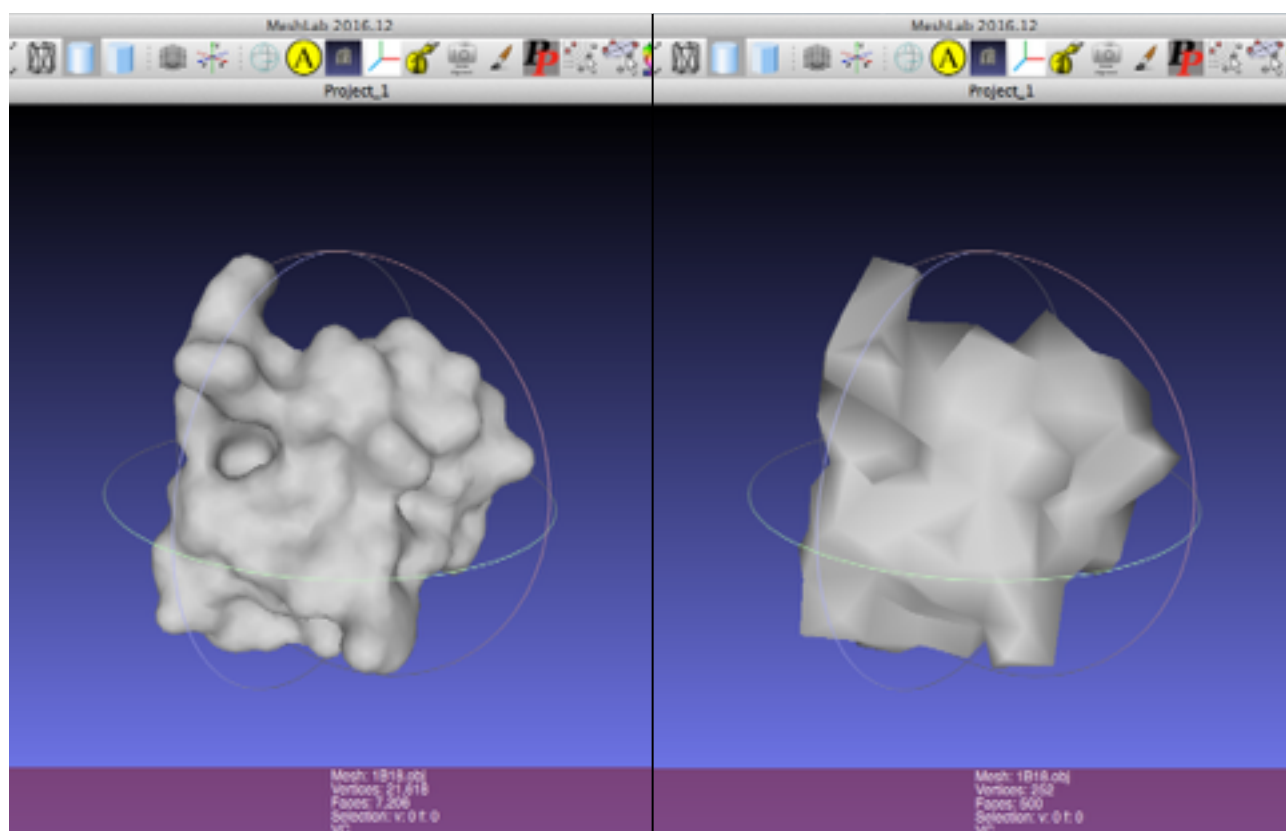


Fig. 4.4 : Réduction polygonale de la représentation 3D de la protéine 1B18. On passe de 21 618 points à gauche à seulement 252 points à sonifier à droite.

Nous avons ainsi à notre disposition dans *Max 7* une représentation simplifiée de l'objet que l'on souhaite sonifier, avec la possibilité de travailler sur les coordonnées de ses points. Attention, l'utilisation en tache de fond de la représentation simplifiée pour la sonification des points n'empêche pas l'utilisateur de charger à côté la représentation complète dans l'interface de visualisation. Le positionnement de la scène sonore obtenue est ainsi synchronisé avec le positionnement de la protéine représentée dans l'interface utilisateur.

4.1.2. *Max 7*

1. *Historique*

Max 7 (anciennement *Max/MSP*) est un logiciel de programmation qui permet notamment de faire de la synthèse sonore, de l'analyse et du traitement du signal, de l'enregistrement ainsi que du contrôle d'instruments MIDI. Développé à l'origine par l'IRCAM (Institut de Recherche et Coordination Acoustique/Musique) dans les années 1980, il a ensuite été racheté par Cycling '74 qui s'occupe désormais de son développement logiciel et de sa commercialisation.

Max est originellement le fruit de l'association de deux logiciels séparés développés par Miller Puckette⁶⁰ : *Max* et *MSP*. *Max* permet de faire des calculs mathématiques et, par extension, de contrôler en temps réel des instruments MIDI. *MSP* est une bibliothèque de fonctions qui, ajoutée à *Max*, permet de travailler en temps réel avec du signal audio.

L'environnement de *Max 7* propose aujourd'hui plusieurs briques logicielles supplémentaires, dont les plus connues sont *Jitter* et *Max for Live*. *Jitter* en particulier est une bibliothèque supplémentaire de fonctions ajoutée à *Max* (comme l'était *MSP*). Elle permet notamment de travailler sur des matrices. Son champ d'application est par conséquent multiple : traitement d'image en temps réel, mais aussi modélisation matricielle 3D en OpenGL. C'est cette dernière fonctionnalité qui nous a notamment intéressé dans *Max 7*. Il est en effet possible avec *Jitter* de lire et d'interagir avec des fichiers .obj dans *Max*.

2. *Fonctionnement*

Max est un logiciel de programmation dit « graphique » ; c'est-à-dire qu'il utilise un langage de programmation graphique dans lequel les programmes sont écrits en assemblant non pas des lignes de code mais des éléments graphiques présentés à l'utilisateur *via* son interface. Sa syntaxe concrète est composée de symboles graphiques (ici des boîtes fonctionnelles qu'on appelle « objets »), qui sont disposés spatialement dans ce qu'on appelle dans *Max* un « patch ».

Un projet *Max* se présente ainsi sous la forme d'au moins deux fenêtres :

⁶⁰ Miller Puckette a ensuite quitté l'IRCAM et créé *Pure Data*, l'équivalent *open source* et donc gratuit de *Max*.

4. Implémentation de notre proposition de sonification

- La première concerne donc le « patch ». Initialement vierge, c'est à l'intérieur de cette fenêtre que l'on va pouvoir concevoir et réaliser son projet en y déposant des objets virtuels. À noter qu'un même projet peut être constitué de plusieurs patches, puisque qu'il est possible à la fois d'implémenter des sous-patches à l'intérieur d'un patch principal et de faire communiquer des patches ensemble (on appellera alors les patches appelés des « abstractions »).
- La seconde est une console qui permet de contrôler les erreurs de programmation contenues dans le patch ; c'est finalement une boîte de dialogue entre le logiciel et l'utilisateur.

Une des raisons qui justifie notre utilisation de *Max 7* est la possibilité d'exporter son patch finalisé comme une application *standalone*⁶¹. Cette option est particulièrement utile puisqu'elle permet de partager son projet à d'autres utilisateurs, sans les contraindre à devoir également acheter et installer *Max*.

Le langage de programmation graphique de *Max* consiste donc à connecter entre elles des boîtes possédant des entrées et des sorties. Ainsi, ces boîtes s'envoient mutuellement des données, chaque boîte représentant un traitement particulier à effectuer sur ou avec ces données. Le principal type de boîtes est appelé « objet » : chaque objet joue un rôle précis et représente de fait une fonction codée en C que l'on va appeler et instancier par rapport à d'autres objets, des données, des messages ou des valeurs numériques.

3. Choix du moteur Ambisonic

Pour spatialiser nos sources, nous utilisons dans *Max* un moteur Ambisonic qu'il nous faut ensuite décoder en binaural. Il existe trois principaux moteurs Ambisonic disposant d'objets *Max* sur le marché :

- Sûrement le plus connu des trois, le *Spat Revolution* (anciennement *Spat*, pour « Spatialisateur ») est une suite logicielle développée par l'IRCAM et commercialisée par Flux.: [Carpentier, 2015a ; Carpentier et al., 2015b]. Nous ne l'avons pas pris en considération car il est payant. Or, si l'on souhaite que l'implémentation reste potentiellement modifiable par d'autres utilisateurs, il est nécessaire d'utiliser une librairie d'objets en libre accès.

⁶¹ Qui peut être utilisée seule, sans forcément posséder *Max 7*.

- L'*ICST Ambisonics* est une bibliothèque d'objets *Max* libre d'accès⁶² développée par l'*Institute for Computer Music and Sound Technology* de la Haute École d'art de Zurich) [Schacher, 2010].
- Enfin, la librairie *HOA (High Order Ambisonics)* est un ensemble de classes C++ et FAUST, disponibles également sous la forme d'objets *Max*, *Pure Data* et VST [Colafrancesco et al., 2103 ; Sèdes et al., 2014]. Elle est issue de deux projets du Labex Arts-H2H de l'université Paris 8. Également libre de droit⁶³, nous avons préféré utiliser cette bibliothèque d'objets car contrairement à l'*ICST Ambisonics*, elle propose nativement un décodeur binaural qui fonctionne directement avec son encodeur Ambisonic. Il y a de plus une vraie maintenance logicielle de la part de ses développeurs - de nombreuses mises à jour et *add-on* ont ainsi été proposés.

2. Patch de visualisation

Notre patch de visualisation de protéine est principalement basé sur des objets *Jitter* (différentes captures d'écran du patch sont disponibles en annexe A).

Il met en relation la représentation graphique de protéine que l'on importe (modélisée avec l'objet [jit.gl.model]) avec une sphère créée à l'intérieur de *Jitter* (grâce à l'objet [jit.gl.gridshape @shape sphere]). L'interaction manuelle avec la protéine *via* l'objet [jit.gl.handle] contrôle également la sphère de filtrage. Ainsi, si l'on déplace, change l'orientation ou zoome dans la molécule, la sphère de filtrage suit exactement les mêmes modifications. Cet objet permet également d'afficher les axes du repère sur lequel sont centrées la protéine et la sphère.

Différentes autres petites fonctionnalités sont ensuite implémentées (voir les figures 4.5 et 4.6). Par exemple, il est possible de changer le mode d'affichage de la protéine pour passer d'une représentation lisse à une représentation en fil de fer ou en nuage de points. On laisse également la possibilité à l'utilisateur de contrôler le niveau d'opacité de la sphère de filtrage.

⁶² Disponible en téléchargement à l'adresse suivante : <https://www.zhdk.ch/downloads-ambisonics-externals-for-maxmsp-5381>.

⁶³ Disponible en téléchargement à l'adresse suivante : <http://hoalibrary.mshparisnord.fr/telechargements/>.

4. Implémentation de notre proposition de sonification

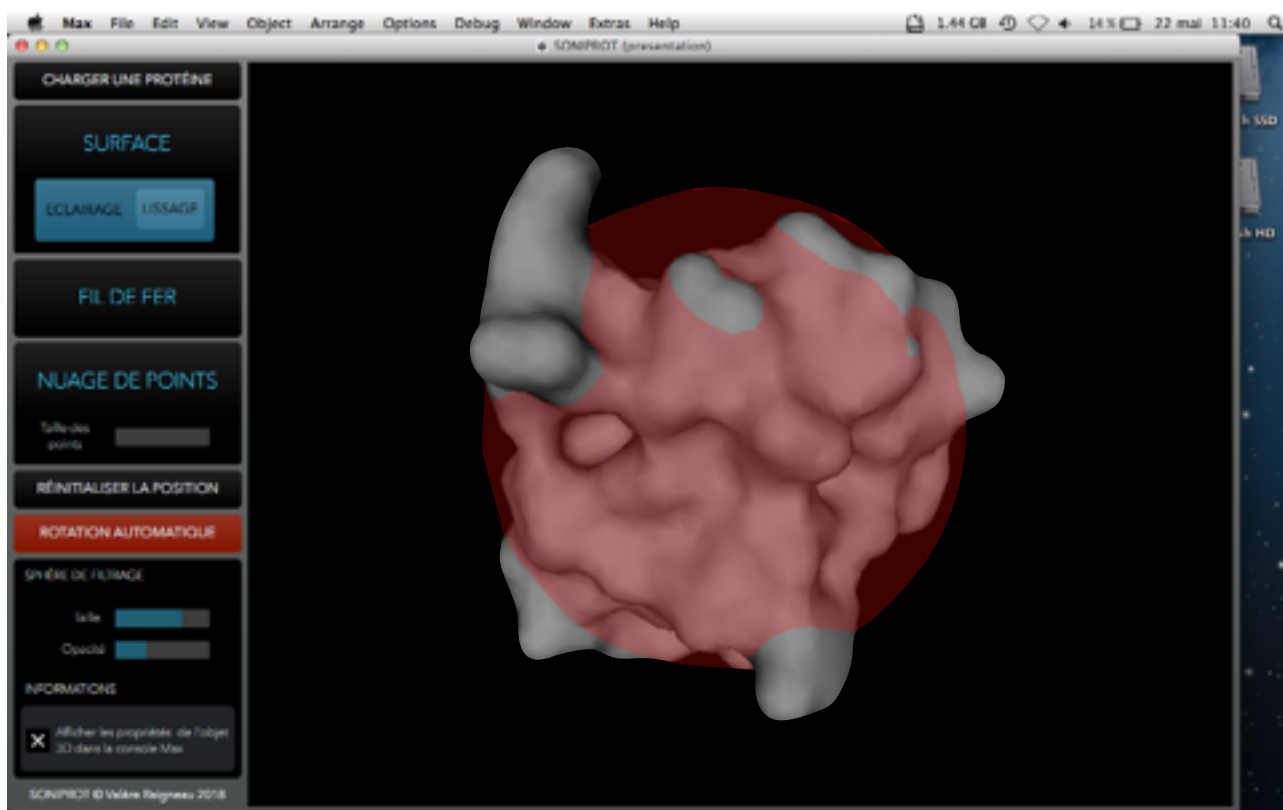


Fig. 4.5 : Vue de l'interface utilisateur du patch de visualisation avec la protéine 1B18. À noter que cette interface n'est pas définitive, car elle ne propose pas encore à l'utilisateur le contrôle des différents paramètres modifiables de la sonification.

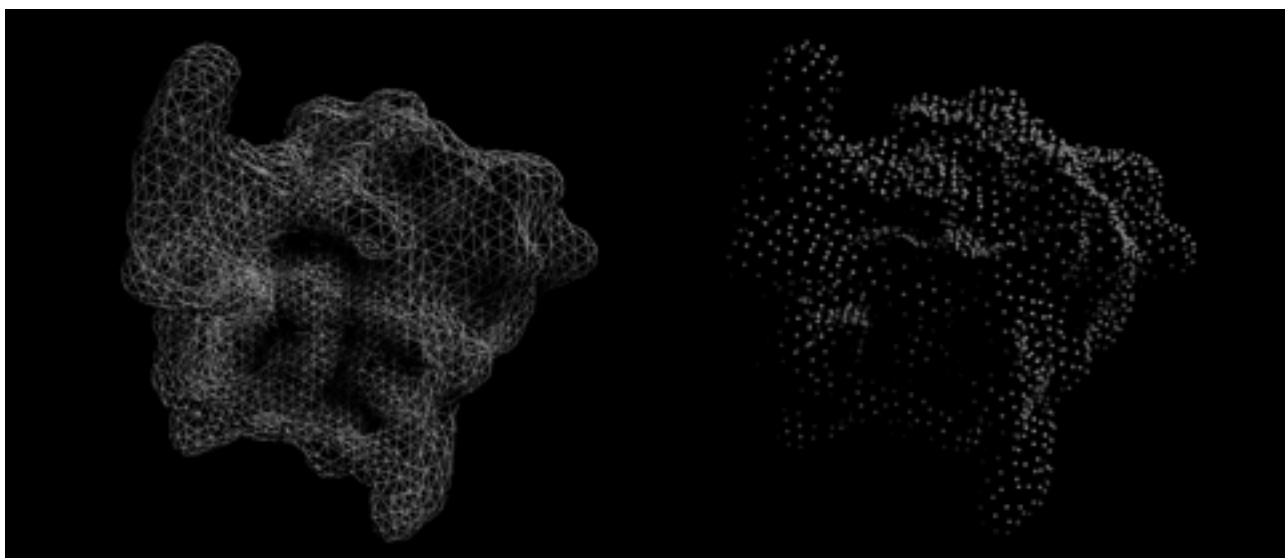


Fig. 4.6 : Représentation de la protéine 1B18 en fil de fer (à gauche) et en nuage de points (à droite) *via* le patch de visualisation

Comme expliqué précédemment, les objets importés sont remis à l'échelle de la *worldbox* grâce à l'attribut `@normalize`, qui permet de modifier toutes les coordonnées (x, y, z) d'un `.obj` pour les faire rentrer dans l'intervalle $[-1 ; 1]$.

3. Patch de sonification

Notre patch de sonification est composé d'un patch principal (figure 4.7) et d'un sous-patch (figure 4.8). Différentes captures d'écran du patch de sonification sont disponibles en annexe B.

1. Patch principal

Une fois que le fichier `.obj` de la figure simplifiée est chargé dans le patch principal, on lui fait subir une nouvelle décimation de points. Cette dernière n'a pas pour but de simplifier la forme de la protéine mais elle sert à s'assurer que toutes les coordonnées de points que nous allons utiliser sont uniques.

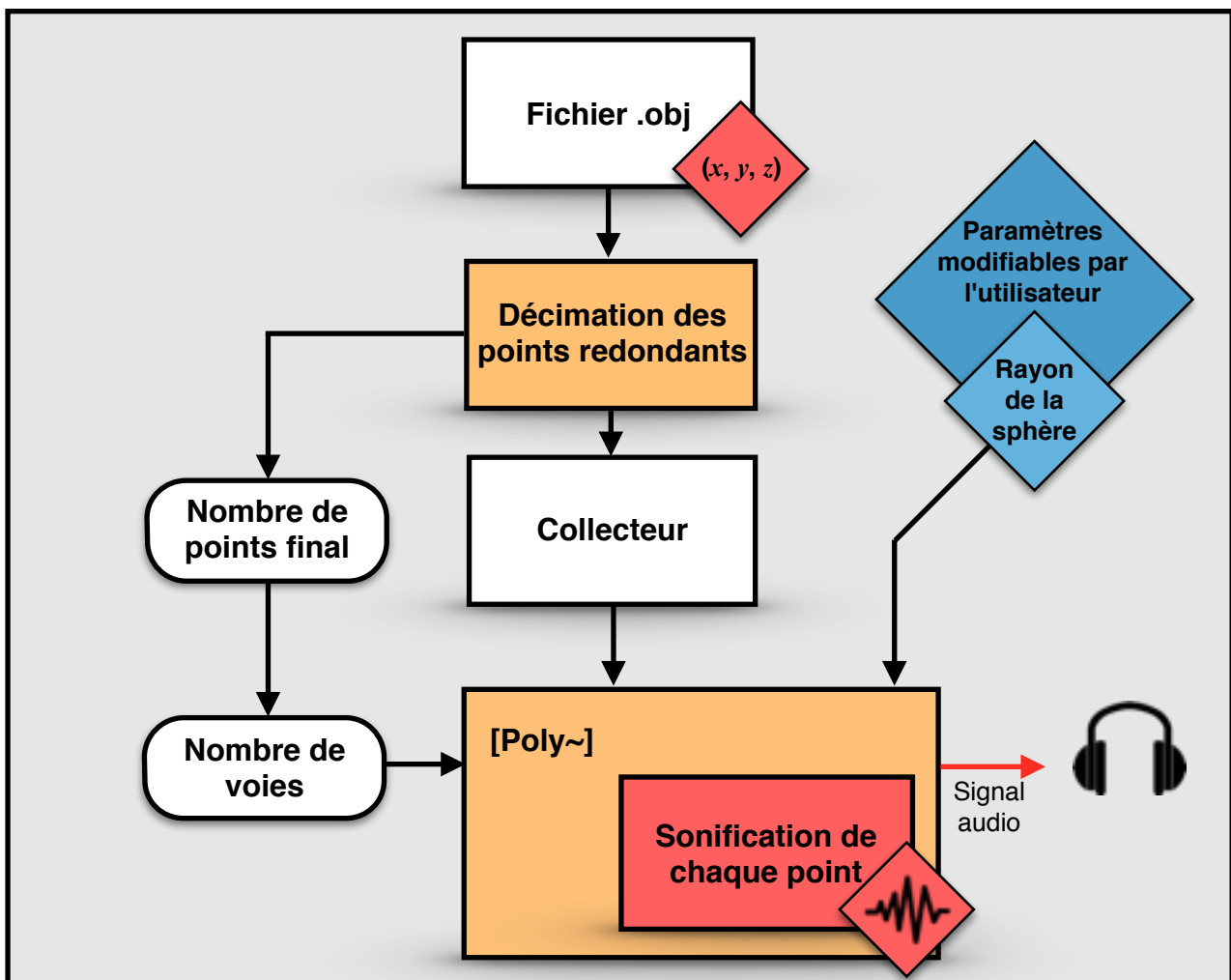


Fig. 4.7 : Schéma fonctionnel du patch principal de sonification

En effet, *Jitter* a tendance à dupliquer de nombreux points lorsqu'il rend un modèle 3D. Pour éviter ces doublons, nous utilisons donc un petit *external*⁶⁴ codé en Java par un membre du forum utilisateurs de Cycling '74 qui a déjà été confronté à ce problème⁶⁵. Cela nous permet donc de déterminer exactement le nombre de points que nous avons à sonifier.

Une fois notre jeu de coordonnées nettoyé des doublons, nous le stockons dans un collecteur qui gardera en mémoire les coordonnées de tous les points. Ce collecteur va servir à alimenter l'objet *Max* [Poly~]. Cet objet sert à faire de la polyphonie. En fixant le nombre de voies qu'il contient (autant de voies que de points donc), on fixe en fait le nombre d'instances de sous-patch que le [Poly~] va pouvoir contrôler. Il peut alors assigner des données différentes à chacune des voies qu'il contient.

Par exemple, si l'objet que j'ai chargé contient 14 points après décimation des points redondants, le [Poly~] va contenir 14 voies et va donc charger en fond 14 sous-patches. Chaque sous-patch va alors recevoir les coordonnées du point qui lui correspond. La voie 1 va donc obtenir les coordonnées du premier point, la voie 2 celle du deuxième, etc. C'est donc dans le sous-patch du [Poly~] que l'on va développer concrètement notre sonification.

2. Sous-patch

Le sous-patch du [Poly~] contient tous les éléments décrits dans la conception de notre *mapping* de sonification (cf. 3.2.3). On retrouve ainsi le synthétiseur de sinus, l'atténuateur en fonction de la distance, le *delay* ainsi que la ligne à retard.

On y reçoit les coordonnées du point associé à la voie. Après avoir calculé son rayon, on le compare avec celui de la sphère. S'il est inférieur ou égal, la voix est automatiquement coupée par le moteur Ambisonic. S'il est supérieur, alors le point est sonifié. À chaque fois que l'utilisateur change la taille de la sphère, la comparaison est de nouveau faite afin de déterminer le nouveau statut de la voie.

⁶⁴ Un *external* (« externe » en français) est un objet *Max* créé spécifiquement par un programme pour réaliser une nouvelle fonction.

⁶⁵ Pour plus de précisions sur la redondance de vertex dans *Jitter* et l'*external* Java en question, voir ce fil de discussion du forum Cycling '74 : <https://cycling74.com/forums/share-eliminate-redundant-replicated-vertices-with-jit-gl-model-gl-multiple>.

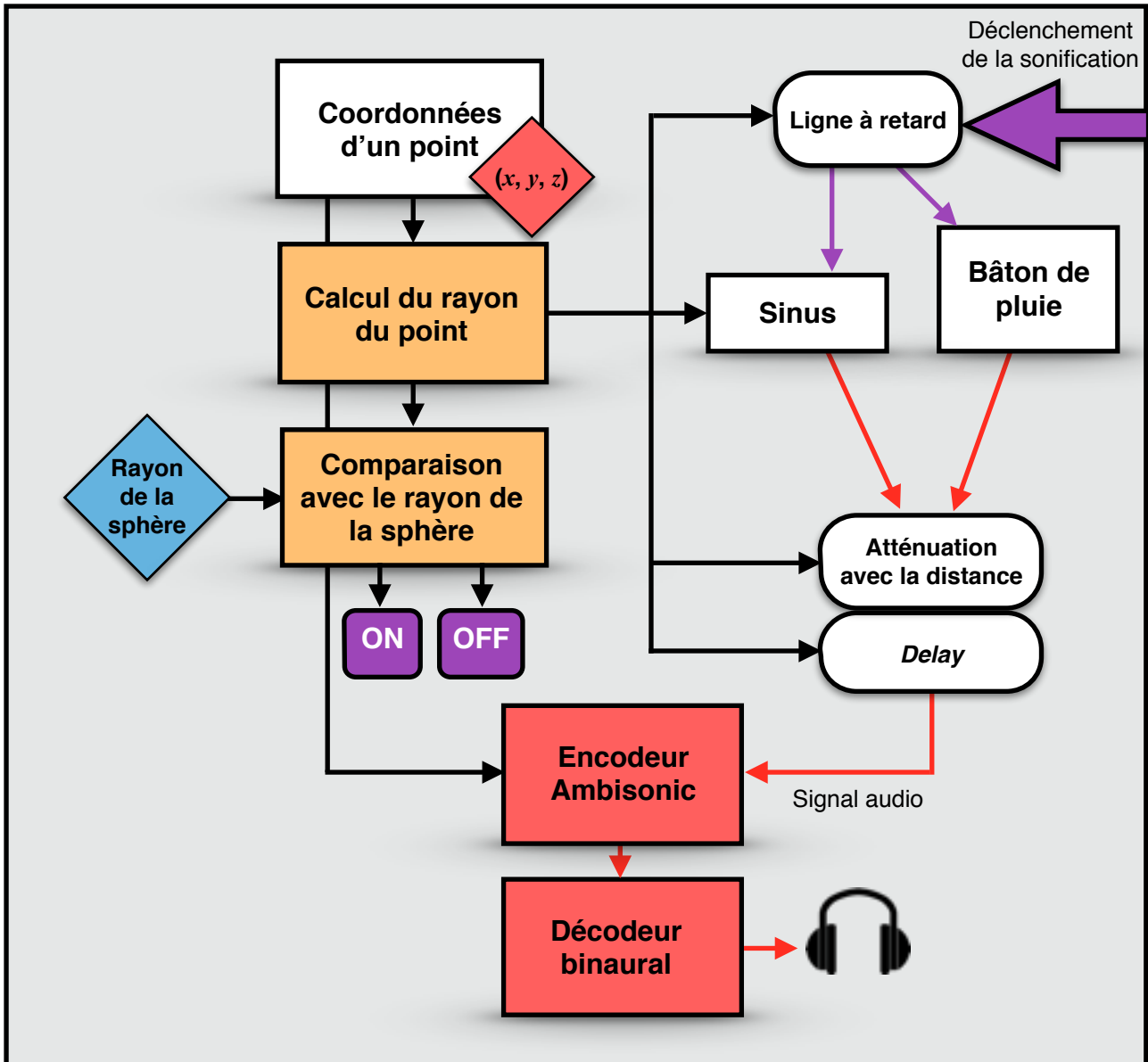


Fig. 4.8 : Schéma fonctionnel du sous-patch de sonification d'un point

Conclusion

À travers l'étude des différents paradigmes et enjeux de la sonification de données, nous avons pu mettre en exergue dans la première partie de ce mémoire les différents champs d'application qui peuvent bénéficier des atouts de la modalité auditive. En particulier, les représentations visuelles impliquent des problématiques et des contraintes auxquelles la sonification a su répondre de manière probante.

Il paraît donc censé d'attaquer le problème de la représentation de protéines par le prisme de la sonification de données. La seconde partie du mémoire nous d'ailleurs permis de constater que, même si le domaine de la sonification de molécules est diversifié autant dans ses approches que ses objectifs, il se restreint majoritairement à l'analyse de données séquentielles et paramétriques. Ainsi, si l'étude de ces contributions a permis d'apporter quelques réponses, elle a surtout soulevé un vide qui subsiste au niveau de la sonification de la forme même des protéines. Nous avons donc du passer par la sonification d'objets 3D pour apporter un éclairage nouveau et dégager quelques propositions intéressantes.

L'une d'elles en particulier, l'utilisation du son 3D pour la spatialisation de points, nous a mené jusqu'à la conception et l'implémentation de notre propre méthode de sonification. Elle met ainsi en jeu les points composant la forme de l'objet 3D, en utilisant leurs coordonnées pour générer des sources localisées dont les différents attributs sonores sont dépendants de sa distance au point d'écoute centré. En cherchant notamment à simplifier la représentation, nous avons pu également travailler autour de l'interactivité de l'utilisateur sur le *mapping*. On lui laisse pour cela la main sur de nombreux paramètres.

Cependant, le travail présenté dans ce mémoire ne fait qu'ouvrir une porte ; car si nous avons réussi à détailler une approche complète des questions relatives à la représentation de données de protéines, la réponse que nous avons formulée en retour n'a pas été mise à l'essai.

Ainsi, nombreuses sont les perspectives d'avenir. Il est tout d'abord primordial de récolter des retours d'expérience sur la proposition de *mapping* implémentée. Cette validation expérimentale peut prendre plusieurs formes : questionnaires, tests de reconnaissance en aveugle de formes 3D

simples comme un cube ou une pyramide, tests de classification de jeux de protéines, etc. Ces retours nous permettraient de remettre en perspective les différents choix effectués lors du *mapping*, pour ensuite dégager des améliorations possibles du modèle de sonification.

Certaines sont déjà apparentes : on peut par exemple imaginer utiliser le *tracking* de la tête de l'utilisateur pour améliorer les indices de localisation, voire même se servir d'un casque de réalité virtuelle. On pense également à l'implémentation d'une nappe sonore en fond, qui pourrait renseigner sur le nombre de points que contient la protéine chargée ou toute autre valeur. La prise en compte de nouveaux paramètres liés à la forme comme le volume, la courbure ou le rayon de courbure est aussi envisageable. On peut enfin penser à laisser la possibilité à l'utilisateur de modifier la position de son point d'écoute dans la scène sonore. Il pourrait ainsi se déplacer dans la molécule sans être centré sur la représentation.

Bibliographie

[Alexjander et Deamer, 1999] Alexjander, S., Deamer, D., *The Infrared Frequency of DNA bases: Science and Art*, IEEE Engineering in Medicine and Biology Magazine, vol. 18, no. 2, pp. 74-79, 1999.

[Alonso-Arevalo et al., 2008] Alonso-Arevalo, M. A., Shelley, S., Hermes, D., Kohlrausch, A., *Evaluating Geometrical Properties of Virtual Shapes Using Interactive Sonification*, presented at the IEEE International Workshop on Haptic Audio Visual Environments and their Applications (HAVE'08), Ottawa, Ontario, Canada, 2008.

[Alonso-Arevalo et al. 2012] Alonso-Arevalo, M. A., Shelley, S., Hermes, D., Hollowood, J., Pettitt, M., Sharples, S., Kohlrausch, A., *Curve Shape and Curvature Perception through Interactive Sonification*, ACM Transactions on Applied Perception, vol.9, no. 4, p. 19, 2012.

[André, 2000] André, E., « The Generation of Multimedia Presentations », in Dale, R., Moisl, H. Somers, H., *Handbook of Natural Language Processing: Techniques and Applications for the Processing of Language as Text*, Dekker, M., pp. 305-327, 2000.

[Baier et al., 2007] Baier, G., Hermann, T., Stephani, U., *Event-based sonification of eeg rhythms in real time*, Clinical Neurophysiology, 118(6), pp. 1377-1386, 2007.

[Ballweg et al., 2016] Ballweg, H., Bronowska, A. K., Vickers, P., *Interactive Sonification for Structural Biology and Structure-Based Drug Design*, Proceedings of ISON 2016, 5th Interactive Sonification Workshop, CITEC, Bielefeld University, Germany, December 16, 2016.

[Begault, 1992] Begault, D., *Perceptual Effects of Synthetic Reverberation on Three-Dimensional Audio Systems*, J. Audio Eng. Soc, 40(11), pp. 895-904, 1992.

[Begault, 1994] Begault, D., *3D Sound for Virtual Reality and Multimedia*, Cambridge: Academic Press, 1994.

[Blattner et al., 1989] Blattner, M., Sumikawa, D., Greenberg, R., *Earcons and Icons : Their Structure and Common Design Principles*, SIGCHI Bull., 21, pp. 123-124, 1989.

- [Blauert, 1997] Blauert, J., *Spatial Hearing: The Psychophysics of Human Sound Localization*, Cambridge, MA, USA: MIT Press, 1997.
- [Bregman, 1994] Bregman, A. S., *Auditory Scene Analysis: The Perceptual Organization of Sound*, MIT Press, 1994.
- [Brewster, 1997] Brewster, S., *Using Non-Speech Sound to Overcome Information Overload*, *Display*, 17, pp. 179-189, 1997.
- [Brewster et Murray, 2000] Brewster, S. A., Murray, R., *Presenting Dynamic Information on Mobile Computers*, *Personal Ubiquitous Comput.*, 4(4), pp. 209-212, 2000.
- [Bronkhorst, 2000] Bronkhorst, A. W., *The Cocktail Party Phenomenon: A Review of Research on Speech Intelligibility in Multiple-Talker Conditions*, *Acta Acustica united with Acustica*, vol. 86, no. 1, pp. 117–128, 2000.
- [Brown et al., 1989] Brown, M. L., Newsome, S. L., Glinert, E. P., *An experiment into the use of auditory cues to reduce visual workload*, *SIGCHI Bull.*, 20, pp. 339-346, 1989.
- [Brown et al., 2003] Brown, L., Brewster, S., Ramloll, S., Burton, R., Riedel, B., *Design Guidelines for Audio Presentation of Graphs and Tables*, *Proceedings of the 9th International Conference on Auditory Display*, 2003.
- [Buxton, 1989] Buxton, W., *Introduction to this Special Issue on Non-Speech Audio*, *Human-computer Interaction*, 4, pp. 1-9, 1989.
- [Bywater et Middleton, 2016] Bywater, R. P., Middleton, J. N., *Melody discrimination and protein fold classification*, *Heliyon*, 2016.
- [Carpentier, 2015a] Carpentier, T., *Récents développements du Spatialisateur*, *Journées d'Informatique Musicale*, Montréal, Canada, 2015.
- [Carpentier et al., 2015b] Carpentier, T., Noisternig, M., Warusfel, O., *Twenty Years of IRCAM Spat: Looking Back, Looking Forward*, *41st International Computer Music Conference (ICMC)*, Sep 2015, Denton, TX, United States, pp. 270-277, 2015.

- [Childs, 2002] Childs, E., *Achorripsis: A Sonification of Probability Distributions*, International Conference on Auditory Display, Kyoto, Japon, 2002.
- [Colafrancesco et al., 2103] Colafrancesco, J., Guillot, P., Paris, E., Sedes, A., Bonardi, A., *La bibliothèque HOA, bilan et perspectives*, JIM, Saint-Denis, France, 2013.
- [Davis et al., 2014] David, M., Lavandier, M., Grimault, N., *Room and head colouration can induce obligatory stream segregation*, J. Acoust. Soc. Am. 136 (1), pp. 5-8, 2014.
- [David et al., 2015] David, M., Lavandier, M., Grimault, N., *Sequential streaming, binaural cues and lateralization*, J. Acoust. Soc. Am. 138 (6), pp. 3500-3512, 2015.
- [David et al., 2017] David, M., Lavandier, M., Grimault, N., Oxenham, A., *Discrimination and streaming of speech sounds based on differences in interaural and spectral cues*, J. Acoust. Soc. Am. 142, pp. 1674-1685, 2017.
- [Delalande, 2009] Delalande, O., Ferey, N., Grasseau, G. & Baaden, M., *Complex molecular assemblies at hand via interactive molecular simulations*, Journal of Computational Chemistry 30(15), pp. 2375–2387, 2009.
- [Delatour, 2000] Delatour, T., *Molecular Music: The Acoustic Conversion of Molecular Vibrational Spectra*, Computer Music Journal, vol. 24, no. 3, pp. 48-68, 2000.
- [Dingler et al., 2008] Dingler, T., Lindsay, J., Walker, B., *Learnability of Sound Cues for Environmental Features: Auditory Icons, Earcons, Spearcons, and Speech*, Methods, pp. 1-6, 2008.
- [Dombois et Eckel, 2011] Dombois, F. et Eckel, G., « Audification », in Hermann, T., Hunt, A., Neuhoff, J. G., éditeurs, *The Sonification Handbook*, Logos Publishing House, Berlin, Germany, 2011.
- [Domnitz et Colburn, 1977] R. H. Domnitz and H. S. Colburn. *Lateral position and interaural discrimination*. Journal of the Acoustical Society of America, 61, pp. 1586-1598, 1977.

- [Dubus et Bresin, 2013] Dubus, G., Bresin, R., *Sonification of physical quantities throughout history: a meta-study of previous mapping strategies*, Proceedings of the 17th International Conference on Auditory Display (ICAD 2011), Budapest, Hungary, 2011.
- [Dunn et Clark, 1999] Dunn, J., Clark, M. A., *Life music: The sonification of proteins*, Leonardo, 32(1), pp. 25-32, 1999.
- [Edworthy et al., 1991] Edworthy, J., Loxley, S., Dennis, I., *Improving auditory warning design : Relationship between warning sound parameters and perceived urgency*, Human Factors, 33, pp. 205-231, 1991.
- [Edworthy, 1998] Edworthy, J., *Does Sound Help Us to Work Better with Machines?*, Interact. Comput., 10, pp. 401-409, 1998.
- [Fitch et Kramer, 1994] Fitch, W. T., Kramer, G., « Sonifying the Body Electric: Superiority of an Auditory over a Visual Display in a Complex, Multivariate System », in G. Kramer (Ed.), *Auditory Display: Sonification, Audification, and Auditory Interfaces*, Reading, MA: Addison-Wesley, pp. 307-326, 1994.
- [Flowers et Hauer, 1993] Flowers, J. H., Hauer, T. A., *"Sound" alternatives to visual graphics for exploratory data analysis*, Behavior Research Methods, Instruments & Computers, 25(2), pp. 242-249, 1993.
- [Frantti et Leverault, 1965] G. E. Frantti L. A. Leverault, *Auditory Discrimination of Seismic Signals from Earthquakes and Explosions*, Bull. Seism. Soc. Am., vol. 55, pp. 1-25, 1965.
- [Friberg et Gärdenfors, 2004] Friberg, J. et Gärdenfors, D., *Audio Games: New Perspectives on Game Audio*, Proceedings of the 2004 ACM SIGCHI International Conference on Advances in computer entertainment technology, pp. 148-154, 2004.
- [Gallace et Spence, 2006] Gallace, A., Spence, C., *Multisensory synesthetic interactions in the speeded classification of visual size*, Attention, Perception, & Psychophysics: 68., (7), pp. 1191-1203, 2006.

- [Gaudy et al., 2006] Gaudy, T., Natkin, S., Archambault, D., *Playing Audio Games without Instructions for Uses : To do without instruction leaflet or without language itself?*, CGAMES'06, Int. Conf. on Computer Games, Dublin, Ireland, pp. 263-268, 2006.
- [Gaver, 1986] Gaver, W., *Auditory Icons: Using Sound in Computer Interfaces*, Hum. Comput. Interact., 2, pp. 167-177, 1986.
- [Gerzon, 1980] Gerzon, M. A., *Practical Periphony, the Reproduction of Full Sphere Sound*. Preprint of the 65th Audio Engineering Society Convention, London, 1980.
- [Godbout et Boyd, 2010] Godbout, A. et Boyd, J. E., *Corrective sonic feedback for speed skating : A case study*, In Proceedings of the 16th International Conference on Auditory Display, 2010.
- [Gomez et Ramirez, 2011] Gomez, I., Ramirez, R., *A data sonification approach to cognitive state identification*, in Proceedings of the 17th International Conference on Auditory Display (ICAD2011), Budapest, Hungary, 2011.
- [Gond et Dall'Antonia, 2008] Grond, F., Dall'Antonia, F., *SUMO - A sonification utility for molecules*, in Proceedings of the 14th International Conference on Auditory Display (ICAD 2008), Paris, France, 2008.
- [Grond et al., 2010] Grond, F., Janssen, S., Schirmer, S., Hermann, T., *Browsing RNA Structures by Interactive Sonification*, in Proceedings of the 3rd Interactive Sonification Workshop, KTH, Stockholm, Sweden, pp. 11-16, 2010.
- [Grond et Berger, 2011] Grond, F., Berger, J. « Parameter Mapping Sonification », in Hermann, T. Hunt, A., Neuhoff J. G. (éditeurs), *The Sonification Handbook*, Logos Publishing House, Berlin, Germany, pp. 363-397, 2011.
- [Haas et Edworthy, 2006] Haas, E., Edworthy, J., « An introduction to Auditory Warnings and Alarms », in Wogalter, M. S., éditeur, *Handbook of Warnings*. Mahwah, NJ: Lawrence Erlbaum, 2006.

- [Hansen et al., 1999] Hansen, M. D., Charp, E., Lodha, S., Meads, D., Pang, A., *PROMUSE: A System for multi-media data presentation of protein structural alignments*, in Proceedings of the Pacific Symposium on Biocomputing, 1999.
- [Hayashi et Munakata, 1984] Hayashi, K., Munakata, N., *Basically musical*, Nature 310, p. 96, 1984.
- [Hayward, 1994] Hayward, C., « Listening to the Earth Sing », in Kramer, G., *Auditory Display: Sonification, Audification, and Auditory Interfaces*, pp. 369-404, 1994.
- [Hermann et Ritter, 1999] Hermann, T., Ritter, H., *Listen to your Data: Model-Based Sonification for Data Analysis*, Advances in intelligent computing and multimedia systems, vol. 8, pp. 189-194, 1999.
- [Hermann, 2008] Hermann, T., *Taxonomy and definitions for sonification and auditory display*, Proceedings of the 14th International Conference on Auditory Display, Paris, France, 2008.
- [Hermann, 2011] Hermann, T., « Model-based sonification », in Hermann, T., Hunt, A. et Neuhoff, J. G., éditeurs, *The Sonification Handbook*, Logos Publishing House, Berlin, Germany, 2011.
- [Hermann et al., 2002] Hermann, T., Meinicke, P., Bekel, H., Ritter, H., Müller, H. M., Weiss, S., *Sonification for eeg data analysis*, Proceedings of the 8th International Conference on Auditory Display (ICAD2002), 2002.
- [Hermann et al., 2011] T. Hermann, A. Hunt, J. G. Neuhoff (éditeurs), *The Sonification Handbook*, Logos Publishing House, Berlin, Germany, 2011.
- [Hershkowitz et Durlach, 1969] Hershkowitz, R. M., Durlach, N. I., *Interaural Time and Amplitude jnds for a 500-Hz Tone*, Journal of the Acoustical Society of America, 46, pp. 1464-1467, 1969.
- [Hofstadter, 1980] Hofstadter, D. R., Gödel, Escher, Bach: An Eternal Golden Braid, Vintage Books, New York, 1980.

- [Hunt et al., 2002] Hunt, A., Wanderley, M. M., Paradis, M., *The Importance of Parameter Mapping in Electronic Instrument Design*, Proceedings of the Conference on New Instruments for Musical Expression, 2002.
- [Jørgensen, 2006] Jørgensen, K., *On the Functional Aspects of Computer Game Audio*, Audio Mostly 2006: A Conf. on Sound in Games, Pitëa, Sweden, pp. 48-52, 2006.
- [Kitagawa et Ichihara, 2002] Kitagawa, N., Ichihara, S., *Hearing visual motion in depth*, Nature, 416, pp. 172-174, 2002.
- [Keating, 1996] Keating, D. A., *The Generation of Virtual Acoustic Environments for Blind People*, Proc. 1st European Conference on Disability, Virtual Reality and Associated Technologies, Maidenhead, UK, pp. 201-207, 1996.
- [King et Angus, 1996] King, R. D., Angus, C. G., *PM - Protein Music*, Computer Applications in Biosciences 12 (3), pp. 251-252, 1996.
- [Kramer, 1994] Kramer, G. (éditeur), *Auditory Display: Sonification, Audification, and Auditory Interfaces*, Reading, MA: Addison-Wesley, 1994.
- [Kramer et al., 1999] K. Kramer, B. Walker, T. Bonebright, P. Cook, J. Flowers, N. Miner, J. Neuhoff, *Sonification Report: Status of the Field and Research Agenda*, Faculty Publications, Department of Psychology, 1999.
- [Marks et al., 1987] Marks, L. E., Hammeal, R. J., & Bornstein, M. H., *Perceiving similarity and comprehending metaphor*, Monographs of the Society for Research in Child Development, 52, (1), pp. 1-102, 1987.
- [McAdams et Bigand, 1993] McAdams, S. et Bigand, E., *Thinking in Sound : The Cognitive Psychology of Human Audition*, Clarendon Press/Oxford University Press, New York, NY, USA, 1993.
- [McGookin et Brewster, 2011] McGookin, D., Brewster, S., « Earcons », in Hermann, T., Hunt, A., Neuhoff, J. G. (éditeurs), *The Sonification Handbook*, Logos Publishing House, Berlin, Germany, pp. 339-361, 2011.

- [Mershon et King, 1975] Mershon, D., King, L., *Intensity and reverberation as factors in the auditory perception of egocentric distance*, Attention, Perception, & Psychophysics, 18(6), pp. 409-415, 1975.
- [Minghim et Forrest, 1995] Minghim, R., Forrest, A. R., *An Illustrated Analysis of Sonification for Scientific Visualisation*, Proceedings of the 6th IEEE Visualization Conference, 1995.
- [Mondloch et Maurer, 2004] Mondloch, C., Maurer, D., *Do small white balls squeak? Pitch-object correspondences in young children*, Cognitive, Affective, & Behavioral Neuroscience, 4, (2), pp. 133-136, 2004.
- [Moore, 1997] Moore, B. C., *An Introduction to the Psychology of Hearing* (4th ed.). Academic Press, San Diego, Calif, 1997.
- [Nicol, 2010] Nicol, R., *Binaural Technology*, AES monograph, 2010.
- [Nielsen, 1992] Nielsen, S., *Auditory distance perception in different rooms*, in Audio Engineering Society Convention 92, 1992
- [Noisternig et al., 2003] Noisternig, M., Musil, T., Sontacchi, A. et Holdrich, R., *3D Binaural Sound Reproduction Using a Virtual Ambisonic Approach*. In International Symposium on Virtual Environments, Human Computer Interfaces, and Measurement Systems, 2003.
- [O'Donoghue et al., 2010] O'Donoghue, S. I., Goodsell, D. S., Frangakis, A. S., Jossinet, F., Laskowski, R. A., Nilges, M., Saibil, H. R., Schafferhans, A., Wade, R. C., Westhof, E., Olson, A. J., *Visualization of macromolecular structures*, Nature Methods, 7(3 Suppl), pp.42–55, 2010.
- [Ohno 1986] Ohno, S., Ohno, M., *The all pervasive principle of repetitious recurrence governs not only coding sequence construction but also human endeavor in musical composition*, Immunogenetics 24, pp. 71-78, 1986.
- [Ohno 1988] Ohno, S., « Of words, genes, and music », In E. Sercarz, editor, *The Semiotics of Cellular Communication in the Immune System*, NATO ISI Series, volume H32, Springer, pp. 131-147, 1988.

[Parseihian et Katz 2012] Parseihian, G., Katz, B. F. G., *Rapid Head-Related Transfer Function Adaptation Using a Virtual Auditory Environment*. Journal of the Acoustical Society of America, Express Letters, 2012.

[Parseihian et Katz, 2012] Parseihian, G., Katz, B. F. G., *Morphocons: A New Sonification Concept Based on Morphological Earcons*, Journal of the Audio Engineering Society, vol. 60, no. 6, pp. 409-418, 2012.

[Pauletto et Hunt, 2009] Pauletto, S., Hunt, A., *Interactive sonification of complex data*, International Journal of Human-Computer Studies, 67(11), pp. 923-933, 2009.

[Picinali, 2012] Picinali, L., Chrysostomou, C., Seker, H., *The sound of proteins*, in Biomedical and Health Informatics (BHI), 2012 IEEE-EMBS International Conference on, pp. 612-615, 2012.

[Pickover, 1992] Pickover, C. A., *There is music in our genes*, in *Mazes for the Mind: Computers and the Unexpected*, St. Martin's Press, 1992.

[Rau et al., 2015] Rau, B., Frieb, F., Krone, M., Muller, C., Ertl, T., *Enhancing visualization of molecular simulations using sonification*, in *Virtual and Augmented Reality for Molecular Science*, 2015 IEEE 1st International Workshop, pp. 25-30, 2015.

[Röber et al., 2005] Röber, N., Deutschmann, E. C. et Masuch, M., *Authoring of 3D virtual auditory environments*, in *Proceedings of the Audio Mostly Conference - a Conference on Sound in Games*, 2006.

[Rossiter et Ng, 1996] Rossiter, D., Ng. W. Y., *A system for the complementary visualization of 3D volume images using 2D and 3D binaurally processed sonification representations*, in *Proceedings of the 7th conference on Visualization*, IEEE Computer Society Press, San Francisco, USA, pp. 351-354, 1996.

[Rovan et al., 1997] Rován, J., Wanderley, M. M., Dubnov, S., Depalle, P., *Instrumental Gestural Mapping Strategies as Expressivity Determinants in Computer Music Performance*, *Proceedings of the Kansei Workshop*, 1997.

- [Röber et al., 2006] Röber, N., Deutschmann, E. C. et Masuch, M., *Authoring of 3d virtual auditory environments*. In Proceedings of the Audio Mostly Conference - A Conference on Sound in Games, 2006.
- [Sanders et McCormick, 1993] M. S. Sanders, E. J. McCormick, *Human Factors in Engineering and Design* (7th ed.). New York, NY, England: Mcgraw-Hill Book Company, 1993.
- [Sansom, 2002] Sansom, C., *DNA makes protein —makes music?*, The Biochemist, p. 40, 2002.
- [Schacher, 2010] J. C. Schacher, *Seven Years of ICST Ambisonics Tools for Max/MSP - A Brief report*, Proc. of the 2nd International Symposium on Ambisonics and Spherical Acoustics, Paris, France, 2010.
- [Schaffert et al., 2009] Schaffert, N., Mattes, K., Barrass, S. et Effenberg, A., *Exploring function and aesthetics in sonifications for elite sports*, Proceedings of the Second International Conference on Music Communication Science, 2009.
- [Sèdes et al., 2014] A. Sèdes, P. Guillot et E. Paris, *The HOA library, review and prospect*, ICMC-SMC, Athens, Greece, 2014.
- [Smith, 1992] Smith, J. O., *Physical modeling using digital waveguides*, Computer Music Journal, 16(4), pp. 74-91, 1992.
- [Sorkin, 1987] Sorkin, R. D., « Design of Auditory and Tactile Display », in Salvendy, G., éditeur : *Handbook of Human Factors*, New York: Wiley & Sons, pp. 549–576, 1987.
- [Speeth, 1961] S. D. Speeth, *Seismometer Sounds*, J. Acoust. Soc. Am., vol. 33, pp. 909-916, 1961.
- [Spence et Driver, 1997] Spence, C., Driver, J., « Audiovisual Links in Attention : Implications for Interface Design », in Harris, D., éditeur, *Engineering Psychology and Cognitive Ergonomics Vol. 2 : Job Design and Product Design*. Hampshire : Ashgate Publishing, 1997.
- [Straebel et Thoben, 2014] Straebel, V., Thoben, W., *Alvin Lucier's Music for Solo Performer: Experimental music beyond sonification*, Organised Sound, vol. 19, no. 1, pp. 17-29, 2014.

- [Stockman et al., 2005] Stockman, T., Valgerour Nickerson, L., Hind, G., *Auditory Graphs: A Summary of Current Experience and Towards a Research Agenda*, Proceedings of the International Conference on Auditory Display (ICAD2005), 2005.
- [Stockman et al., 2007] Stockman, T., Rajgor, N., Metatla, O., Harrar, L., *The design of interactive audio soccer*, Proceedings of the 13th International Conference on Auditory Display, 2007.
- [Stockmann et al., 2008] Stockman, T., Berndt, A., Röber, N., *A Musical Instrument based on 3D Data and Volume Sonification Techniques*, Proc. of the Audio Mostly Conference, pp. 72-79, 2008.
- [Strybel et al., 1992] Strybel, T., Manligas, C., Perrott, D., *Minimum audible movement angle as a function of the azimuth and elevation of the source*, Human Factors, 34(3), pp. 267-275, 1992.
- [Stryer, 1988] Stryer, L., *Biochemistry, Third edition*, W. H. Freeman and Company, p. 270, 1988.
- [Suied et al., 2008] Suied, C., Susini, P., McAdams, S., *Evaluating Warning Sound Urgency with Reaction Times*, J Experimental Psychology : Applied, 14(3), pp. 201-212, 2008.
- [Thoret et al., 2012] Thoret, E., Aramaki, M., Kronland-Martinet, R., Velay, J. L. et Ystad, S., *From shape to sound : sonification of two dimensional curves by reenaction of biological movements*, 9th International Symposium on Computer Music Modeling and Retrieval, London, 2012.
- [Verron et al., 2010] Verron, C., Aramaki, M., Kronland-Martinet, R. et Pallone, G., *A 3D Immersive Synthesizer for Environmental Sounds*. IEEE Transactions on Audio, Speech, and Language Processing, 18(6), pp. 1550-1561, 2010.
- [Walker et Kramer, 1996] Walker, B. N., Kramer, G., *Mappings and metaphors in auditory displays : An experimental assessment*, Frysinger, S. P., Kramer, G., éditeurs, Proceedings of the 3rd International Conference on Auditory Display (ICAD96), 1996.
- [Walker et Kramer, 2004] Walker, B., Kramer, G., « Ecological Psychoacoustics and Auditory Displays: Hearing, Grouping, and Meaning Making », in Neuhoff, J., éditeur, *Ecological Psychoacoustics*, New York: Academic Press, pp. 150-175 2004

- [Walker et al., 2006] Walker, B. N., Nance, A., Lindsay, J., *Spearcons: Speech-Based Earcons Improve Navigation Performance in Auditory Menus*, Proceedings of the International Conference on Auditory Display (ICAD 2006), London, pp. 63–68, 2006.
- [Walker et Nees, 2011] Walker, B. N., Nees, M. A., « Theory of Sonification ». in Hermann, T., Hunt, A. et Neuhoff, J. G., éditeurs, *The Sonification Handbook*, chapitre 2, pp. 9–39. Logos Publishing House,, Berlin, Germany, 2011.
- [Wenzel, 1992] Wenzel, E. M., « Three-Dimensional Virtual Acoustic Displays », in Blattner, M. M., Dannenberg, R. B. (éditeurs), *Multimedia Interface Design*, Addison-Wesley, 1992.
- [Wenzel et al., 1993] Wenzel, E. M., Arruda, M., Kistler, D. J. et Wightman, F. L., *Localization using non-individualized head-related transfer functions*, J. Acoust. Soc. Am., 94(1), pp. 111-123, 1993.
- [Wenzel, 1994] Wenzel, E. M., « Spatial sounds and sonification », in Gregory Kramer, (ed.), *Auditory Display: Sonification, Audification, and Auditory Interfaces*, Reading, MA: Addison-Wesley, pp. 127-150, 1994.
- [Wier et al., 1977] Wier, C. C., Jesteadt, W., Green, D. M., *Frequency Discrimination as a Function of Frequency and Sensation Level*, Journal of the Acoustical Society of America, vol. 61, pp. 178-183, 1977.
- [Wightman et Kistler, 1993] Wightman, F. L., Kistler, D. J., *Multidimensional scaling analysis of head-related transfer functions*, in IEEE Workshop on Applications of Signal Processing to Audio and Acoustics, 1993.
- [Yeo et Berger, 2005] Yeo, W., Berger, J., *A framework for designing image sonification methods*, Proceedings of the 11th International Conference on Auditory Display (ICAD 2005), pp. 323-327, 2005.
- [Yost, 1981] Yost, W. A., *Lateral position of sinusoids presented with intensive and temporal differences*, Journal of the Acoustical Society of America, 70, pp. 397-409, 1981.

Annexes

A. Patch de visualisation

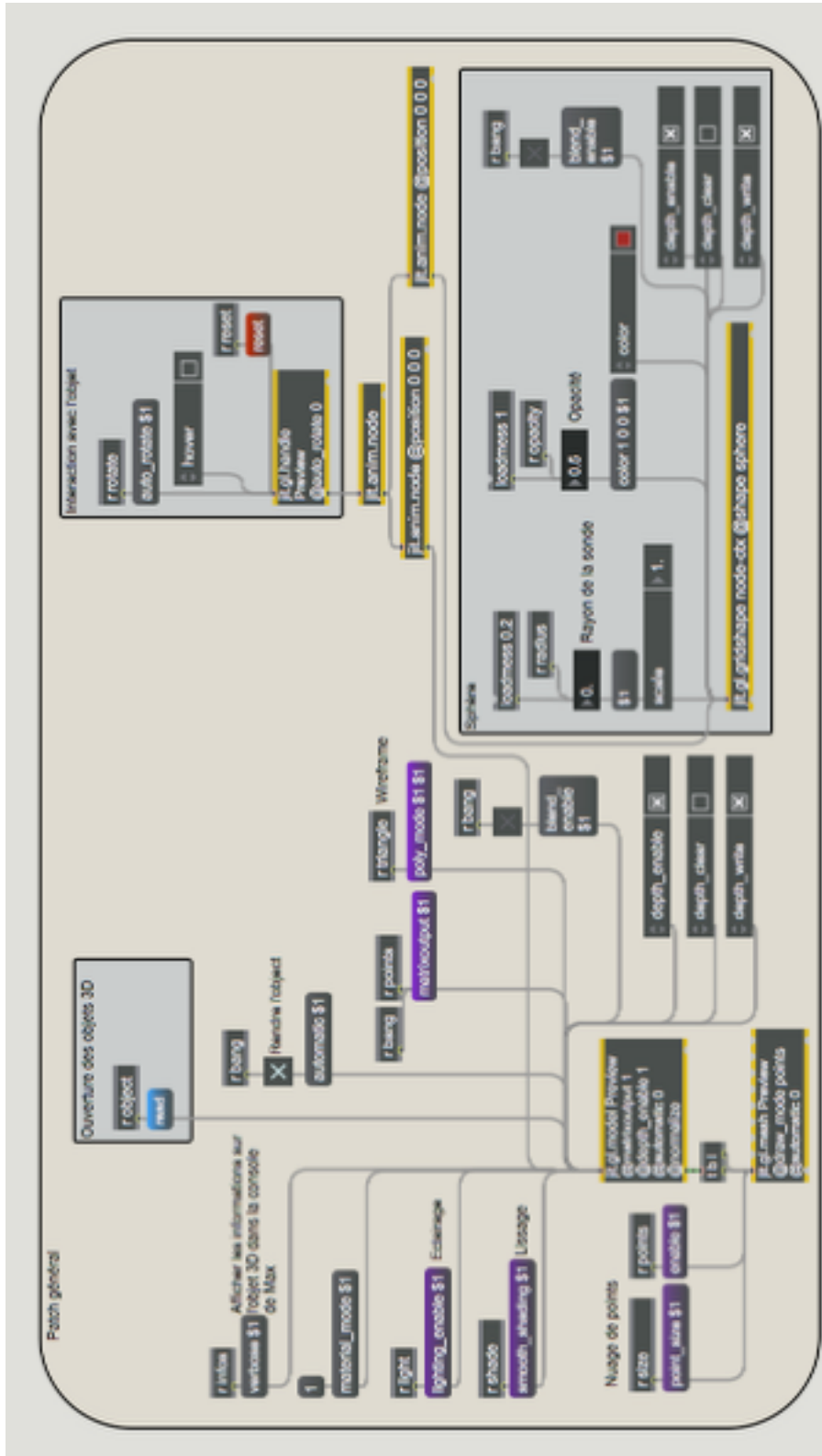


Fig. A.1 : Partie principale du patch de visualisation

B. Patch de sonification

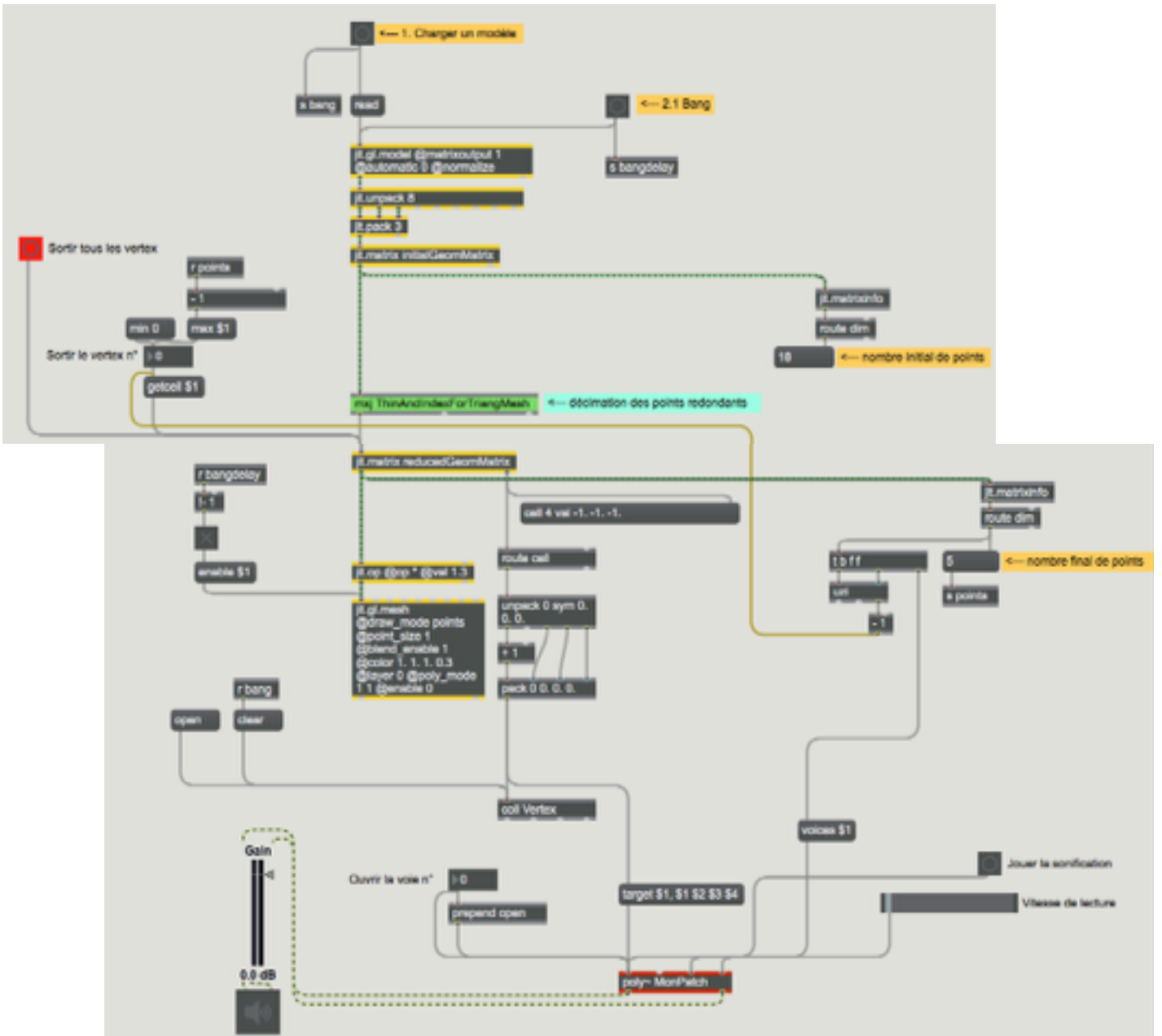


Fig. B.1 : Patch principal de sonification

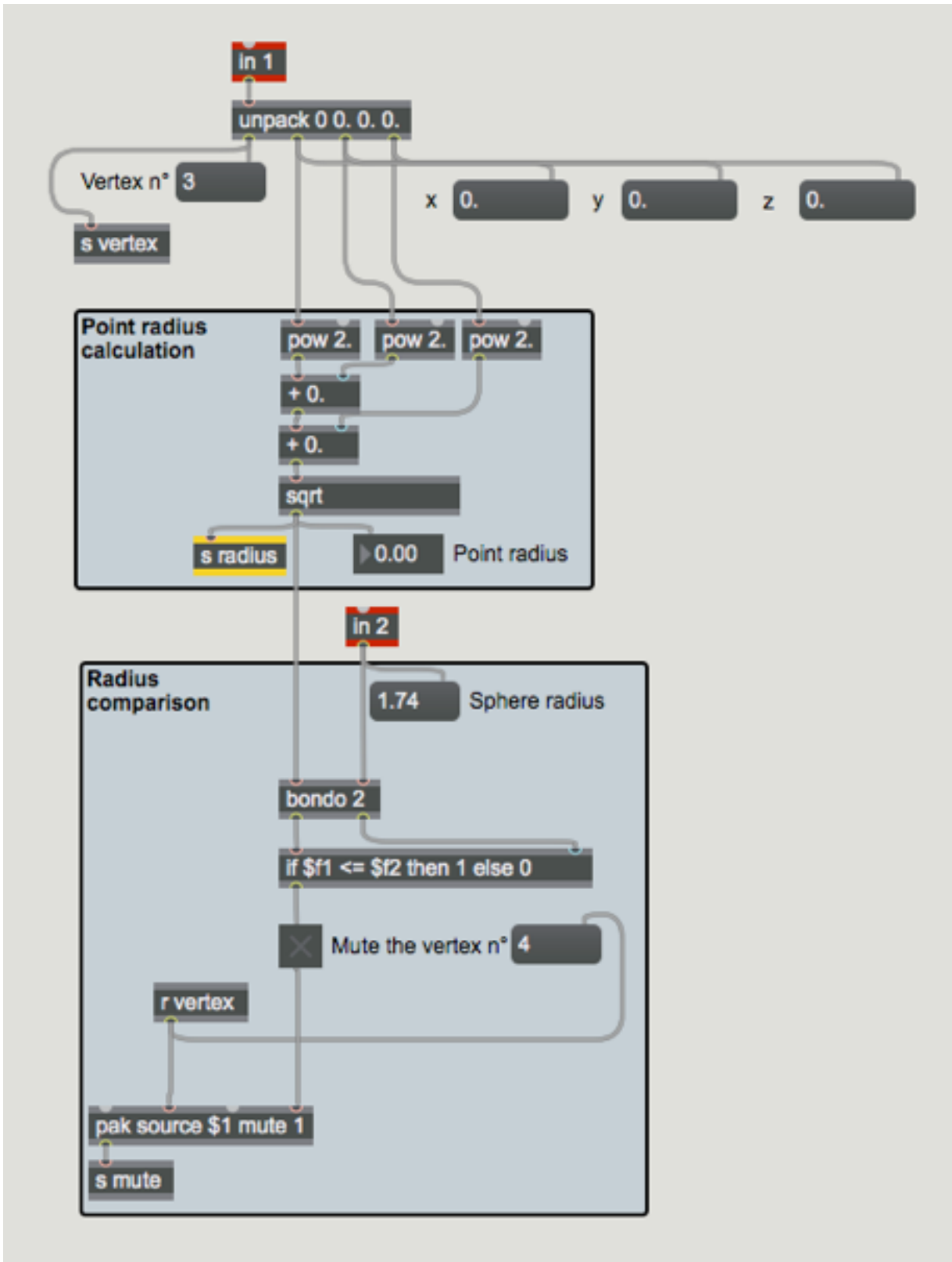


Fig. B.2 : Calculs sur le rayon, avec la brique logicielle de « mute »

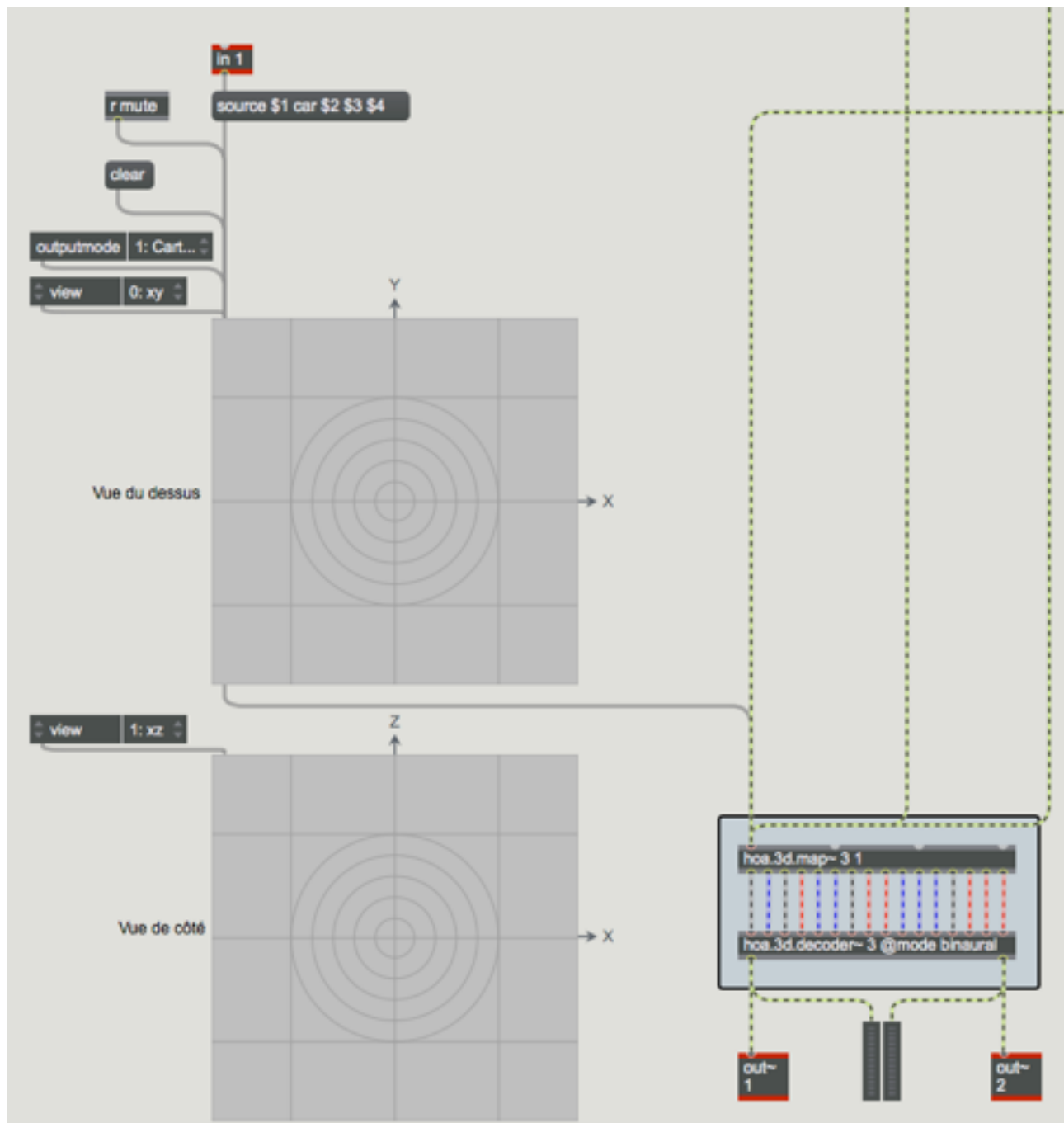


Fig. B.3 : Partie du patch qui gère la spatialisation Ambisonic

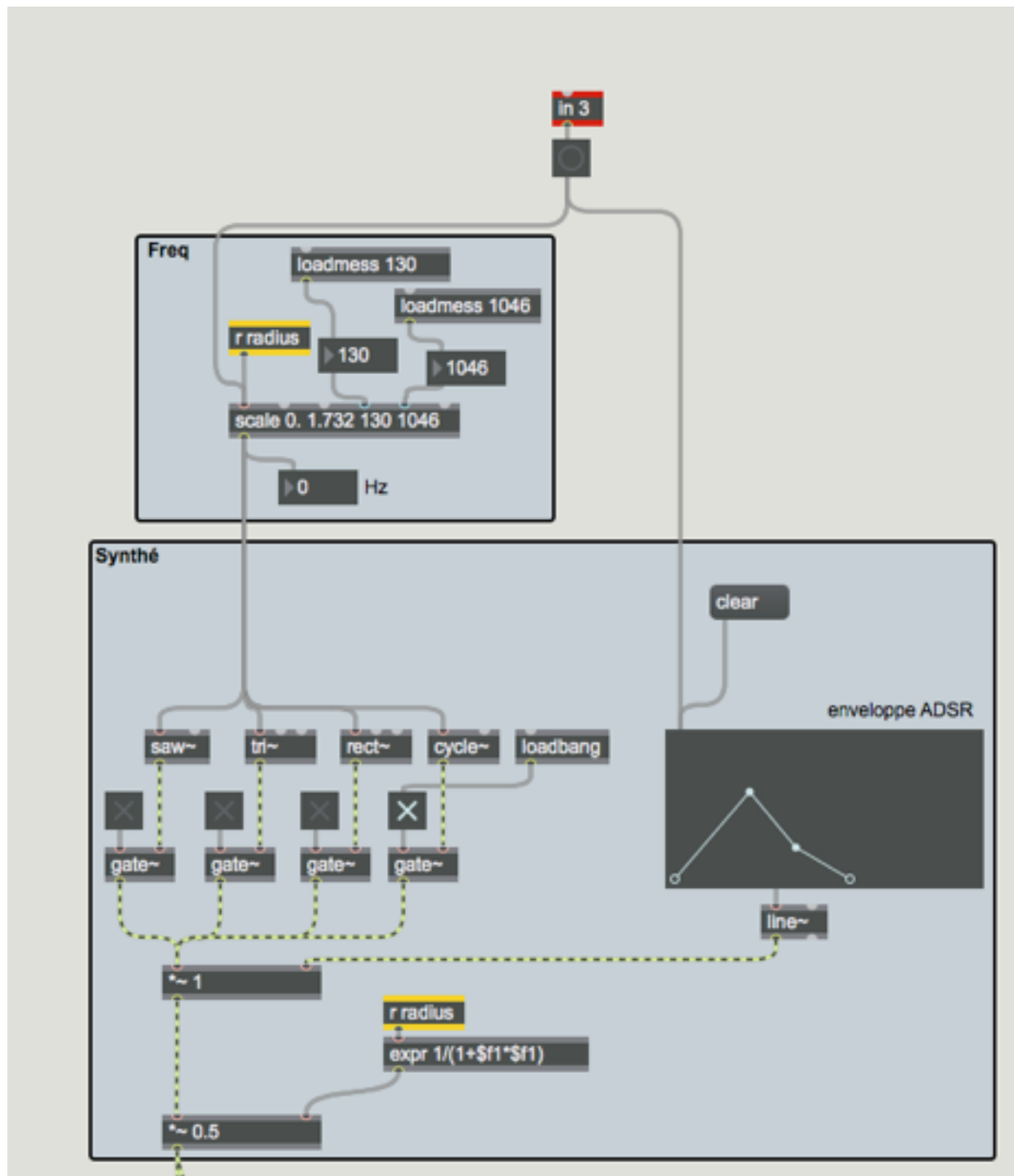


Fig. B.4 : Synthèse du son pur

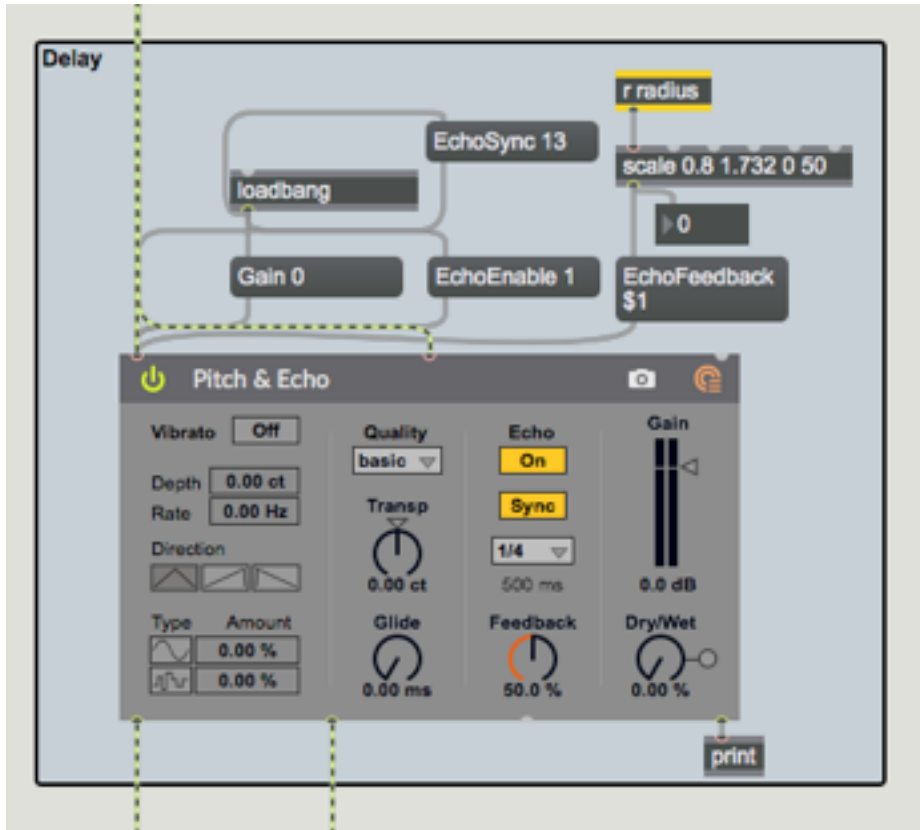


Fig. B.5 : Brique logicielle « Écho »

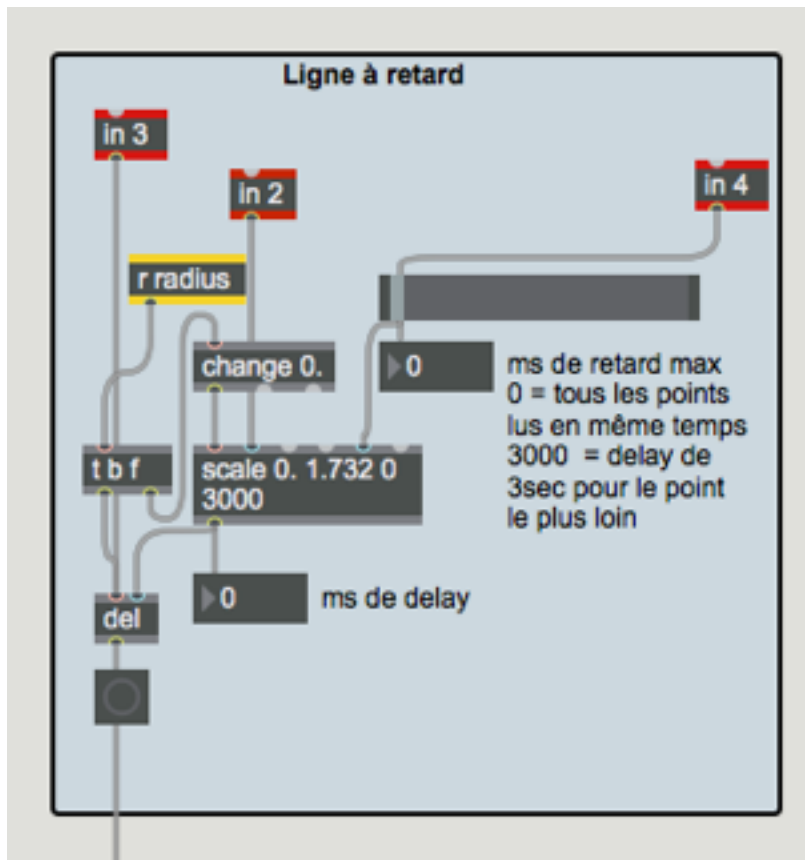


Fig. B.6 : Ligne à retard