



Stage réalisé au Laboratoire Informatique Acoustique Musique [LIAM]
de la Faculté de musique de l'Université de Montréal
sous la direction de Caroline Traube

**Réalisation d'une application informatique
pour l'analyse des échelles musicales
de chants traditionnels du Sud de l'Italie**

par

Patrice Guyot

Etudiant du Master ATIAM

Acoustique, traitement du signal et Informatique appliqués à la musique
de l'Université Pierre et Marie Curie Paris VI

Rapport de stage

Mars - août 2010

Remerciements

Je tiens à remercier tout particulièrement ma directrice de recherche Caroline Traube pour m'avoir donné l'opportunité d'effectuer ce stage sur un thème qui me tenait à coeur, pour son accueil au sein de la Faculté de musique de l'Université de Montréal, pour son enthousiasme, sa disponibilité, sa proximité et son soutien.

J'adresse un grand merci sincère à Flavia et Attilio pour leur accueil attentionné lors de mon séjour à Martano au Sud de l'Italie, pour leur hospitalité et leur gentillesse. Je remercie particulièrement Flavia Gervasi avec qui la collaboration fut passionnante lors ce séjour.

Je remercie les chercheurs qui ont pris le temps de répondre à mes questions et qui, chacun à leur manière, m'ont permis d'avancer dans ce travail : Nathalie Fernando, Jean During, Jérôme Barthelemy, Grégoire Carpentier, Geoffroy Peters

J'aimerais conclure ces remerciements par une attention toute particulière pour mes collègues de travail de Montréal, du LIAM à la Faculté de musique et du BRAMS, qui, en plus de m'assister dans ce travail avec disponibilité et compétence, ont largement contribué à me faire passer un agréable séjour au Québec : Michel Bernays, Nicolas Riche, Clément Lagarrigue, Sébastien Bel, Régis Trapeau, Frédéric Chiasson, Roxane Campeau, Jean-Michel Dumas et Maryse Lavoie.

Résumé

Le concept d'échelle musicale constitue l'un des points de rencontre historique entre recherche scientifique et musicale. Dans le cas du chant traditionnel, les analyses récentes s'appuient sur des progrès techniques issus du domaine du traitement du signal. Une démarche de l'ethnomusicologie contemporaine consiste par ailleurs à valider par une approche perceptive et interactive les hypothèses établies par l'analyse acoustique. Le présent stage s'inscrit dans le contexte d'une collaboration avec une chercheuse en ethnomusicologie, Flavia Gervasi, dont le projet de doctorat est consacré à l'analyse esthétique de chants traditionnels du Salento (région du Sud de l'Italie) qui se caractérisent par un fort mouvement de "revivalisme" (réinterprétation du répertoire traditionnel par des chanteurs de la nouvelle génération). Cette collaboration s'est concrétisée par la mise en place sur le terrain d'un test de perception et la réalisation d'une application informatique en vue de l'analyse des échelles musicales et du geste vocal (inflexions, micro-intonations, ...) utilisés par les chanteurs.

Déroulement

Ce stage de Master ATIAM s'est déroulé au Laboratoire Informatique Acoustique Musique (LIAM) de l'Université de Montréal, sous la direction de Caroline Traube. Le LIAM est affilié à la fois au secteur composition électroacoustique et au secteur musicologie de la Faculté de musique et accueille des chercheurs et étudiants provenant des différents domaines de la recherche consacrée à la musique. Ce stage fut l'objet d'une collaboration interdisciplinaire avec une doctorante en ethnomusicologie, Flavia Gervasi. Un travail de terrain de deux semaines a été effectué en début de stage à Martano au Sud de l'Italie.

What would you think if I sang out of tune ?

John Lennon et Paul McCartney

Table des matières

Introduction	1
1 Les échelles musicales du chant traditionnel	1
1.1 Contexte de travail : l'ethnomusicologie computationnelle	2
1.1.1 L'extraction automatique d'information musicale	2
1.1.2 L'ethnomusicologie	2
1.1.3 L'ethnomusicologie computationnelle	3
1.2 Les algorithmes d'extraction automatique d'information musicale	3
1.2.1 Détection de la fréquence fondamentale	4
1.2.2 Segmentation de notes	6
1.2.3 Quelques logiciels d'analyse audio	7
1.3 Les échelles musicales et leurs modes	8
1.3.1 Les échelles musicales	9
1.3.2 Les modes	11
1.4 La perception des intervalles musicaux	12
1.5 Le chant de tradition orale	13
1.5.1 Contexte d'exécution	13
1.5.2 Transmission orale	14
1.5.3 Les systématiques musicales	14
1.6 État de l'art : les études d'échelle musicales de chants traditionnels	15
1.6.1 Hypothèses	15
1.6.2 Les études	16
2 Le chant traditionnel du Salento	19
2.1 La musique du Sud de l'Italie	19
2.2 Etude comparative du chant traditionnel du Salento	20
2.2.1 Contexte général	20
2.2.2 Les hauteurs de notes	22
2.3 Le travail de terrain réalisé dans le cadre du stage	24
2.3.1 Test de perception des hauteurs	24
2.3.2 Mise en place d'une base de données	27
3 Réalisation d'une application informatique	29
3.1 Description du projet	29
3.1.1 Extraction des caractéristiques	30
3.1.2 Analyse et comparaison	30
3.1.3 Choix du langage de développement	31
3.1.4 Nomenclature des fichiers de données	31
3.2 Implémentation de la phase d'extraction des caractéristiques	32

3.2.1	L'extraction de la fréquence fondamentale	32
3.2.2	L'extraction de l'énergie	33
3.2.3	Traitements ultérieurs à l'extraction de la fréquence fondamentale	34
3.2.4	Segmentation de notes	37
3.2.5	L'extraction de la tonique	38
3.3	Implémentation des outils d'analyse et de comparaison	39
3.3.1	Histogramme d'intervalles	39
3.3.2	Affichages multiples	39
3.3.3	Affichage des intervalles utilisés	39
3.3.4	Lecteurs audio	40
3.3.5	Exportation	40
3.4	Interface graphique	40
3.5	Résultats	41
3.5.1	Résultats sur les échelles musicales	41
3.5.2	Le geste vocal	41
3.5.3	Utilisations alternatives de l'application	41
3.6	Perspectives de développement	42
3.6.1	Détection de similarités	42
3.6.2	Détection et élimination de la déviation	42
3.6.3	Etirement du temps	43
3.6.4	Prolongation par continuité	43
3.6.5	Modification du diapason	43
3.7	Discussion sur le développement	43
	Conclusion	44
	4 Annexes	46
4.1	Extraction de la fréquence fondamentale et traitements ultérieurs	46
4.2	Les outils réalisés	53
4.3	Les interfaces graphiques	56
4.4	Comparaison avec le logiciel Melodyne.	57
4.5	Exemples de modes musicaux	58
4.6	Manuel d'utilisation	59
	Références	61

Introduction

Si la partie la plus visible des recherches scientifiques liées à la musique réside dans la création d'environnements de production musicale, une autre plus discrète se révèle dans l'analyse de la musique. L'étude du chant traditionnel, qui appartient au domaine de l'ethnomusicologie, constitue un exemple de champ d'interaction. Cette étude est parfois soumise aux problématiques liées aux répertoires de transmission orale, et nécessite dans ce cas une méthodologie adaptée à l'analyse des caractéristiques et de l'évolution de ces répertoires.

La collaboration entre informatique et ethnomusicologie est un domaine de recherche peu développé. Cependant, grâce à une approche objective basée sur l'étude acoustique, elle permet la validation de certaines hypothèses de recherche et parfois l'élaboration de nouvelles. Ce travail est le fruit d'une telle collaboration, dont l'objet de recherche est le chant traditionnel du Sud de l'Italie dans la région du Salento.

Nous nous sommes, dans cette étude, focalisés sur l'analyse des hauteurs de notes, ce qui soulève des problématiques liées au concept historique d'échelle musicale. Ce travail a débouché sur la réalisation d'un test perceptif mis en place sur le terrain du Salento. Par la suite, la majeure partie de ce stage fut tournée vers la création d'une application informatique autonome, afin d'analyser et de comparer les hauteurs de notes utilisées dans le chant traditionnel du Salento.

1 Les échelles musicales du chant traditionnel

Ce chapitre décrit quelques bases techniques et théoriques préalables à la présentation de notre travail dans le Salento (Chapitre 2) et à la réalisation d'une application informatique (Chapitre 3) en vue de l'analyse des hauteurs de notes et des échelles musicales utilisées dans le chant traditionnel du Sud de l'Italie.

Après avoir précisé le domaine de recherche dans lequel s'inscrit ce travail (section 1.1), nous détaillerons d'un point de vue technique les algorithmes utilisés lors de l'extraction automatique de notes de musique (section 1.2). Nous présenterons ensuite quelques notions relatives aux échelles musicales (section 1.3), à leur perception (section 1.4), et d'autres liées au contexte du chant traditionnel (section 1.5). Nous citerons enfin les principaux travaux répertoriés traitant de l'analyse

d'échelles musicales dans le chant traditionnel (section 1.6).

1.1 Contexte de travail : l'ethnomusicologie computationnelle

1.1.1 L'extraction automatique d'information musicale

La recherche en informatique appliquée à la musique a fait émerger un domaine de recherche : l'extraction automatique d'information musicale, plus connue en anglais sous le nom de *Music Information Retrieval* ou *MIR* (Typke et al., 2005). Ce champ de recherche, dont les problématiques relèvent de l'extraction d'informations caractéristiques d'une musique à partir de son enregistrement audio, est au cœur de grandes campagnes de recherche actuelles, telles que le projet *Quaero*. Les caractéristiques recherchées dans les morceaux audio sont pour certaines applications liées à la transcription automatique, comme la détection de hauteurs de notes ou de rythme. Des applications grand public de ces travaux sont aujourd'hui disponibles et connaissent un succès important.

On peut pourtant regretter que la plupart des applications de recherche d'information musicale se destinent aux musiques occidentales, dont certaines caractéristiques sont assez uniformes comme l'utilisation de l'échelle tempérée. La recherche d'information musicale présente pourtant un potentiel très intéressant dans les musiques extra-occidentales, notamment les musiques traditionnelles, et pourrait fournir des outils très utiles aux ethnomusicologues.

1.1.2 L'ethnomusicologie

L'ethnomusicologie est une science humaine ayant vu le jour à la fin du XIXe siècle sous le nom de musicologie comparée. Cette discipline s'est vu proposer de nombreuses définitions depuis son apparition, aussi nous nous contenterons de constater que son champ d'étude s'applique aux musiques traditionnelles de transmission orale, contrairement à la musicologie qui rend compte de la musique écrite occidentale (notons qu'aujourd'hui la distinction entre les deux disciplines tend à s'estomper).

Cette discipline a su rapidement prendre conscience du potentiel qu'offrait la technologie. Les chercheurs, par exemple, utilisèrent rapidement des enregistreurs portables, qu'ils préféraient à une

transcription manuelle jugée plus réductrice. Ainsi en 1951, Charles Seeger décrivait une machine de transcription automatique de la fréquence fondamentale, utilisée par des musicologues plutôt que par des physiciens, afin d'étudier les musiques traditionnelles (Seeger, 1951). Un demi-siècle plus tard, les algorithmes de traitement du signal semblent potentiellement permettre la réalisation de ce type d'outil.

1.1.3 L'ethnomusicologie computationnelle

Il semble donc de nos jours qu'un nouveau champ de recherche se dessine au croisement de ces deux domaines, que nous appellerons en référence à Tzanetakis et al. (2007), *ethnomusicologie computationnelle*. Ce champ de recherche fait collaborer les scientifiques et les musicologues afin de mieux déterminer les caractéristiques de la musique traditionnelle. Il peut s'appliquer à la symbolique ou aux modèles utilisés pour décrire la musique, à la transcription automatique, ou encore à l'archivage (Proutskova, 2007). L'application des technologies informatiques permet de travailler sur des corpus importants et d'utiliser des outils statistiques, qui peuvent contribuer à une vérification objective des hypothèses ethnomusicologiques.

Ainsi, c'est dans le champ de l'ethnomusicologie computationnelle que nous plaçons ce travail de recherche né d'une collaboration avec Flavia Gervasi, doctorante en ethnomusicologie.

1.2 Les algorithmes d'extraction automatique d'information musicale

Nous nous sommes concentrés, dans le cadre de ce stage, sur l'étude des hauteurs de notes dans la voix chantée, et nous présentons dans cette partie les principaux algorithmes utilisés pour extraire de façon automatique les notes et leurs hauteurs à partir d'un enregistrement audio.

Rappelons tout d'abord que tout signal périodique, et notamment les voyelles d'un chant, peut être décomposé en une somme de sinusoides, appelées harmoniques ou partiels, dont les fréquences sont toutes multiples d'une fréquence particulière, appelée fréquence fondamentale. La détermination de cette fréquence fondamentale est un problème préalable à la détermination des hauteurs de notes. L'étape suivante consiste à délimiter des frontières afin de segmenter la fréquence fondamentale en zones stables qui caractériseront les notes (Isikhan, 2008). D'autres caractéristiques du signal peuvent être également évaluées pour contribuer à sa segmentation, et en particulier son énergie.

1.2.1 Détection de la fréquence fondamentale

De nombreuses méthodes et algorithmes ont été proposés pour l'extraction de la fréquence fondamentale, ou *pitch* en anglais. Aujourd'hui encore, la détection polyphonique de fréquence fondamentale, souvent appelée en anglais *multipitch*, est un thème central des campagnes d'extraction d'informations musicales. Dans le cas de la monodie, où une seule hauteur de note est présente à chaque instant, le problème de détection de fréquence fondamentale a fait des progrès considérables depuis vingt ans. Toutefois, nous verrons dans notre application pratique que la détection de fréquence fondamentale présente tout de même des erreurs, même pour les algorithmes réputés robustes. Pour utiliser la fréquence fondamentale détectée, il est donc souvent nécessaire de procéder à des traitements ultérieurs afin de réduire les erreurs. Une erreur de détection courante est par exemple l'estimation du double ou de la moitié de la fréquence fondamentale réelle, qui correspondent à des sauts d'octave.

Autocorrélation

Dans le cadre de ce stage, nous avons utilisé dans un premier temps la fonction d'autocorrélation. Cette fonction utilise une multiplication du signal par le signal lui-même décalé dans le temps par un facteur τ , le tout sommé sur une fenêtre d'analyse de taille W .

$$r(\tau) = \sum_{j=1}^W x_j x_{j+\tau} \quad (1)$$

En cherchant le premier maximum de cette fonction, nous devons théoriquement tomber sur la première période du signal, soit l'inverse de la fréquence fondamentale. Nous pouvons ainsi trouver pour chaque fenêtre de calcul de taille W une fréquence fondamentale et utiliser une fenêtre glissante pour analyser tout le signal.

Yin

Nous avons par la suite utilisé l'algorithme Yin (De Cheveigné, 2002). Cette méthode rapide et efficace s'appuie sur l'autocorrélation, mais utilise la différence entre le signal et le signal décalé, le tout élevé au carré.

$$r(\tau) = \sum_{j=1}^W (x_j - x_{j+\tau})^2 \quad (2)$$

La recherche de la première période correspond ici à la détection du premier minimum de cette fonction. Dans la définition de l'algorithme Yin, suivent des étapes de normalisation par la moyenne cumulée, seuillage et interpolation. Cet algorithme très utilisé aujourd'hui notamment dans les traitements en temps réel, donne de bons résultats. Il est beaucoup moins sensible aux différences d'amplitude que la simple fonction d'autocorrélation.

SHR (Subharmonic-to-Harmonic Ration)

Nous avons enfin utilisé une dernière méthode, cette fois-ci spectrale, dédiée à l'extraction de fréquences fondamentales dans la voix (Sun, 2002). Selon Sun, les signaux vocaux sont caractérisés par des cycles alternés d'amplitude dus à une instabilité du système des cordes vocales (Sun, 2000). Ces cycles font apparaître dans le spectre du signal des fréquences sous-harmoniques.

Dans l'algorithme SHRP, la somme SH des amplitudes des harmoniques est calculée. Ici $A(f)$ est le spectre d'amplitude du signal, f_0 sa fréquence fondamentale, et N le nombre maximum d'harmoniques :

$$SH = \sum_{n=1}^N A(nf_0) \quad (3)$$

De même la somme des fréquences sous-harmoniques SS, en supposant que la plus basse est égale à la moitié de la fréquence fondamentale :

$$SS = \sum_{n=1}^N A((n - \frac{1}{2})f_0) \quad (4)$$

Le rapport entre sous-harmoniques et harmoniques peut ainsi être calculé :

$$SHR = \frac{SS}{SH} \quad (5)$$

Sun utilise ce rapport, ainsi qu'un seuil déterminé de façon perceptive, afin d'évaluer la présence de fréquences sous-harmoniques dans la voix. Le calcul nécessite l'évaluation de maxima locaux. L'évaluation de la présence de sous-harmoniques dans la voix permet d'évaluer la fréquence fondamentale en réduisant considérablement les sauts d'octave.

Traitements ultérieurs

Des traitements ultérieurs sont souvent souhaitables pour améliorer les résultats de l'algorithme d'extraction de fréquence fondamentale. McNab utilise par exemple le concept de "construction d'île" (McNab et al., 1996). Des zones de stabilité, contenant plusieurs échantillons de même hauteur sont détectées, puis étendues en corrigeant les sauts d'octave. Un filtrage par la moyenne est également utilisé.

Un autre type de filtrage courant consiste à utiliser un filtre médian. Le principe du filtre médian est d'ordonner les échantillons d'une fenêtre par ordre croissant, et de retourner la valeur de l'échantillon placé au milieu de l'ensemble (position médiane). Les points aberrants, comme les sauts d'octave, sont ainsi éliminés, car placés en début ou en fin de file. Ce filtre, comme le filtre moyeneur, est appliqué sur l'ensemble du signal par une fenêtre glissante. De la taille de la fenêtre dépend donc l'ampleur de l'effet du filtre. Une grande fenêtre aura tendance à lisser le signal de façon importante alors qu'une petite n'aura que peu d'effet.

1.2.2 Segmentation de notes

L'étape de segmentation de notes consiste à délimiter des frontières dans le flux de fréquence fondamentale, en utilisant les caractéristiques du signal. Ces caractéristiques, ou "fonctions d'observation", peuvent décrire les variations d'énergie ou de fréquence fondamentale, mais aussi de contenu spectral (Rossignol, 2000). Les résultats de ces fonctions d'observation peuvent être utilisés de manière indépendante ou être fusionnés pour obtenir une décision finale sur la segmentation.

L'utilisation des sauts d'énergie peut être bien adaptée à la segmentation de notes séparées par des silences ou des attaques. Une méthode courante consiste à utiliser des seuils appliqués à l'énergie ou à la variation d'énergie afin d'estimer des frontières.

La segmentation peut également être effectuée à l'aide des variations de fréquence fondamentale. Un algorithme proposé par McNab groupe des trames d'échantillons de fréquence fondamentale en calculant leur moyenne. Un groupe d'une durée supérieure à 100ms devient une note. La trame suivante est intégrée à la note si sa fréquence moyenne ne dépasse pas un quart de ton de différence avec celle de la note.

Rossignol propose de nombreuses fonctions d'observation pour la segmentation dont deux utilisent la fréquence fondamentale. L'une est basée sur un modèle statistique appliqué à chaque partiel du signal. Dans l'autre le signal est modélisé par un processus autorégressif, et le changement de note est détecté par un point de rupture. Par ailleurs, Rossignol propose une détection suivie d'une élimination du vibrato afin d'améliorer la segmentation en zones stables.

Attribution d'une étiquette

Une fois la fréquence fondamentale segmentée en notes, il reste à attribuer une hauteur à chacune des notes. Cette hauteur peut être attribuée de façon absolue, par exemple en Hertz, mais il est souvent plus adapté dans les applications musicales d'utiliser comme étiquette une note de musique, ce qui implique une quantification et soulève la question d'échelle musicale dont nous parlerons plus loin. Une autre difficulté réside dans la variation de la fréquence fondamentale au cours de la note. Cette variation est très courante, elle est bien sûr présente lors d'un vibrato, mais l'est aussi souvent aux extrémités des notes. Dans le cas du chant, il est très possible que la fréquence fondamentale ne soit jamais totalement stable. Les vibratos réguliers semblent toutefois être perçus à leur hauteur moyenne (Sundberg, 1979). Ainsi McNab utilise-t-il la moyenne de la fréquence fondamentale sur la durée totale de la note pour en déterminer la hauteur telle qu'elle est perçue.

1.2.3 Quelques logiciels d'analyse audio

Les logiciels d'analyse audio utilisent des algorithmes d'extraction automatique d'information musicale afin, d'une part de proposer des traitements de plus haut niveau et d'autre part, de permettre une utilisation facile et intuitive. Nous présentons dans cette partie les principaux logiciels utilisés lors de ce stage.

Praat Praat¹ est un logiciel libre conçu pour la manipulation, le traitement et la synthèse de sons vocaux. Il est couramment utilisé dans les recherches scientifiques liées à l'analyse de la voix.

Melodyne Melodyne² est un logiciel d'analyse et de modification audio dédié aux professionnels de l'enregistrement. Il permet une visualisation de la fréquence fondamentale et des notes jouées. Ce logiciel peut être utilisé pour modifier les hauteurs de notes jouées, par exemple pour réajuster un enregistrement vocal, sans en modifier le timbre ou les inflexions.

Audiosculpt Audiosculpt³ est un logiciel d'analyse et de modifications audio. Basé sur la transformée de Fourier, ce logiciel permet divers traitements de haute qualité. En outre, Audiosculpt permet d'exporter les données d'analyse, ce qui permet de les utiliser avec le logiciel OpenMusic⁴ adapté à l'analyse musicale.

1.3 Les échelles musicales et leurs modes

Nous traiterons dans cette section des notions d'échelles et de modes. L'échelle musicale est définie comme une segmentation d'un intervalle cadre, en général l'octave, et représente l'ensemble des notes possibles à l'intérieur de ce cadre. Le mode représente les notes choisies sur cette échelle pour constituer de façon partielle ou totale une pièce musicale (Fernando, 2003). Ainsi dans le système occidental le plus couramment employé, souvent appelé tempérament égal, l'échelle est composée de douze demi-tons qui découpent l'octave de façon logarithmiquement égale, et le mode de *Do majeur* est constitué des notes *Do Ré Mi Fa Sol La Si*.

¹Praat a été conçu à l'Institut de sciences phonétiques de l'Université d'Amsterdam par Paul Boersma et David Weenink.

²Melodyne Studio, Copyright Celemony Software, Conçu par Peter Neubaecker, Carsten Gehle, Christoph Kreutzkamp, Uwe Granzow.

³AudioSculpt, Copyright IRCAM. Conçu et développé par l'équipe Analyse-Synthèse.

⁴OpenMusic, Copyright IRCAM. Conçu et développé par l'équipe Représentation musicales.

1.3.1 Les échelles musicales

Historique et problématiques

Depuis l'Antiquité jusqu'à nos jours, de nombreux musiciens et scientifiques se sont interrogés sur la façon de définir les échelles musicales. Les systèmes proposés ont été régulièrement remis en question en fonction des genres et des formes musicales dominantes à chaque époque.

La création d'une échelle passe en effet par la définition précise de ses intervalles. La définition des intervalles dits justes ou naturels repose sur la résonance naturelle des sons harmoniques, dont les partiels sont des multiples entiers de la fréquence fondamentale. L'octave naturelle est ainsi définie comme le rapport entre la fréquence fondamentale et le premier harmonique (rapport de fréquence $\frac{1}{2}$), alors que la quinte naturelle sera définie par le rapport entre la première et la deuxième fréquence harmonique (rapport de fréquence $\frac{2}{3}$)(Asselin, 1984).

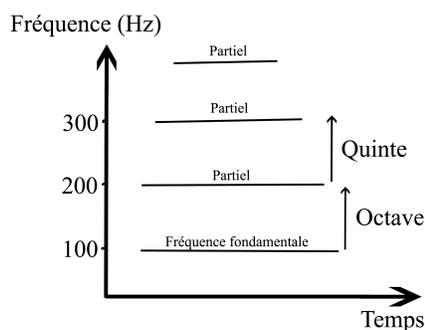


FIG. 1 – Spectre harmonique d'une note stable.

Ces intervalles naturels se retrouvent par divisions successives d'une corde vibrante. Par exemple, la quinte obtenue par le rapport de corde $\frac{2}{3}$ est rigoureusement équivalente à la quinte obtenue grâce aux harmoniques naturelles. Ainsi les Pythagoriciens en 500 av J.C définissaient les intervalles consonants comme des intervalles obtenus par division d'une corde par des rapports particuliers utilisant les chiffres 1, 2, 3 et 4 (connus sous le nom de Tétraktys), ce qui leur permettait d'obtenir des octaves, quintes et quarts pures. En utilisant une succession de quintes pures ramenées à la même octave, la généralisation de ce système permet de construire une gamme chromatique appelée gamme *pythagoricienne*. La tierce obtenue dans le système pythagoricien (408 cents) est par contre plus grande que celle du tempérament égale (400 cents) et que la tierce juste

(386 cents) (Devaney, 2008). De plus, douze quintes consécutives ne donnent pas l'octave et font apparaître un comma dit *pythagoricien*. Au XVI^e siècle, Zarlino propose de réduire cette imperfection et construit une gamme s'appuyant sur l'accord majeur (Betten, 2001). Cette gamme fut peu à peu remplacée par l'adoption du tempérament égal, qui sépare l'octave en douze demi-tons égaux. Même si ce tempérament égal semble régler le problème de l'accordage, il n'en reste pas moins un compromis où la plupart des intervalles sont différents des intervalles naturels. La quinte définie dans le système tempéré (700 cents), utilisée dans la grande majorité des musiques occidentales, est ainsi légèrement plus petite que la quinte naturelle définie sur la base des harmoniques (702 cents).

Alternatives au système tempéré

Si le tempérament égal est utilisé en Occident dans la grande majorité des musiques, il est important de mentionner d'autres approches plus singulières, dans un premier temps au sein de la musique savante occidentale. Il existe par exemple en Europe un débat esthétique visant à faire jouer les œuvres anciennes sur des tempéraments de l'époque, dont une application pratique est la restauration d'orgues anciens (Bougeret, 1989). Il existe aussi dans la création contemporaine une volonté d'utiliser d'autres échelles variant du système tempéré, par exemple dans la musique microtonale.

L'application directe de la construction d'échelles se retrouve évidemment dans la facture d'instruments de musique. Certains instruments ont par nature une échelle fixée à l'accordage (le piano par exemple). D'autres instruments, et nous pouvons ici inclure la voix, possèdent des aptitudes plus ou moins grandes à utiliser l'ensemble du continuum sonore de leur registre. Il est clair que les premiers cadrent les seconds quand ces deux types d'instruments sont amenés à jouer ensemble. Les chanteurs sont par exemple priés de chanter "juste", c'est-à-dire de chanter les tons de l'échelle tempérée, lorsqu'ils ont accompagnés par un piano. Il semble néanmoins que si les chanteurs ne sont pas "contraints" par un piano, ils se dirigent naturellement vers des intervalles naturels, ce qui peut avoir comme conséquence la modification du diapason du morceau au cours du temps (Howard, 2007), (Devaney, 2008).

Dans les musiques non occidentales, différentes échelles sont employées. Par exemple, le système indien utilise selon la tradition une échelle de 22 intervalles par octave appelés *shrutis*. Par ailleurs en Iran, on utilise depuis des siècles le *ton neutre*, intervalle correspondant à la moitié d'une tierce mineure. Il existe ainsi de nombreux systèmes différents de par le monde, et les ethnomusicologues

ont souvent essayé de caractériser ces échelles musicales. Le premier grand travail effectué dans ce sens date probablement de Ellis (1885), qui a d'ailleurs défini la notion de *cent* que nous utilisons dans ce rapport pour caractériser les intervalles (un cent étant par définition égale à un centième de demi-ton). D'autres chercheurs ont mené des études perceptives et ethnomusicologiques pour discuter des origines possibles des échelles, et de la potentielle universalité de certains intervalles (Burns, 1999).

Les échelles utilisées dans différentes cultures sont également l'objet d'une évolution qui est sans doute accélérée par les rencontres entre musiques d'origines géographiques diverses. Nous pouvons parfois observer cette évolution, en Turquie par exemple où nous assistons à un déplacement des frettes des instruments à cordes pincées au cours du XXe siècle. Ce processus vise la plupart du temps à rendre compatible ces instruments avec d'autres qui utilisent la gamme tempérée, par exemple le piano (During, 2004). Cette uniformisation des échelles vers le tempérament égal peut être accélérée de nos jours avec l'arrivée des instruments électroniques et du protocole MIDI.

1.3.2 Les modes

Les modes musicaux représentent une sélection des degrés de l'échelle utilisée dans un morceau. Même s'il peut faire référence au mode majeur ou mineur dans la musique occidentale, le terme mode est également utilisé pour décrire les gammes utilisées par les Grecs anciens, les gammes majeures ou mineures étant des cas particuliers de ces modes. Il est toutefois difficile de savoir avec exactitude les intervalles qui caractérisaient ces modes dans l'Antiquité. Les appellations de ces modes, telles que les termes *lydien, éolien etc...*, sont pourtant largement utilisés de nos jours dans la théorie musicale et notamment en jazz. Ces modes sont décrits de façon plus précise en annexe à la page 58.

Dans la musique traditionnelle, le terme "mode" peut être utilisé pour généraliser un type de mélodies (Powers, 2004). Ces modes, aux caractères parfois exotiques, ont influencé de nombreux compositeurs, notamment Béla Bartók. Dans la musique indienne, le mode est également utilisé pour caractériser un sentiment à transmettre ou un moment particulier de la journée.

1.4 La perception des intervalles musicaux

Dans le cadre de ce travail qui vise à analyser et quantifier la trace acoustique de chants traditionnels, il faut également prendre en compte la dimension perceptive qui intervient autant chez le chanteur que chez le chercheur. Ainsi, même si l'objet de cette section n'est pas une étude exhaustive du domaine, il est important d'avoir quelques pistes de réflexion pour évaluer les facteurs psychoacoustiques rentrant en jeu dans la perception des échelles musicales.

La perception des échelles semble intimement liée à la perception des intervalles qui la composent. Cette dernière peut être influencée par des facteurs très différents, comme le reflètent les nombreuses études du domaine. Ainsi selon Russo, la taille subjective des intervalles est influencée par l'entraînement, la direction montante ou descendante des intervalles, ou encore le registre (Russo, 2005). Concernant ce dernier point, il semble en effet que les intervalles perçus dans les sons aigus soient plus petits que les intervalles effectifs, comme le montrent certains travaux sur les échelles de mel.

D'autres études concernent les relations entre timbre et fréquence fondamentale. Selon Warrier, si beaucoup d'études sont contradictoires, il semble tout de même que le timbre ait une influence sur la fréquence fondamentale perçue (Warrier, 2002). Ce résultat fait écho aux enquêtes de Simha Arom et son équipe, qui constatèrent que les notions de fréquence fondamentale et de timbre étaient intimement liées en Afrique centrale.

D'autres études s'intéressent à l'influence de la culture sur la perception des intervalles et des mélodies. Selon Schellenberg, la connaissance implicite de l'échelle occidentale requiert des années d'exposition (Schellenberg, 1999). La perception des intervalles pourrait aussi être influencée par la langue maternelle (Xu et al., 2005).

Finalement, la perception des échelles chez le musicologue peut également être perturbée par des connaissances induites. Dans une étude d'Ambrzevicius (2008) sur la perception des intervalles dans le contexte de chants traditionnels lithuaniens, l'auteur modifie les hauteurs de notes des chants pour les ajuster sur la gamme tempérée. Un test perceptif est ensuite mis en place auprès d'un groupe de participants musicologues, priés d'effectuer la transcription des chants dont certains sont originaux alors que d'autres ont été modifiés à leur insu. Il est surprenant de constater que, dans le cas des chants originaux comme dans le cas des chants modifiés, les musicologues utilisent des symboles employés dans la musique microtonale. Ainsi semblent-ils influencés par

des connaissances implicites de ces chants qui *ne devraient pas* utiliser la gamme tempérée. Par cette étude, Ambrazevicius remet en question l'origine couramment attribuée de la musique traditionnelle lithuanienne aux modes de la Grèce antique.

1.5 Le chant de tradition orale

Ce travail est lié à un corpus de chants traditionnels transmis oralement. La pratique de ce type de chant est commune à de nombreuses cultures de par le monde. De nos jours, bien qu'ils restent très présents sur d'autres territoires, ces chants ont pratiquement disparu dans la plupart des pays d'Europe occidentale. Il en subsiste pourtant des traces dans certaines régions européennes, le plus souvent en milieu rural. Cette forme musicale est actuellement mise en valeur et profite d'une diffusion sous l'appellation "Musiques du Monde". Son étude musicale est effectuée par les ethnomusicologues depuis l'avènement de leur discipline.

Les caractéristiques de ces chants traditionnels sont très différentes d'une culture à l'autre, mais nous pouvons néanmoins y relever les quelques grandes lignes directrices suivantes.

1.5.1 Contexte d'exécution

Le chant traditionnel peut être utilisé comme musique fonctionnelle. Dans ce cas, sa pratique est reliée à un événement précis, vécu de manière collective ou individuelle. On peut ainsi trouver au sud de l'Italie des chants de travail liés aux activités agricoles, des berceuses, des chants de deuil ou encore des chants d'amour (Gervasi, 2010b). La nature de ces chants dépend fortement de leur contexte d'exécution. Leur étude est ainsi confrontée à un dilemme :

- Il est nécessaire d'une part, afin d'extraire des caractéristiques objectives et précises, d'étudier le musicien hors de son contexte. Une étude sur le timbre demande par exemple des enregistrements effectués dans des conditions rigoureusement identiques telles que celles du studio d'enregistrement.
- D'autre part, ces chants étant intimement liés à leurs contextes, les en priver risquerait de les dénaturer profondément et donc de modifier la source même des caractéristiques que l'on voulait objectives. Il peut ainsi être difficile pour un interprète de musique traditionnelle de

considérer cette musique en dehors de son contexte habituel, qui peut s'avérer très éloigné de la situation où un chanteur professionnel se produit dans un lieu de spectacle ou enregistre en studio. Ainsi la démarche de l'ethnomusicologue est plutôt d'étudier et d'enregistrer ces chants dans leur contexte traditionnel d'exécution.

1.5.2 Transmission orale

Les chants traditionnels peuvent être de transmission orale, c'est-à-dire transmis au fil des générations par la voix et surtout sans partitions. En général, les interprètes n'utilisent pas les notions théoriques de la musique occidentale. Ce type de transmission entraîne une certaine variabilité et une évolution du répertoire et de son interprétation qu'il est très difficile d'évaluer. Les concepts de théorie musicale utilisés dans l'apprentissage et la description de la musique occidentale, tels que le rythme, les hauteurs de notes, ou le timbre ne sont pas forcément utilisés et peuvent même être, dans certains cas complètement amalgamés car perçus de façon globale, comme le montre par exemple des études d'échelles musicales effectuées en Afrique centrale (Voisin, 1994). L'étude de cette musique par l'ethnomusicologue demande ainsi une approche sensiblement différente de la démarche musicologique basée sur la partition.

1.5.3 Les systématiques musicales

Une procédure d'analyse consiste à caractériser le chant par des paramètres isolés, et d'étudier la récurrence de ces paramètres au sein des chants afin de savoir s'ils caractérisent ou non la tradition étudiée.

Cette approche est donc orientée vers l'étude des paramètres qui caractérisent le chant traditionnel, que nous appellerons *systématiques musicales*. Les systématiques musicales peuvent être par exemple de l'ordre du timbre et de la qualité de la voix (Le Bagousse, 2008), du rythme ou des hauteurs. Pour chacun de ces axes de travail, une étude doit être effectuée afin de déterminer si les paramètres peuvent être déterminants pour définir les systèmes musicaux implicites adoptés par la communauté culturelle étudiée.

A titre d'exemple, une méthodologie d'ethnomusicologie contemporaine, utilisée par l'équipe de Simha Arom (Arom et al., 2007) consiste à :

- Extraire des chants des paramètres précis
- Modifier individuellement ces paramètres au sein des chants
- Mettre en place des tests de perception utilisant des chants originaux et des chants modifiés
- Faire passer les tests aux personnes considérées comme garantes de la tradition
- Evaluer si les paramètres sont déterminants pour définir les systèmes musicaux implicites

Le test de perception appliqué au corpus étudié au cours de ce stage s'inspire de cette méthodologie.

1.6 État de l'art : les études d'échelle musicales de chants traditionnels

1.6.1 Hypothèses

Dans le cadre du chant traditionnel, nous pouvons constater des récurrences dans les hauteurs de notes et les intervalles utilisés. Certains chercheurs ont ainsi supposé que les chanteurs étudiés suivent un modèle implicite guidé par la tradition, qui pourrait être une échelle musicale. Toutefois, l'étude de ces échelles se révèle beaucoup plus compliquée que la mesure de distance séparant les frettes d'un instrument à cordes pincées. L'analyse acoustique du chant est une technique d'étude récente pour déterminer ces échelles. L'analyse des échelles peut en outre devenir très complexe si les ornements, largement utilisés dans la musique traditionnelle, sont prises en compte. Ainsi est-il difficile de caractériser les notes utilisées par un chanteur lorsqu'il effectue des glissandi, ou autres ornements, souvent caractérisés par un passage très court sur des notes étrangères à l'échelle. Cependant, au delà des échelles musicales utilisées, ces ornements sont une caractéristique importante de l'interprétation qu'il est important d'analyser et de documenter. Par ailleurs, la question de l'évolution des échelles vers le système tempéré peut également être posée dans le cadre des échelles utilisées dans le chant traditionnel.

1.6.2 Les études

Nous répertorions dans cette partie les différents travaux effectués dans le domaine de la recherche automatique d'échelles musicales sous-jacentes à des corpus de chants traditionnels. Si nous énonçons ici plusieurs exemples de travaux, nous devons préciser que les recherches dans ce domaine semblent assez rares.

En 1976, Askenfelt présente un programme de transcription automatique et son application à la musique suédoise (Askenfelt, 1976). Si sa méthode lui semble prometteuse, Askenfelt évoque des développements nécessaires, en particulier en matière de détection de fréquence fondamentale. Cet article semble en tout cas rendre compte de la première tentative de transcription automatique utilisée sur de la musique de tradition orale, ce qui lui procure une dimension historique. Si Askenfelt a été limité par des contraintes techniques, d'autres chercheurs ont su utiliser les améliorations scientifiques pour mettre en place leurs études. Ainsi Simha Arom, qui figure parmi les grands ethnomusicologues du XXe siècle, a-t-il su s'entourer lors de différentes enquêtes de terrain d'une équipe multidisciplinaire, incluant notamment des scientifiques et des chercheurs de l'Ircam. Les études de son équipe portent notamment sur les échelles utilisées en Afrique centrale, dans des régions où les concepts musicaux abstraits sont rarement verbalisés.

Lors d'une première mission effectuée en République Centrafricaine entre 1989 et 1993, Simha Arom et son équipe utilisèrent un synthétiseur Yamaha DX7 pour simuler le son d'un xylophone (Voisin, 1994). Pour s'approcher de conditions de jeux réels, ils utilisèrent des languettes de bois collées sur les touches, ce qui permit aux musiciens de jouer de ce synthétiseur avec des baguettes. Les musiciens pouvaient ainsi tester en condition de jeu les échelles proposées et avaient la possibilité de modifier leurs hauteurs dans les cas où ils les rejetteraient. Les chercheurs ont pu voir au cours de cette étude des accordages très différents acceptés de manière égale par la communauté.

Lors d'une deuxième mission en Afrique centrale, son équipe amena un synthétiseur DX7 pour étudier les échelles utilisées dans les polyphonies Aka (Fürniss, 2006). Ils utilisèrent le synthétiseur pour modifier les échelles enregistrées et employées dans les chants afin de voir si les chants ainsi modifiés étaient ou non acceptés par la communauté. Cette expérience ne fut pas satisfaisante, probablement à cause de moyens techniques limités, qui entraînaient par exemple une synthèse de la voix des chanteurs manquant de réalisme.

En 2007, Simha Arom et son équipe publient une synthèse des méthodes développées pour étudier

les échelles africaines (Arom et al., 2007). Ils affirment à nouveau que la seule façon de valider une hypothèse de recherche est de permettre aux musiciens de modifier les propositions d'échelle qui leur sont faites (Arom, 1991). Cette étude décrit le mode opératoire utilisé dans l'étude des polyphonies pygmées Bedzan ainsi que chez les flûtistes Ouldémé. Ainsi, pour l'étude des chants Bedzan, le système proposé analyse la fréquence fondamentale, puis les intervalles qui séparent les divers degrés de la gamme, afin de transformer les échelles utilisées sans changer les autres paramètres du chant. Les applications utilisées s'appuient sur les logiciels développés par l'Ircam: Audiosculpt et OpenMusic, associés à l'application Scala (Marandola, 2004), et au patch Pareto⁵.

Ces études montrent d'une part que le concept abstrait d'échelles musicales fixes n'est pas significatif pour ces cultures d'Afrique centrale, car il est dépendant d'autres paramètres, comme le timbre. Elles soulèvent d'autre part de nouvelles hypothèses sur la conception des intervalles et des échelles (Pelletier-Ortiz, 1991), (Fernando, 2003). Notons également que Simha Arom et son équipe ont su adapter leurs outils à la population étudiée.

Si nous avons vu que le concept d'échelle fixe n'est pas utilisé en Afrique centrale, il peut être aussi remis en question dans d'autres régions du monde. Ainsi Akkoç, en 2002, s'interroge-il sur la modélisation mathématique des échelles employées dans la musique traditionnelle turque (Akkoç, 2002). Cette étude évoque le caractère non déterministe des notes utilisées par certains maîtres de la musique turque, ce qui s'oppose aux théories conventionnelles utilisant des échelles fixes. Akkoç nous propose le concept de forme de distribution pour caractériser les échelles employées .

D'autres études sont plus adaptées au concept d'échelles fixes. Ambrazevicius publie en 2004 une étude des échelles employées dans le contexte de chant traditionnel monodique de Lituanie. Cette étude prend en compte la variation de tonalité pouvant exister dans le chant, où le chanteur n'a pas de repères harmoniques donnés par un instrument. Une première étape de travail consiste ainsi à reconstruire la tonalité par minimisation de sa variance. L'échelle utilisée peut ensuite être évaluée, et montre l'utilisation d'intervalles de 180 cents en moyenne.

En 2005, deux chercheurs de l'Ircam présentent les échelles pondérées (Carpentier, 2005). Cet article présente un outil qui permet de représenter l'échelle utilisée dans les pièces musicales, afin d'étendre les recherches par contenus dans le cadre du standard MPEG-7. La note fondamentale,

⁵Patchs d'Analyse et de Resynthèse des Echelles dans les musiques de Tradition Orale. Patchs réalisés sur le logiciel Open Music de l'IRCAM par Fabien Levy, avec l'aide technique et les conseils de Jean Bresson, Carlos Agon, Gérard Assayag, Fabrice Marandola, Hans Tutschku et Frédéric Voisin pour une mission ethnomusicologique en juillet 2001.

qui servira de référence à l'échelle, est recherchée dans ce projet comme étant la note la plus importante du morceau, par exemple celle qui apparaît le plus souvent. Les chercheurs utilisent l'algorithme Yin pour extraire la fréquence fondamentale, et le programme attribue un poids à chaque note en fonction de son énergie et de sa durée. Une échelle peut ainsi être construite afin de caractériser un morceau. Le projet n'a malheureusement pas abouti à une implémentation effective pour le standard MPEG-7, mais la méthode décrite reste une source d'inspiration.

En 2008, Emilia Gomez décrit une méthode de transcription automatique de chant flamenco (Gómez, 2008). Cette méthode prend en compte le caractère microtonal du flamenco ainsi que toutes les ornementsations utilisées par les chanteurs. La méthode exploite des caractéristiques de bas niveau, comme l'énergie par bande de fréquence, et la fréquence fondamentale détectée par la méthode d'autocorrélation. Une estimation de la hauteur du morceau est effectuée à l'aide d'une minimisation d'erreurs par rapport à un tempérament égal. Le signal est segmenté en notes courtes par un algorithme de programmation dynamique et la détermination de l'échelle effectivement utilisée est effectuée de manière récursive. Le vibrato est également estimé.

2 Le chant traditionnel du Salento

Le présent stage a fait l'objet d'un travail de terrain de quinze jours dans le Sud de l'Italie. Nous présentons dans cette partie le contexte géographique et culturel de cette région (section 2.1), puis nous détaillerons les problématiques de l'étude de Flavia Gervasi ainsi que nos hypothèses de travail (section 2.2). Nous finirons en exposant le travail réalisé sur le terrain, principalement la mise en place d'un test de perception (section 2.3).

2.1 La musique du Sud de l'Italie

Les régions du Sud de l'Italie possèdent un patrimoine musical très riche. Nous nous intéressons dans cette étude au Salento situé dans le "talon de la botte" italienne, et plus particulièrement à la région de la *Grecia Salentina*. Dans cette petite région d'une dizaine de villages située au cœur du Salento, subsiste un très vieux dialecte proche du grec ancien, le *griko*, alors que le dialecte le plus courant de la région est le *salentin*. On assiste depuis quelques dizaines d'années dans ces régions à un phénomène de revivalisme, qui place la valorisation des traditions locales au cœur d'objectifs touristiques, politiques ou artistiques.



FIG. 2 – Le Salento.



FIG. 3 – La Grecia Salentina.

Cet intérêt manifeste peut s'expliquer dans une certaine mesure par l'attrait exercé par le tarentisme (Caroli, 2009), rituel de transe disparu depuis quelques dizaines d'années. Dans ce rituel, le malade, ou *tarentulé*, était soigné de la morsure de la tarentule au moyen de la danse, de la musique et des couleurs. Ce rituel fut l'objet d'une importante étude de terrain, orchestré par Ernesto de Martino qui constitua à la fin des années cinquante une équipe multidisciplinaire comportant des spécialistes

en histoire des religions, psychiatrie, psychologie, anthropologie et ethnomusicologie (De Martino, 1966).

La musique traditionnelle du Salento est ainsi étudiée le plus souvent dans les ouvrages relevant du tarentisme ou de la transe. La musique liée au tarentisme est pourtant le plus souvent instrumentale (Rouget, 1980, p. 158). Il existe néanmoins, en marge de la musique destinée au rituel du tarentisme, d'autres musiques traditionnelles, et notamment des chants provenant de la culture paysanne, comme les chants de travail ou les berceuses.

Ces derniers chants ont été enregistrés en parallèle du phénomène du tarentisme, ou comme en étant une composante. Les premiers enregistrements remontent aux enquêtes d'Alan Lomax (1954) dans les années cinquante, suivis par les enregistrements de l'équipe de De Martino. De nombreux autres disques de musique traditionnelle ont été édités, sans présenter systématiquement la rigueur méthodologique nécessaire à un travail scientifique. Ces enregistrements, de qualité très variable, constituent néanmoins une source d'information intéressante. Toutefois, même si de nombreux documents sonores présentent des transcriptions, il existe peu d'études approfondies sur les caractéristiques du chant du Salento.

2.2 Etude comparative du chant traditionnel du Salento

2.2.1 Contexte général

La problématique

Le présent projet de stage est lié aux recherches effectuées par Flavia Gervasi dans le cadre de son doctorat en ethnomusicologie à l'Université de Montréal. Ces recherches ont pour thème l'esthétique du chant traditionnel du Salento et du phénomène de revivalisme qui le caractérise. Elles consistent notamment à comparer la tradition portée par les anciens chanteurs au renouveau proposé par les plus jeunes. L'hypothèse sous-jacente est que le chant et ses pratiques d'exécution aient sensiblement évolué ces dernières décennies, en dépit du fait que les jeunes chanteurs revendiquent leur statut de *chanteur traditionnel* et soient identifiés comme tels. C'est dans ce cadre qu'une étude comparative sur le timbre a été menée en 2008, à partir d'enregistrements effectués en studio par deux générations de chanteurs (Gervasi, 2010b),(Le Bagousse, 2008).

Le travail de terrain

Lors d'un récent travail de terrain, Flavia Gervasi a effectué une analyse interactive auprès de chanteurs du Salento. Une étape concrète résidait dans l'organisation de *séances d'apprentissage* où sont confrontés anciens et jeunes chanteurs, afin de comparer leurs interprétations d'un même chant. Le but de ces séances est d'induire les anciens chanteurs à formuler leurs processus de production sonore, ou stratégies opérationnelles, qui en dehors de ces cours de chants resteraient complètement implicites. Ces interactions entre chanteurs présentent ainsi un potentiel très important pour caractériser le style et l'esthétique du chant salentin de tradition orale (Gervasi, 2008). Les anciens chanteurs n'utilisent pourtant pas les notions théoriques de la musique occidentale, ce qui pose un problème de vocabulaire. De plus, ils portent le plus souvent un avis global sur un chant, difficilement exploitable pour mener une étude précise. Il est donc nécessaire de caractériser les chants par des paramètres isolés et de tester la pertinence de ces paramètres (*méthodologie décrite dans la section 1.5.3*).

Les chanteurs participant à la recherche

A Martano vit Cosimino Chiriatti, né en 1926, considéré par la communauté comme une référence locale du chant traditionnel. Cosimino est une figure emblématique du village, qui est par exemple sollicité implicitement par la communauté pour chanter pendant la fête de Pâques. Ce chanteur issu de la culture paysanne avait déjà été enregistré en studio en 2008. Son chant de prédilection est celui des charretiers de jadis, *l'Aria del trainieri*.

Lors des séances d'apprentissage, Cosimino est souvent accompagné d'un autre chanteur, Antonio Costantini, né en 1939, qui s'est souvent avéré précieux pour verbaliser les aspects théoriques du chant.

Nous avons également travaillé avec Lucia De Pascalis, née en 1924, qui chante encore tous les jours. Nous possédons également des enregistrements plus anciens de Lucia De Pascalis, effectués entre les années 1970 et 2000 par son fils Luigi Chiriatti, ce dernier ayant édité plusieurs disques de musique traditionnelle.

Par ailleurs, Flavia Gervasi a eu l'occasion de travailler avec de nombreux chanteurs du revivalisme. Parmi eux, Anna Cinzia Villani et Enza Pagliara sont deux chanteuses reconnues qui participent aux séances d'apprentissage. Deux jeunes chanteurs ont également été enregistrés pendant l'étude

de 2008, Emanuele Licci et Antonio Castrignano.

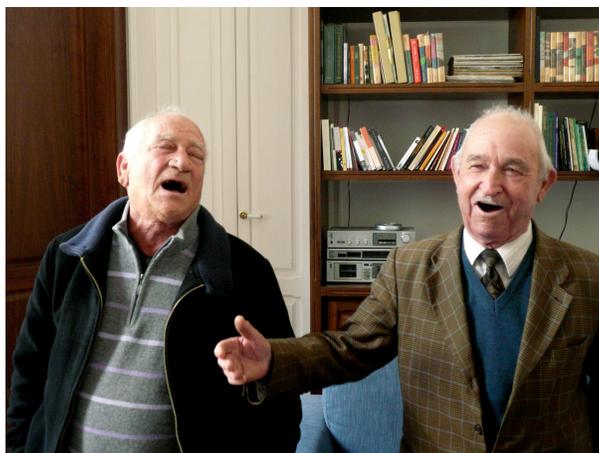


FIG. 4 – De gauche à droite : Antonio Costantini et Cosimino Chiriatti.
"Chant d'accueil pour le français" effectué lors de notre première rencontre.

2.2.2 Les hauteurs de notes

Notre travail collaboratif s'est concentré sur l'analyse précise des hauteurs de notes, des modes et des échelles musicales utilisées. Selon Flavia Gervasi, les intervalles présents dans les chants du répertoire seraient assez éloignés de ceux du tempérament égal. Nous avons commencé à travailler sur cette hypothèse en étudiant dans un premier temps les échelles musicales utilisées.

Méthodologies et problématiques relatives à l'analyse des échelles musicales

Les échelles musicales peuvent être étudiées d'une façon globale dans tout le Salento. Cette analyse peut aussi être précisée en considérant une partie de cette région, comme la Grecia Salentina, un village particulier ou encore en s'intéressant à un chanteur donné. Elle peut enfin s'effectuer de manière très locale en s'attachant à un chant particulier, voire à une partie précise de ce chant.

Néanmoins, les chanteurs susceptibles aujourd'hui de représenter cette tradition sont peu nombreux et assez âgés. Ces deux facteurs compliquent sensiblement l'étude des systématiques musicales propres à ce répertoire. D'une part, le manque de chanteurs représentatifs pourrait nous amener à considérer comme générales des caractéristiques individuelles, et d'autre part, l'âge de certains

chanteurs, malgré le respect que la communauté locale et nous-mêmes leur portons, pourrait être facteur de modifications importantes du chant et fausser ainsi l'étude.

Par ailleurs, selon une hypothèse personnelle, les divers enregistrements collectés depuis les années cinquante peuvent nous aider à effectuer une étude des systématiques utilisées. Ces enregistrements ayant été effectués dans des conditions très différentes, il est difficile de les utiliser dans une étude sur la qualité de la voix. Par contre, l'analyse de la fréquence fondamentale est assez robuste pour ne pas être perturbée de manière significative par les variations de condition d'enregistrement. Ces enregistrements constituent donc une source complémentaire d'information intéressante.

D'autre part, rappelons que le but de cette démarche n'est pas l'établissement formel d'un système musical, tel une échelle musicale précise, qu'utiliseraient implicitement tous les chanteurs. Notre objectif est plutôt de déterminer l'existence de systématiques liées aux hauteurs de notes et aux échelles utilisées. De plus, l'établissement d'une échelle par la seule utilisation des enregistrements serait une démarche réductrice par rapport à une approche ethnomusicologique qui vise à tester systématiquement la pertinence des systématiques étudiées (Gervasi, 2010a). Le but de cette analyse est donc plutôt d'établir des hypothèses de travail qui devront par la suite être validées sur le terrain par une approche interactive.

Le geste vocal

Le travail de terrain de Flavia Gervasi révèle qu'au delà des hauteurs précises de notes, il est important de s'attacher également aux attaques, aux inflexions, aux micro-intonations, et à toutes les ornementsations, que nous désignerons d'une façon globale par *geste vocal*. Ce geste peut être considéré au niveau *macro-structurel*, par l'analyse de la mélodie. Nous pouvons ainsi considérer le découpage en strophes, et le contour de la mélodie au sein de ces strophes. Par ailleurs, le geste peut être étudié de manière plus précise, au niveau *micro-structurel*, en considérant par exemple le type d'inflexion produit lors de l'attaque d'une note.

Il est complémentaire, au-delà du geste vocal utilisé concrètement par le chanteur, d'étudier son intention. Par exemple, un mouvement ascendant de forte amplitude pourrait, pour les chanteurs, mieux caractériser un extrait que les notes atteintes lors de ce mouvement. D'autre part, le facteur contextuel doit être pris en considération. L'environnement du chanteur au moment de l'enregistrement est un facteur important de modification de ses intentions et donc des chants.

2.3 Le travail de terrain réalisé dans le cadre du stage

Dans le cadre de ce stage, nous avons travaillé à Martano avec le logiciel Melodyne qui nous a permis d'effectuer des analyses de fréquence fondamentale, de hauteurs de note, et du geste vocal. Dans un second temps, nous avons modifié les hauteurs des notes pour mettre en place un test de perception. Nous avons cependant relevé des insuffisances du logiciel par rapport à notre étude. Par exemple, Melodyne ne permet pas l'affichage simultané des fréquences fondamentales de différents extraits, ce qui aurait permis une comparaison pratique et rapide. De plus, le logiciel ne détermine pas les échelles utilisées.

2.3.1 Test de perception des hauteurs

Création du test

Le test s'est basé sur différentes versions de *L'Aria del trainieri*, interprété par Cosimino. En étudiant ce morceau, nous avons remarqué les caractéristiques suivantes :

- c'est un morceau en mode mineur;
- les notes pivots, c'est-à-dire les notes les plus longues et les plus intenses autour desquelles s'articule le chant, sont la tonique, la tierce mineure et la quinte;
- la septième oscille entre une septième majeure et une septième mineure;
- la seconde est plutôt mineure, ce qui nous a orientés sur l'hypothèse de l'utilisation d'un mode phrygien.

A l'aide du logiciel Melodyne et à partir de quatre enregistrements différents de strophes du même morceau, nous avons tout d'abord effectué un réajustement par rapport à la gamme tempérée, et ensuite plusieurs modifications plus importantes vers :

1. le mode mineur naturel
2. le mode mineur harmonique

3. le mode mineur mélodique
4. le mode majeur
5. une cadence parfaite

Nous avons ainsi créé dix extraits à partir des quatre originaux. Ces extraits ont été répartis dans un test, contenant des extraits originaux et d'autres modifiés, certains extraits étant répétés. Le test était divisé en trois parties. Les deux premières parties s'appuyaient respectivement sur les deux questions suivantes :

- *Aimez-vous cet extrait ?*
- *Cet extrait ressemble-t-il au chant que vous connaissez ?*

A chacune de ces questions, et pour chaque extrait, le participant devait choisir une case à cocher, les réponses allant de "non" à "beaucoup".

	NO	NON TROPPO	UN PO'	Si	MOLTO
TRACCIA 1					

FIG. 5 – Les cases à cocher du test de perception.

Le troisième test était comparatif. A chaque écoute, deux versions de la même strophe, une originale et une modifiée, étaient proposées. Les participants étaient priés de commenter les versions qu'ils entendaient et de choisir la version qu'ils préféraient. Les réactions des participants étaient enregistrées.

Mode opératoire et résultats

Les extraits étaient lancés à la demande depuis un ordinateur et diffusés sur un petit haut-parleur. Les participants étaient accompagnés d'un expérimentateur pour les aider à remplir leur grille d'évaluation.

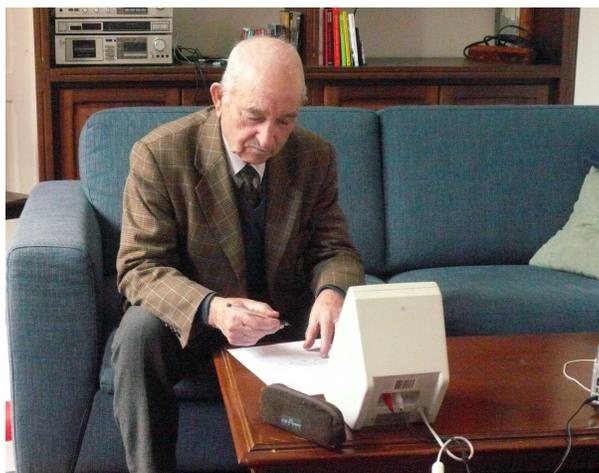


FIG. 6 – Cosimino en train de passer le test.

Trois participants ont effectué le test. Il s'agit d'Antonio, Cosimino et Lucia, cette dernière ayant seulement participé à la partie comparative.

L'analyse des tests a été effectuée de manière détaillée. Voici une synthèse des résultats. Pour Cosimino on peut noter une très grande cohérence dans l'évaluation des extraits. En général, les meilleures appréciations correspondent aux extraits originaux. D'autres jugements nous donnent une piste de travail intéressante, sur l'utilisation de la 7e majeure par rapport à la 7e mineure, intervalles tout deux utilisés dans les enregistrements de Cosimino. Par ailleurs, les extraits transformés en mode mineur mélodique ou majeur sont ceux qui ont été jugés le plus négativement. Pour Antonio, les résultats sont difficilement exploitables car, mis à part quelques extraits très fortement modifiés, Antonio a uniquement utilisé les cases "oui". Lucia n'a en général pas perçu la modification des extraits.

Interprétation et perspectives

La participation de Cosimino nous a ouvert des perspectives intéressantes, comme la possibilité d'affiner la recherche sur les échelles et les modes. Soulignons que pour ce chanteur peu habitué aux technologies musicales, entendre sa voix modifiée artificiellement par le logiciel semble avoir été une expérience troublante.

Pour Antonio, les résultats sont plus difficilement exploitables. Nous supposons qu'entendre la

voix de Cosimino soit déjà un indice de qualité, ce qui oriente ses appréciations sur une base plutôt positive. Il en va de même pour les réponses de Lucia aux comparaisons auxquelles nous l'avons soumise. Ainsi les participants, y compris Cosimino, peuvent situer leurs critères de perception au-delà du système musical, et s'attacher par exemple à la diction ou à la qualité de la voix, ce qui n'est pas pertinent pour ce test.

Dans l'objectif de la poursuite d'un travail avec ce test perceptif, je proposais à Flavia Gervasi les deux directions de travail suivantes. Pour donner des résultats exploitables, ce test devrait en effet être généralisé à un plus grand nombre de participants. Une perspective pourrait être de le faire passer aux jeunes chanteurs du revivalisme, sachant que leurs critères d'acceptation puissent être totalement différents de ceux des anciens chanteurs. Une autre approche pourrait être de généraliser ce test à des personnes de la génération des anciens chanteurs, qu'ils soient eux-mêmes chanteurs ou non. Par ailleurs, dans le cas de la poursuite de tests, il serait souhaitable de définir les conditions d'expérimentation plus rigoureusement, par exemple en limitant le biais induit par la présence des expérimentateurs. Toutefois, certains participants étant assez âgés, il est difficile d'imaginer un protocole de test qui laisserait le participant totalement autonome.

2.3.2 Mise en place d'une base de données

Nous avons mis en place une base de données, contenant la plupart des enregistrements édités ayant été effectués dans le Salento. Nous avons ainsi répertorié une cinquantaine d'enregistrements de monodies datant de 1954 à nos jours, provenant de huit documents (collections de disques, disques associés à des livres, etc...). Nous avons classé ces enregistrements dans une base de données assez conséquente, qui utilise les principaux champs suivants :

- Nom de l'extrait
- Nom du chant
- Nom du chanteur
- Age du chanteur
- Langue

- Année d'enregistrement
- Lieu d'enregistrement
- Présence de polyphonie
- Présence de bruit

Il faut noter que de nombreux enregistrements comprennent des passages polyphoniques. D'autres présentent également des bruits (petites percussions, battements du pied ou de la main, bruit de charrette, etc). De plus, assez peu d'enregistrements sont chantés en griko, la plupart étant chantés en salentin. Nous avons complété cette base de données au fur et à mesure de notre étude en y insérant des informations extraites sur les échelles utilisées dans les chants.

3 Réalisation d'une application informatique

Le test perceptif mis en place sur le terrain dans le Salento nous ayant incités à réaliser des études plus approfondies sur les hauteurs de notes, nous avons décidé de concevoir une application informatique pour effectuer des analyses plus poussées. Cette application devrait être utilisée par Flavia Gervasi pour lui permettre d'analyser les échelles musicales des chants du corpus, ainsi que les gestes vocaux caractéristiques des chanteurs. Nous décrivons dans ce chapitre les besoins et les stratégies choisies pour implémenter cette application (section 3.1), puis nous détaillons plus précisément les algorithmes implémentés pour l'extraction automatique d'information (section 3.2) et les outils d'analyse (section 3.3). Nous décrivons ensuite l'interface graphique développée pour faciliter l'utilisation des outils (3.4), et présentons finalement les premiers résultats découlant des analyses (section 3.5).

3.1 Description du projet

Les objectifs principaux de cette application étaient d'une part d'extraire automatiquement l'échelle musicale utilisée dans un chant et d'autre part de donner la possibilité d'étudier de façon précise et intuitive le geste vocal du chanteur. Ces deux fonctionnalités impliquent une étape préalable d'extraction de caractéristiques des chants telles que la fréquence fondamentale. L'étape d'analyse du morceau par l'utilisateur, incluant la visualisation de l'échelle musicale, l'étude du geste et d'autres outils que nous avons développés, intervient donc dans un second temps. Nous pouvons donc décrire notre application en deux phases distinctes : une phase d'extraction des caractéristiques suivie d'une phase d'analyse et de comparaison.

Par ailleurs, les chants de notre corpus sont le plus souvent caractérisés par une succession de strophes dont la mélodie est quasiment identique. Nous avons fait le choix d'étudier les strophes indépendamment, en supposant que l'analyse d'une strophe est suffisante pour nous donner des informations caractéristiques du morceau, telles que l'échelle employée. Nous avons d'autre part choisi de ne pas implémenter les fonctions d'édition, comme le découpage du morceau en strophes par exemple, qui devront être réalisées préalablement à l'aide d'un logiciel d'édition et de montage sonore.

3.1.1 Extraction des caractéristiques

Les principales étapes de la phase d'extraction automatique des caractéristiques du morceau ont été définies de la manière suivante :

- Extraction de la fréquence fondamentale et de l'énergie
- Segmentation des notes
- Reconnaissance éventuelle de la tonique

Les différentes données extraites sont stockées dans un fichier après chaque étape, ce qui permet de ne les extraire qu'une seule fois. Nous avons par ailleurs opté pour une approche semi-automatique. L'utilisateur peut ainsi choisir d'analyser un morceau automatiquement et dans ce cas toutes les étapes sont réalisées de manière successive, les algorithmes étant configurés avec des paramètres par défaut. Il peut également choisir de recommencer l'analyse, totalement ou partiellement, en modifiant chacun des paramètres utilisés ou même en appliquant des filtres optionnels permettant d'améliorer l'extraction. Ces deux possibilités permettent d'une part le choix d'une analyse simple et rapide et d'autre part la minimisation des erreurs d'extraction par l'utilisation de traitements adaptés au morceau étudié.

3.1.2 Analyse et comparaison

La seconde phase consiste donc pour l'utilisateur à analyser et comparer des morceaux, le logiciel travaillant directement sur les données extraites et stockées dans des fichiers. Outre l'affichage de la fréquence fondamentale et des notes extraites, les outils suivants ont été mis en place :

- Affichage d'histogrammes d'intervalles permettant de visualiser les échelles et les modes utilisés
- Affichage simultané des fréquences fondamentales de différents extraits
- Affichage des intervalles utilisés

- Lecture audio des notes extraites
- Exportation des données extraites

3.1.3 Choix du langage de développement

La programmation de l'application a été effectuée avec le logiciel Matlab qui offre la souplesse d'un développement expérimental et la possibilité de munir l'application d'une interface graphique. De plus, le logiciel permet la création d'applications exécutables qui peuvent être utilisées sur une autre machine de manière indépendante sans nécessiter l'installation de Matlab. Cette condition était nécessaire pour envoyer mon travail à Flavia Gervasi en Italie et pour permettre éventuellement à d'autres musicologues et ethnomusicologues d'utiliser l'application pour étudier divers corpus vocaux.

3.1.4 Nomenclature des fichiers de données

L'application à réaliser devrait être capable de traiter un grand nombre d'extraits, et d'en stocker les caractéristiques dans des fichiers. Le corpus important et la récurrence de certains chants a posé le problème de la désignation des données. Nous avons ainsi opté pour une nomenclature adaptée à notre corpus, afin d'identifier chaque extrait de façon unique et significative.

Les noms de fichiers possèdent tous le même nombre de caractères et sont composés des champs suivants :

Champs	Taille	Exemple
Profil du chanteur : ancien ou jeune	1	A ou J
Nom du chanteur	3	LUC ou COS ou 000
Collecte : historique ou étude	1	H ou E
Date : année, mois, jour (00 si inconnu)	6	120308
Style du chant	5	STISA ou TRAIN
Début du texte (00000 si pas de texte)	5	FERMA ou ALLE1
Numéro de version	1	1 ou 2
Numéro de strophe	2	01 ou 02 ou 03

exemple de nom : A_LUC_DI_121105_STISA_FERMA_101

3.2 Implémentation de la phase d'extraction des caractéristiques

Nous détaillons dans cette partie les algorithmes implémentés pour l'extraction de caractéristiques acoustiques des chants. Nous pouvons distinguer les caractéristiques de bas niveau telles que la fréquence fondamentale et l'énergie des caractéristiques de plus haut niveau telles que les notes ou la tonique.

3.2.1 L'extraction de la fréquence fondamentale

Au vu des erreurs résultant de l'extraction de la fréquence fondamentale et afin de déterminer les algorithmes correspondant le mieux aux besoins de ce projet, nous avons effectué la comparaison de trois méthodes d'extraction: l'autocorrélation, le Yin, et l'algorithme SHRP. Voici un exemple significatif de comparaison effectuée sur un chant enregistré du corpus. Dans cet exemple, les algorithmes Yin et SHRP ont été calibrés à une bande de détection de fréquence adaptée, entre 150 et 600 Hertz. Pour un extrait de 25 secondes, les méthodes ont nécessité respectivement 13, 10.3, et 6.5 secondes de traitement.

Comme nous le voyons à la figure 7, le résultat de l'autocorrélation, qui était le seul algorithme non borné en fréquence, présente de nombreux sauts d'octave. En effet, les points en dehors du graphique se trouvent pour la plupart au double de la fréquence fondamentale de chant.

Nous pouvons également observer une très forte instabilité de l'algorithme Yin autour de 5 secondes, due à une respiration du chanteur entre deux strophes. Cette instabilité est prévisible et ce passage pourrait être supprimé en fonction de l'énergie du signal. Par contre, nous observons également d'autres instabilités, plus petites, aux alentours de 1.6 et 6.3 secondes. A l'écoute, ces instabilités semblent correspondre à des brusques changements morphologiques d'émission du son. Il est difficile de dire s'ils correspondent effectivement à un très court changement de fréquence fondamentale, ou à des erreurs d'estimation. Il semble en tout cas que, pour notre travail d'analyse sur les hauteurs de notes, il soit préférable de ne pas prendre en compte ces brusques variations.

L'algorithme SHRP donne de bons résultats et semble plus stable que le Yin. On peut toutefois

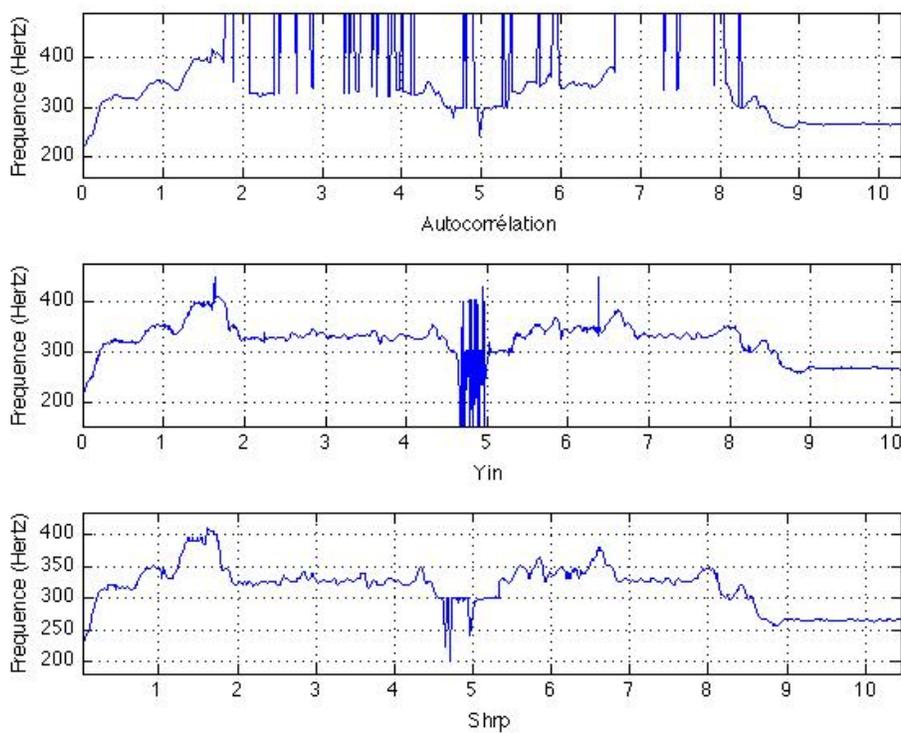


FIG. 7 – Comparaison de trois algorithmes d'extraction de fréquence fondamentale.

regretter le manque de précision au niveau de la fréquence, qui semble échelonnée tous les 50 cents, ce qui n'est pas vraiment suffisant pour effectuer une étude approfondie sur les échelles. Aussi, nous avons choisi d'utiliser le Yin dans l'implémentation actuelle de l'application.

3.2.2 L'extraction de l'énergie

L'extraction de l'énergie a été effectuée de manière à avoir une donnée d'énergie pour chaque échantillon contenu dans le vecteur de fréquence fondamentale (chaque échantillon correspondant à une fenêtre). Cette énergie étant calculée sur une portion du son, elle correspond plutôt à une puissance.

La configuration par défaut du Yin pour un morceau échantillonné à 44100 Hz est :

Taille de la fenêtre : 300 échantillons

Décalage entre deux fenêtres, *hop ou overlap* : 32 échantillons

Ce qui nous donne environ 1400 échantillons d'énergie par seconde.

Nous calculons ainsi la valeur absolue du signal s , puis pour chacune des fenêtres de taille W , la moyenne des échantillons, ce qui nous donne l'énergie e d'un échantillon.

$$e = \sum_{j=1}^W |s(j)|^2 \quad (6)$$

Pour des calculs ultérieurs, il peut être utile d'exprimer l'énergie en décibel e_{dB} normalisée par rapport à l'énergie maximale du signal e_{max} .

$$e_{dB} = 20 \log \left(\frac{e}{e_{max}} \right) \quad (7)$$

3.2.3 Traitements ultérieurs à l'extraction de la fréquence fondamentale

Les algorithmes d'extraction de fréquence fondamentale ne nous donnant pas de résultats directement exploitables, nous avons décidé de proposer des traitements pour améliorer les résultats. Le principe retenu est de passer les points aberrants détectés par les traitements à la valeur *NAN* (*Not A Number*), ce qui permet de ne pas les afficher. Voici les différentes solutions et traitements implémentés :

1. Modifier les paramètres des algorithmes de détection de la fondamentale

Nous pouvons observer à la page 46, figure 4.1, une extraction de fondamentale effectuée par l'algorithme Yin à l'intérieur d'une bande de fréquence de 100 à 700 Hertz. Cette extraction présente de nombreux sauts d'octave, qui peuvent être corrigés en prenant un meilleur calibrage de la bande de fréquence. Ainsi, après une première extraction, il peut être souhaitable

de refaire l'analyse en améliorant les bornes de détection. Les autres paramètres des algorithmes sont également modifiables. On pourrait par exemple rehausser le seuil de détection de minimum du Yin (fixé à 0.1) pour éviter les erreurs de sauts d'octave. Ces paramètres semblent toutefois moins intuitifs que les bornes de la bande de fréquence.

Une nouvelle extraction est observable à la page 47, figure 4.1, effectuée en choisissant une bande de fréquence de 200 à 500 Hertz. Comme nous pouvons le voir, la plupart des sauts d'octave semblent ici supprimés, les discontinuités correspondant le plus souvent à des respirations ou à des attaques. L'algorithme implémenté utilise par défaut une gamme allant de 150 à 600 Hertz.

2. Taux de variation

Un filtrage basé sur le taux de variation a été mis en place. A partir de la valeur absolue de la dérivée du signal s' du signal, nous calculons une moyenne sur une fenêtre de taille W .

$$\Delta = \frac{1}{W} \sum_{j=1}^W |s'(j)| \quad (8)$$

Nous supprimons ensuite les échantillons des fenêtres qui correspondent à une moyenne supérieure à un seuil donné. Notons que la dérivée calculée point par point donne de bien moins bons résultats que le calcul utilisant la moyenne sur une fenêtre glissante.

La courbe de la page 48 nous montre l'effet de ce filtrage. On peut observer en vert, l'affichage de la fondamentale calculée à l'origine par l'algorithme Yin.

Taille de la fenêtre glissante: 10 ms (soit ici 13 échantillons)

Seuil : 50 cents

3. Energie

Un autre filtrage évoqué précédemment consiste à supprimer les points dont l'énergie est inférieure à un seuil. Ici encore, l'utilisation de la moyenne sur une fenêtre glissante donne de meilleurs résultats. Les résultats sont visibles à la figure 4.1. Le seuil est donné en décibels normalisés par rapport à l'énergie maximum du signal.

Taille de la fenêtre : 30 millisecondes

Seuil : -80 dB

4. Energie et taux de variation combinés

Les erreurs de détection étant souvent dues à des silences, soit des moments de basse énergie caractérisés par un taux de variation important, nous avons également mis en place un traitement considérant les fenêtres de basse énergie et de taux de variation important. Toutefois, même avec des seuils plus larges que précédemment, ce traitement semble moins efficace que l'utilisation séparée de ces deux filtres.

5. Filtre médian

Le filtre médian utilisé avec une fenêtre glissante permet de supprimer les points trop éloignés de la majorité des points considérés, et ainsi de lisser la courbe. Comme une taille de fenêtre trop importante supprimera les variations de la courbe (dues au vibrato par exemple), ce filtrage doit être utilisé avec parcimonie. Les résultats sont présentés à la figure 4.1, page 50.

Taille de la fenêtre : 30 millisecondes

6. Suppression d'intervalles

La plupart des extraits du corpus n'ayant pas toujours été enregistrés dans des conditions de studio d'enregistrement, ils présentent souvent des bruits parasites, nuisibles à l'analyse des morceaux. L'utilisateur peut ainsi décider de supprimer certaines parties du chant qu'il juge inutiles. Par exemple cet extrait présente un bruit, vers 13.5 secondes, que nous avons supprimé, de même qu'une courte variation vers les 5 secondes. Le résultat est présenté à la page 51.

Les traitements de filtrage par le taux de variation, de filtrage par l'énergie, et de filtrage médian sont effectués par défaut dans la phase d'extraction automatique. Les seuils et tailles de fenêtre choisis sont ceux présentés dans les exemples décrits.

Traitements non implémentés

Nous aurions pu envisager un algorithme de reconstruction de la courbe par continuité, ce qui aurait amélioré l'affichage des courbes. Une autre amélioration pourrait être apportée par un algorithme de suppression des sauts d'octave. Ces erreurs sont toutefois évitables en calibrant convenablement notre algorithme d'extraction de la fréquence fondamentale, d'autant plus que la plupart des chants du corpus ne couvrent qu'un octave de tessiture.

3.2.4 Segmentation de notes

L'algorithme de segmentation de notes que nous avons développé s'appuie sur les variations de la fréquence fondamentale, ce qui semble le mieux adapté à une analyse des hauteurs de notes. La segmentation de notes repose donc sur une détection des variations importantes de la fréquence fondamentale que nous appellerons *attaques*. Nous supposons que ces attaques séparent les notes qui sont caractérisées par une fréquence fondamentale plus stable.

Une première étape consiste donc à détecter ces attaques. Nous effectuons un calcul de la dérivée s' . Nous calculons ensuite la moyenne de cette dérivée à l'intérieur d'une fenêtre glissante, puis nous prenons la valeur absolue de cette moyenne.

$$\Delta = \frac{1}{W} \left| \sum_{j=1}^W s'(j) \right| \quad (9)$$

Le fait d'utiliser la valeur absolue après la moyenne permet, à la différence du filtrage par le taux de variation, de ne pas considérer comme des attaques les oscillations qui seraient dues par exemple à un vibrato.

L'algorithme utilise par défaut les valeurs suivantes, à une fréquence d'échantillonnage de 44100 Hertz:

Taille de la fenêtre de détection d'attaque : 100 ms

Seuil de détection entre deux échantillons: 0.9 cents

Cette variation moyenne entre deux échantillons correspond, sur la fenêtre entière, à une variation de plus de 140 cents. Les attaques étant détectées, nous pouvons séparer le signal en deux zones : les zones d'attaques, et les autres zones considérées comme des notes tenues. Aux notes tenues est affectée une hauteur qui correspond à la moyenne de la note.

Cette première détection est suivie d'un traitement récursif qui permet d'affiner le résultat. Au-dessous d'une durée minimale, les notes sont supprimées et remplacées par des attaques. Les attaques suffisamment courtes entre deux notes de fréquences proches sont également supprimées. A partir de nouveaux vecteurs d'attaque, nous recalculons les notes. Le processus s'arrête quand il

n'y a plus de modifications effectuées. Les résultats de l'étape d'extraction de notes sont visibles à la figure 4.1, page 51.

Les seuils suivants sont utilisés :

Durée minimale de note : 10 ms

Seuil entre deux notes proches : 50 cents

Durée d'attaques maximale pour considérer deux notes comme proches : 0.3 seconde

La segmentation de notes peut également être l'objet de traitements ultérieurs. L'utilisateur peut en effet choisir de refaire la segmentation, ce qui lui donne accès la paramétrisation des données ainsi qu'à un affichage des différentes courbes d'attaque.

En outre, un outil de *fusion de notes* a été implémenté, ce qui permet à l'utilisateur d'éliminer les attaques entre deux notes. Ainsi les deux notes choisies, ainsi que les éventuelles notes intermédiaires, sont fusionnées en une seule dont la hauteur est recalculée.

3.2.5 L'extraction de la tonique

Une hauteur de référence est nécessaire pour étudier l'échelle musicale utilisée. Considérant le corpus étudié, nous avons choisi d'utiliser la référence de la tonique, facilement détectable à l'oreille, pour calculer les intervalles présents. Nous avons dans un premier temps mis au point une méthode automatique de détection de tonique qui, pour le corpus étudié, présente de bons résultats. Cette détection s'appuie sur la recherche d'une note de forte énergie et de durée importante en fin de strophe. Les chants du corpus présentent en effet pour la plupart une tenue sur la tonique à la fin des strophes. Toutefois, l'étape de détection implémentée actuellement est manuelle. La tonique est soumise au choix de l'utilisateur, ce qui permet plus de souplesse. Le choix s'effectue par une interface graphique.

3.3 Implémentation des outils d'analyse et de comparaison

3.3.1 Histogramme d'intervalles

Dans le but d'analyser les échelles musicales utilisées, nous avons conçu un outil d'affichage d'histogrammes d'intervalles. La figure 4.2 nous indique la hauteur des notes en cents par rapport à la tonique sur l'axe des abscisses, et la durée totale d'occurrence de chaque note sur l'axe des ordonnées. L'histogramme d'intervalles avait été implémenté dans un premier temps à partir du vecteur de fréquence fondamentale. Toutefois, l'histogramme réalisé à partir des notes (qui correspondent à la fréquence fondamentale moyennée sur des parties stables, dont on aurait supprimé les variations trop importantes) donne des résultats plus lisibles.

Il a également été réalisé une implémentation dans laquelle les notes sont pondérées par leur énergie. Ainsi on affectera des poids moins importants aux notes de faible énergie, ce qui va avoir pour effet, au niveau de l'affichage, de diminuer leur temps d'occurrence.

3.3.2 Affichages multiples

Un outil d'affichage multiple a été réalisé afin de permettre une comparaison immédiate des inflexions vocales présentes dans différents extraits. Un exemple est présenté à la figure 4.2 page 54. Nous avons également mis en place une option permettant d'effectuer des translations de courbe, en temps ou en hauteur, afin de procéder à des comparaisons plus précises, par juxtaposition des courbes.

3.3.3 Affichage des intervalles utilisés

L'affichage des intervalles utilisés est un outil très utile pour l'analyse des chants. Cette dernière requerrait en effet de connaître les intervalles entre toutes les notes. Il est particulièrement intéressant d'étudier les petits intervalles présents dans les ornementsations, ou au contraire les sauts importants. Cet outil utilise une interface graphique qui permet à partir d'une note sélectionnée d'afficher les intervalles en cents la séparant des autres notes choisies. Un exemple est présenté à la figure 4.2.

3.3.4 Lecteurs audio

En plus d'un petit lecteur permettant à tout moment d'écouter les chants étudiés, une fonction a été développée pour synthétiser un son à partir des notes reconnues. Si le son de synthèse a été réalisé de façon très basique par synthèse additive, l'utilisation de ce lecteur est très intéressante car elle focalise l'auditeur sur les hauteurs de notes. Ces hauteurs sont en effet les hauteurs absolues des notes reconnues, et ne sont pas quantifiées pour correspondre au tempérament égal.

3.3.5 Exportation

Les données peuvent être exportés dans un fichier texte. Les fréquences fondamentales maximales, minimales, et moyennes y sont indiquées. Les informations concernant les notes sont décrites par fréquence, temps de début de note et durée.

3.4 Interface graphique

Pour des raisons de simplicité et pour permettre une utilisation intuitive, la mise en place d'une interface graphique était nécessaire. Deux des fenêtres réalisées sont présentées à la figure 4.3, page 56.

- Choix des morceaux et des répertoires de travail.
- Fenêtre principale de travail.
- Fenêtre de modification de la fréquence fondamentale.

La fenêtre de modification de la fréquence fondamentale permet de réaliser tous les filtres décrits ci-dessus. Il est également possible via cette interface de revenir à la fréquence fondamentale extraite à l'origine par l'algorithme, ou de refaire l'extraction. Des fonctions d'annulation de la dernière action et de sauvegarde ont également été implémentées.

Nous avons également prévu différents dispositifs d'affichage, ce qui permet par exemple dans le cas de l'extraction de note, de ne visualiser que les notes, ou alors de visualiser également la fréquence fondamentale ou encore les attaques détectées.

3.5 Résultats

3.5.1 Résultats sur les échelles musicales

Les phases d'analyse systématique des échelles musicales des morceaux du corpus et d'interprétation des résultats n'ont pas été effectuées au cours de ce stage. Certaines hypothèses ont néanmoins pu être dégagées, et demandent à être validées par des analyses plus globales. Il apparaît par exemple de manière récurrente des notes situées entre la quarte et la quinte diminuée, à un intervalle d'environ 550 cents de la tonique.

3.5.2 Le geste vocal

La réalisation de cette application informatique constitue un outil efficace pour étudier les chants de façon individuelle. En outre, par rapport à une application commerciale existante comme Melodyne, en plus de proposer des nouveaux traitements et outils plus adaptés à l'ethnomusicologie, notre application donne des résultats plus précis et permet ainsi une analyse du geste musical plus fine. Ainsi le logiciel Melodyne semble moins adapté à l'analyse, car il calcule les hauteurs de notes de façon plus grossière sans prendre en compte les petites ornementsations telles que les trilles, souvent présentes dans le chant du Salento. Ce nouvel outil permet donc un suivi et une comparaison du geste musical bien plus précise qu'avec l'utilisation de Melodyne, comme le montre la figure comparative, à la page 57.

3.5.3 Utilisations alternatives de l'application

La présente application a par ailleurs été utilisée par Roxane Campeau et Mariève Corbeil du BRAMS ⁶ de Montréal lors d'un travail de recherche sur le chant maternel. Notre application a permis d'analyser un corpus de chants maternels, et d'en extraire les caractéristiques principales. L'outil d'exportation a d'ailleurs été développé comme un besoin de cette étude. Selon Roxane Campeau, la prise en main rapide et simple de l'application a permis d'effectuer un travail précis et efficace.

⁶BRAMS : International Laboratory for Brain, Music and Sound Research. Université de Montréal / Université McGill

Ainsi, l'application réalisée, même si elle a été conçue par rapport à une problématique très précise, peut être utilisée pour d'autres études, et particulièrement dans le domaine de l'ethnomusicologie. Un manuel d'utilisation a été rédigé dont nous présentons des extraits en annexe, page 59.

3.6 Perspectives de développement

3.6.1 Détection de similarités

Dans l'objectif de comparer des chants, nous aurions souhaité mettre en place une fonction de détection automatique de similarité. Une application de cette fonctionnalité pourrait par exemple être la comparaison automatique de deux strophes d'un même chant exécuté par des interprètes différents. Cette fonction aurait pu s'appuyer sur des algorithmes de programmation dynamique afin de faire correspondre au mieux chacune des notes des deux chants. Ainsi, les différences de hauteur entre deux chants ainsi que les notes supplémentaires, telles que celles présentes dans les ornements, auraient pu être repérées automatiquement. Cette fonction de détection de similarités aurait également pu être utilisée pour découper automatiquement un morceau en strophes similaires.

3.6.2 Détection et élimination de la déviation

Dans les corpus traditionnels où l'interprète chante a capella, nous pouvons assister à un phénomène de déviation de hauteur du morceau. En effet, même si le morceau est composé de strophes similaires, la note de référence d'une strophe, ou tonique, dévie au fur et à mesure des strophes. Le chanteur peut ainsi garder la même échelle localement à l'intérieur d'une strophe, mais changer de référence au cours du morceau. Certains chants du corpus sont représentatifs de cet effet. L'utilisation de notre application permet de limiter fortement cet effet en se limitant à l'analyse de strophes indépendantes. Toutefois, afin d'étudier les morceaux en entier, il aurait été intéressant d'implémenter un algorithme capable de corriger cette déviation. Cet algorithme s'appuierait sur une recherche de similarité entre strophes (découpées de manière automatique ou manuelle) et effectuerait des traitements (typiquement transposer la strophe de quelques dizaines de cents).

3.6.3 Étirement du temps

La fonction d’affichage multiple permet de comparer différents extraits. La lecture de ces comparaisons est parfois rendue difficile par les vitesses différentes auxquelles sont exécutés les chants. Un algorithme d’étirement du temps, plus connu en anglais sous le nom de *time-stretching*, permettrait de ramener deux morceaux exécutés à des vitesses différentes à une vitesse commune. La comparaison de différents chants pourrait ainsi être facilitée.

3.6.4 Prolongation par continuité

La plupart des courbes de fréquence fondamentale présentent des discontinuités, dues le plus souvent à de mauvaises estimations remplacées par des points sans valeurs (NaN). La prolongation de ces courbes par continuité pourrait en améliorer la lecture, lors de l’utilisation de la fonction d’affichage multiple.

3.6.5 Modification du diapason

Pour une meilleure visualisation des notes chantées, nous aurions pu proposer une modification optionnelle du diapason utilisé. Cette modification pourrait à l’affichage, par un décalage de la grille correspondant au tempérament égal, permettre d’aligner certaines des notes chantées avec la grille afin d’améliorer la lecture.

3.7 Discussion sur le développement

Si on considère la perspective de l’utilisation régulière d’un tel outil par différents ethnomusicologues, nous sommes amenés à considérer la reprogrammation de l’application dans un autre langage informatique. Le langage Matlab, en plus de d’être un langage propriétaire qui induit un coût important, est en effet plus orienté sur la recherche que sur le développement informatique. Par exemple, malgré un effort particulier effectué sur ce point, les applications exécutables que nous avons créées manquent de robustesse et ne peuvent être portées d’un système d’exploitation à un autre. Nous pourrions par exemple envisager un développement en langage C, qui offrirait en

outre de meilleures performances en terme de rapidité de traitement.

Par ailleurs le temps nécessaire à l'implémentation d'une détection de fréquence fondamentale robuste s'est révélé conséquent, ce qui nous a laissé moins de temps pour réaliser des traitements de plus haut niveau. Une approche alternative aurait été l'implémentation d'un sous-programme (ou patch en anglais) s'appuyant sur une application existante, comme par exemple le logiciel Audiosculpt. Le recours à un système efficace et validé d'extraction de la fréquence fondamentale, aurait éventuellement pu nous faire économiser un temps précieux et nécessaire à d'autres développements. Toutefois, la réalisation d'une application autonome présente d'autres avantages indéniables, comme la simplicité d'installation, d'utilisation ou le coût.

Conclusion

Parcours et résultats obtenus

Ce stage de recherche m'a permis, dans un premier temps, à partir d'un travail collaboratif, de contribuer à la mise en place d'un test de perception. Les résultats de ce test m'ont ensuite conduit à la réalisation d'une application informatique dans le but d'approfondir nos recherches. Cette application devait utiliser des algorithmes d'extraction automatique d'information musicale afin de permettre l'analyse et la comparaison des échelles musicales utilisées, ainsi que du geste vocal.

La réalisation de cette application s'est déroulée de manière progressive, chaque étape étant validée par des tests sur les extraits du corpus. Ce processus m'a amené à implémenter différentes options et à choisir les paramètres de manière empirique. L'application finalement réalisée permet d'une part une grande simplicité d'utilisation, et d'autre part des analyses robustes et précises.

Grâce aux différents outils réalisés, notre application informatique présente donc un intérêt évident pour l'étude des échelles musicales et du geste vocal des chanteurs. Les premiers résultats semblent ainsi tout à fait prometteurs. Nous espérons que son utilisation pourra enrichir de manière significative l'étude sur le chant traditionnel du Salento, et peut-être d'autres recherches.

Apport personnel

L'approche interdisciplinaire entre ethnomusicologie et informatique m'a paru passionnante. J'ai travaillé en effet, au cours de cette étude, sur des problématiques liées aux hauteurs de notes dans des perspectives historiques, musicologiques, perceptives et techniques.

Je me suis d'une part initié à l'étude perceptive par la mise en place d'un test, de sa réalisation et l'étude des résultats, ce qui m'a donné, malgré le nombre restreint de participants, une vision globale de la méthodologie et une première approche de recherche dans le domaine de la perception.

D'autre part, j'ai pu dans la majeure partie de ce stage, perfectionner ma pratique du traitement du signal et du développement informatique, notamment avec le logiciel Matlab. J'ai ainsi conçu et réalisé des fonctions de bas et de haut niveau, des interfaces graphiques et des applications externes. La réalisation d'une application autonome m'a permis d'avoir une vision d'ensemble de la réalisation d'un logiciel de traitement audio. Par ailleurs, je me suis initié durant ce stage à l'utilisation d'autres logiciels, comme Audiosculpt et OpenMusic.

J'espère ainsi que ces compétences théoriques et techniques et cet attrait pour l'interdisciplinarité pourront s'avérer utiles pour d'autres travaux dans le domaine de la recherche.

4 Annexes

4.1 Extraction de la fréquence fondamentale et traitements ultérieurs

Les graphiques suivants sont issus de l'analyse de *l'Aria del trainieri* interprété par Cosimino Chiriacchi.

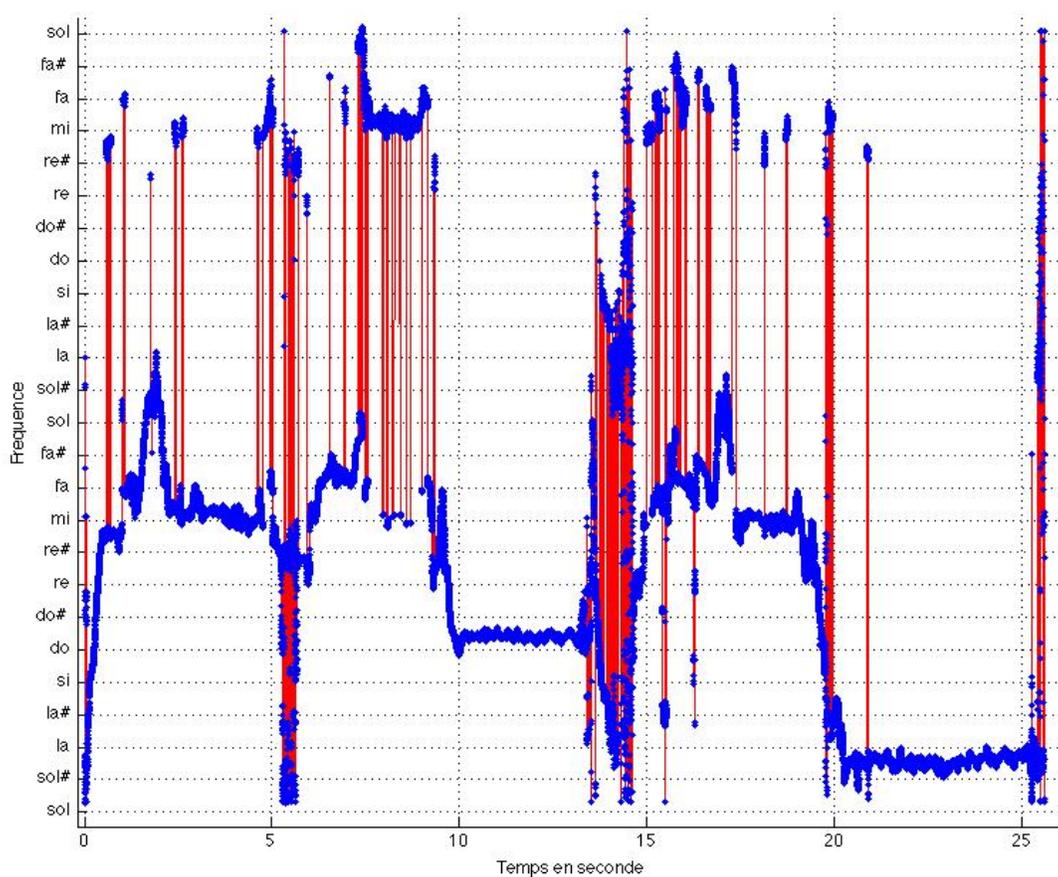


FIG. 8 – Extraction de la fréquence fondamentale par l'algorithme YIN.

Paramètres

Bande de fréquence : de 100 à 700 Hertz

En rouge: la courbe de fréquence fondamentale.

En bleu: les points du vecteur de la fréquence fondamentale.

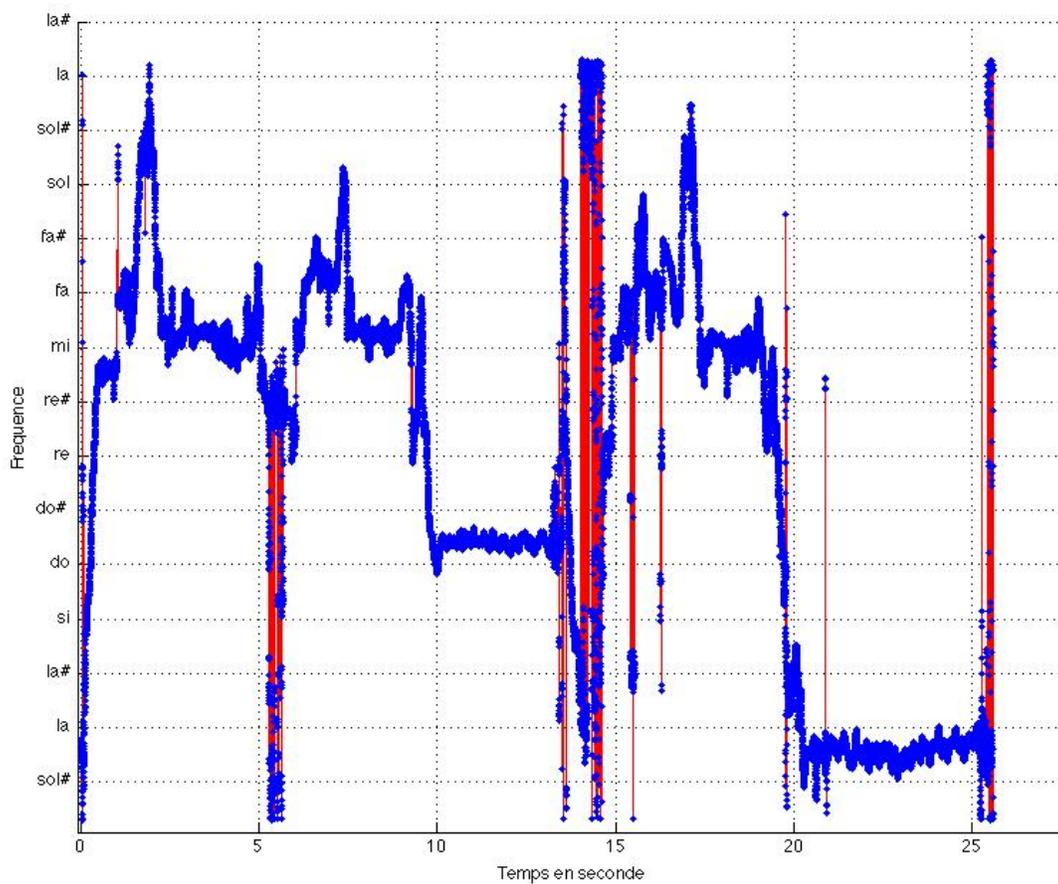


FIG. 9 – Extraction de la fréquence fondamentale par l’algorithme YIN.

Paramètres

Bande de fréquence : 200 à 500 Hertz

En rouge: la courbe de fréquence fondamentale.
En bleu: les points du vecteur de la fréquence fondamentale.

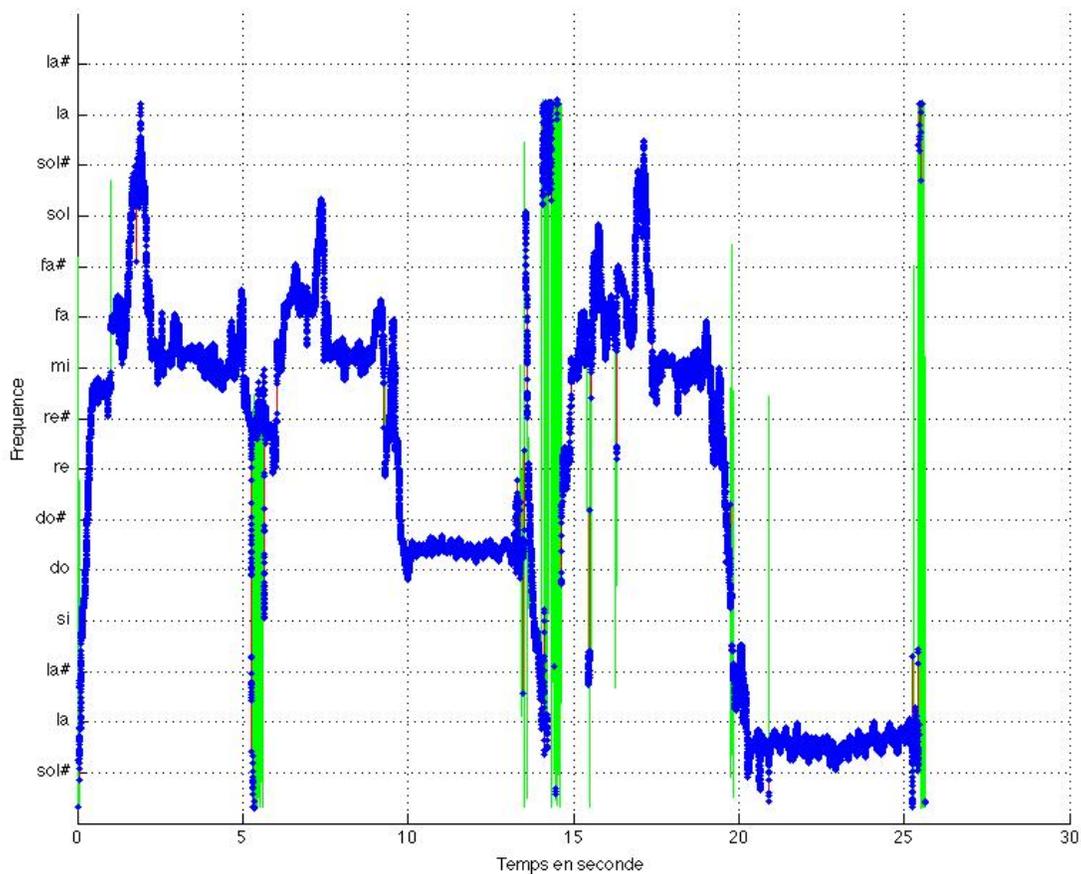


FIG. 10 – Utilisation du filtrage par le taux de variation.

Paramètres

Taille de la fenêtre : 10 ms

Seuil : 50 cents

En rouge: la courbe de fréquence fondamentale.
En bleu: les points du vecteur de la fréquence fondamentale.
En vert: la fréquence fondamentale d'origine issue du YIN.

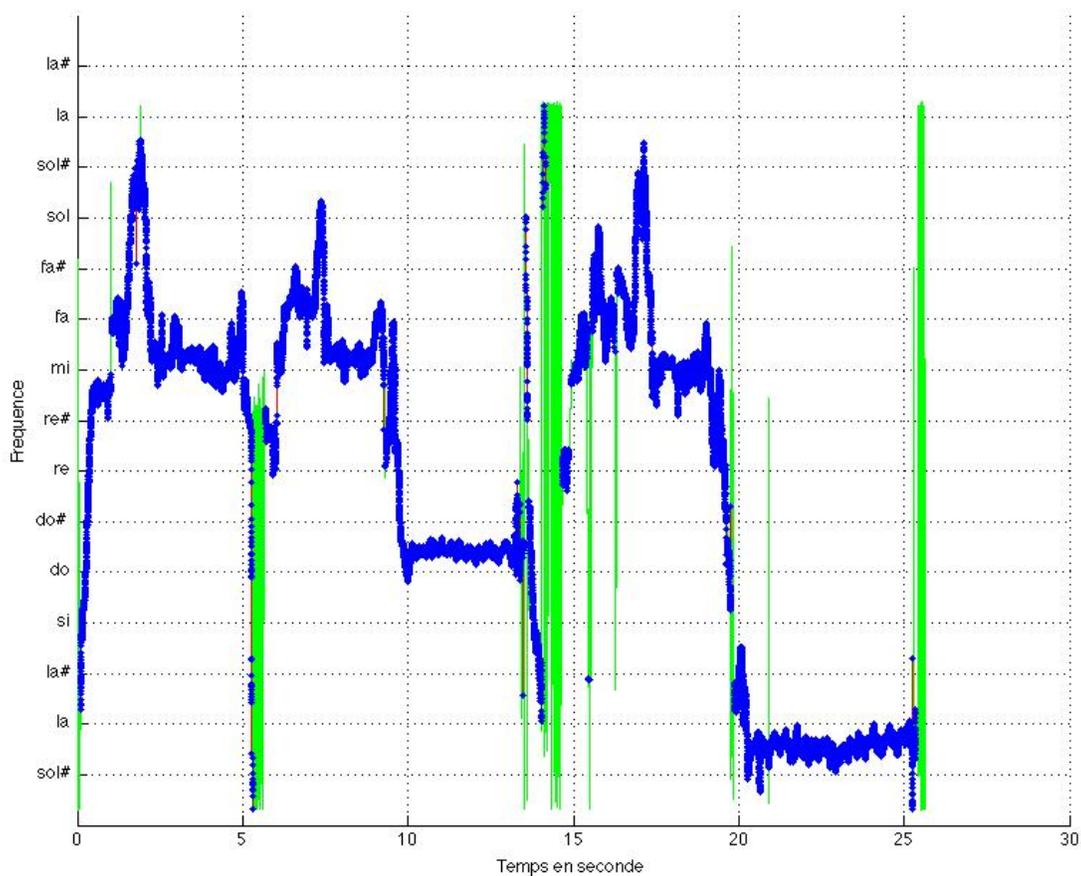


FIG. 11 – Utilisation du filtrage par l'énergie.

Paramètres

Taille de la fenêtre : 30 ms

Seuil : -80 dB

En rouge: la courbe de fréquence fondamentale.
En bleu: les points du vecteur de fréquence fondamentale.
En vert: la fréquence fondamentale d'origine issue du YIN.

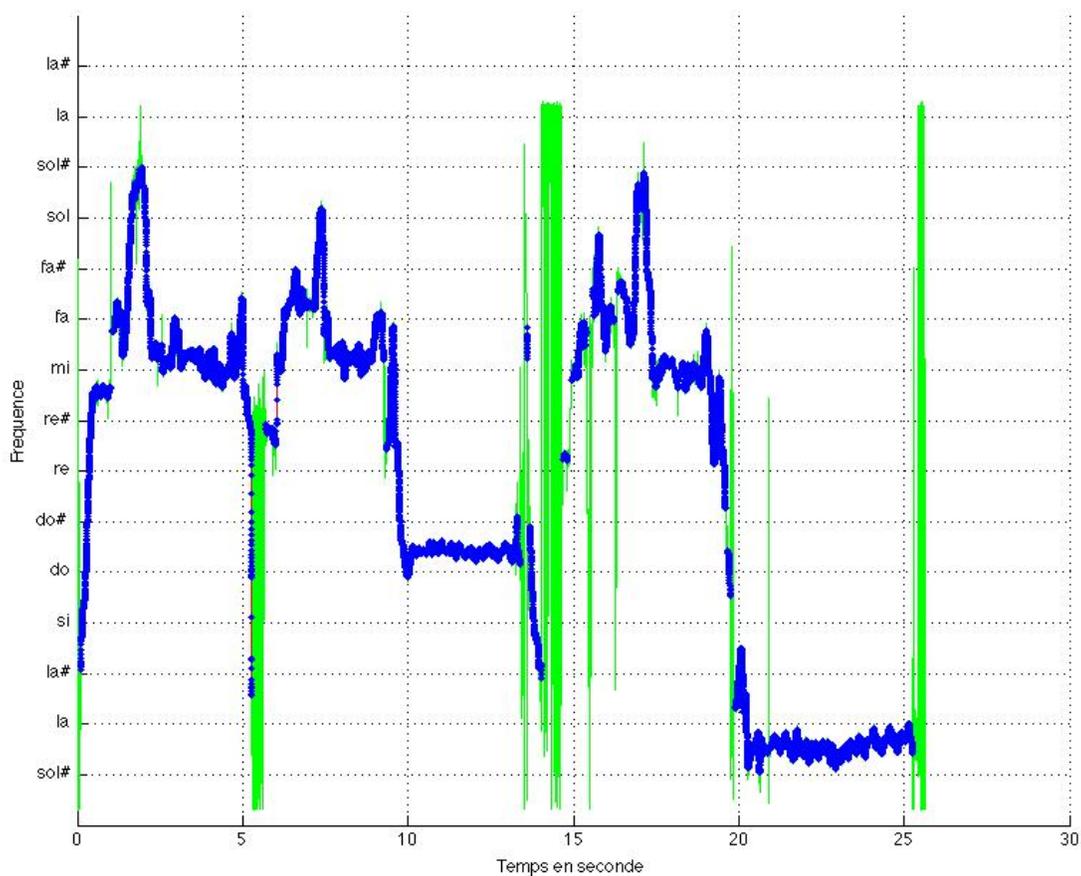


FIG. 12 – Utilisation du filtrage médian.

Paramètres

Taille de la fenêtre : 30 ms

En rouge: la courbe de fréquence fondamentale.
En bleu: les points du vecteur de la fréquence fondamentale.
En vert: la fréquence fondamentale d'origine issue du YIN.

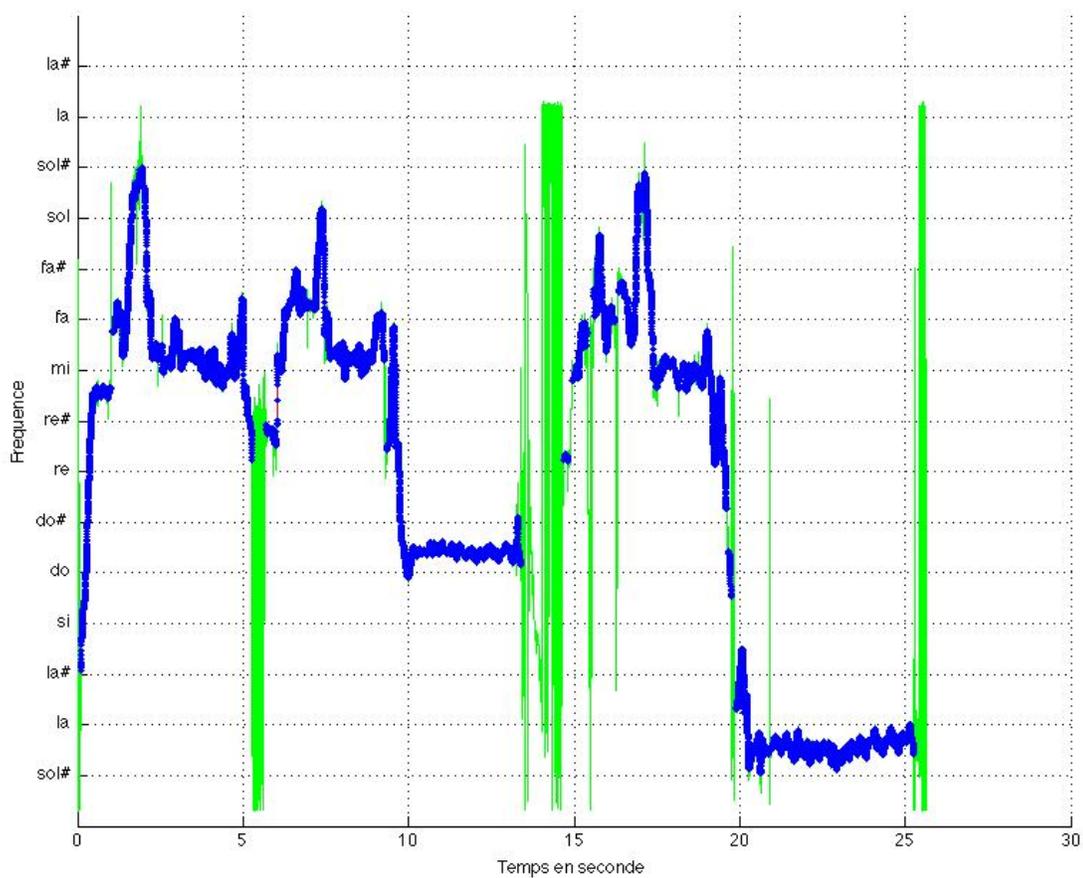


FIG. 13 – Suppression d’intervalles.

En rouge: la courbe de fréquence fondamentale.
En bleu: les points du vecteur de fréquence fondamentale.
En vert: la fréquence fondamentale d’origine issue du YIN.

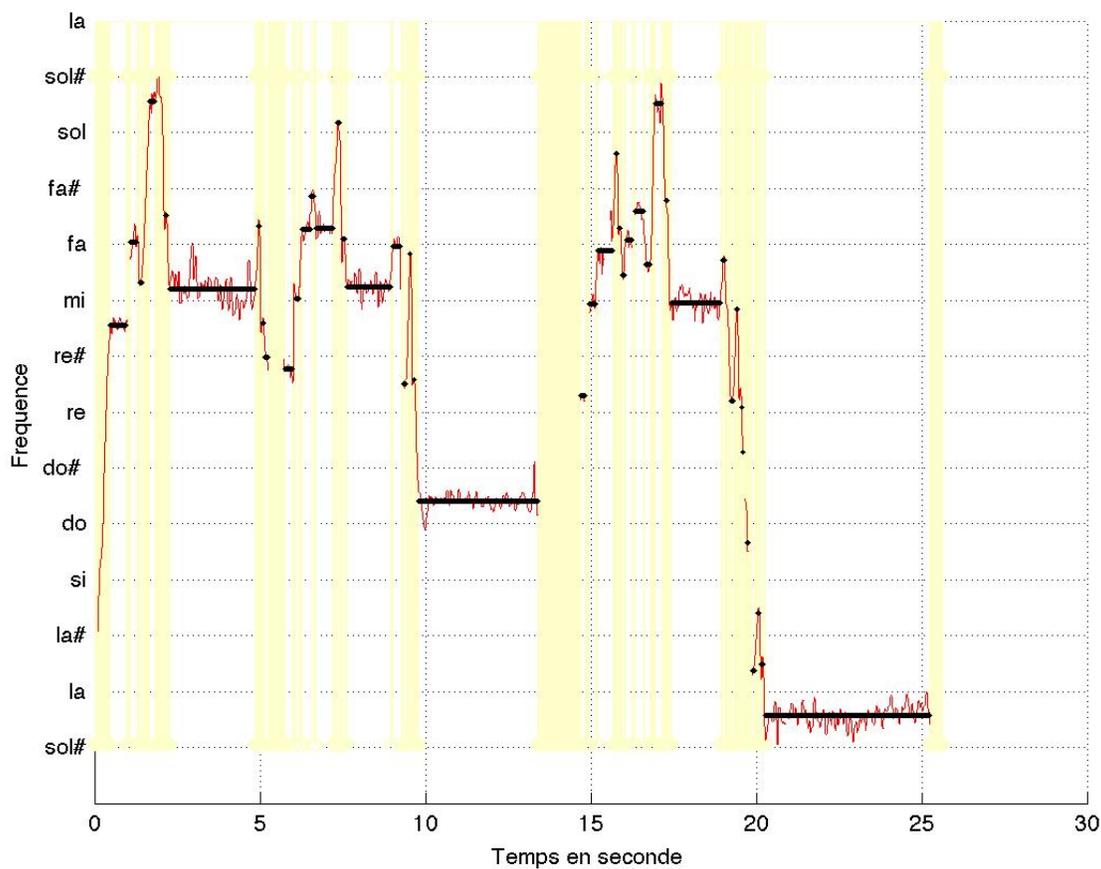


FIG. 14 – Détection des notes.

Paramètres

Durée minimale de note : 10 ms

Seuil entre deux notes proches : 50 cents

Durée d'attaques maximales en deux notes proches : 300 ms

En rouge: la courbe de fréquence fondamentale.

En jaune: les attaques détectées.

En noir: les notes détectées.

4.2 Les outils réalisés

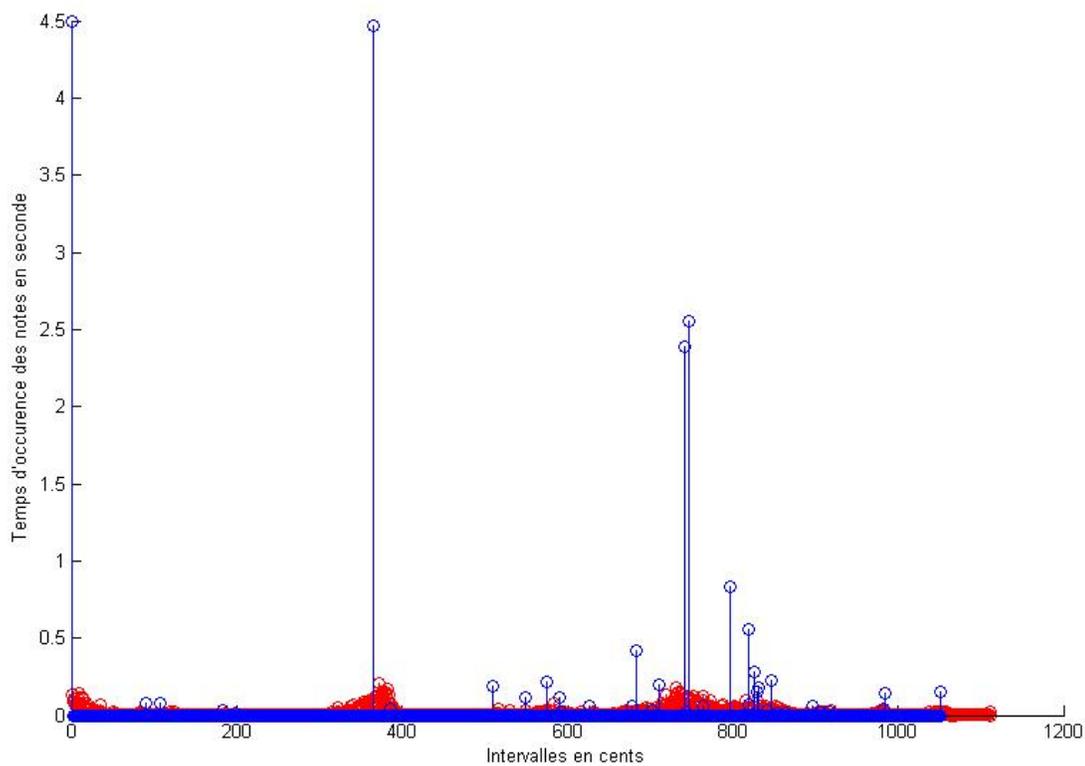


FIG. 15 – Histogramme d'intervalles.

Les agrandissements permettent une visualisation précise des hauteurs de notes.
On peut remarquer l'utilisation d'une ornementation utilisant la seconde mineure (100 cents).

En bleu: le temps d'occurrence des hauteurs de notes reconnues.
En rouge: le temps d'occurrence de chaque hauteur d'échantillon de la fréquence fondamentale.

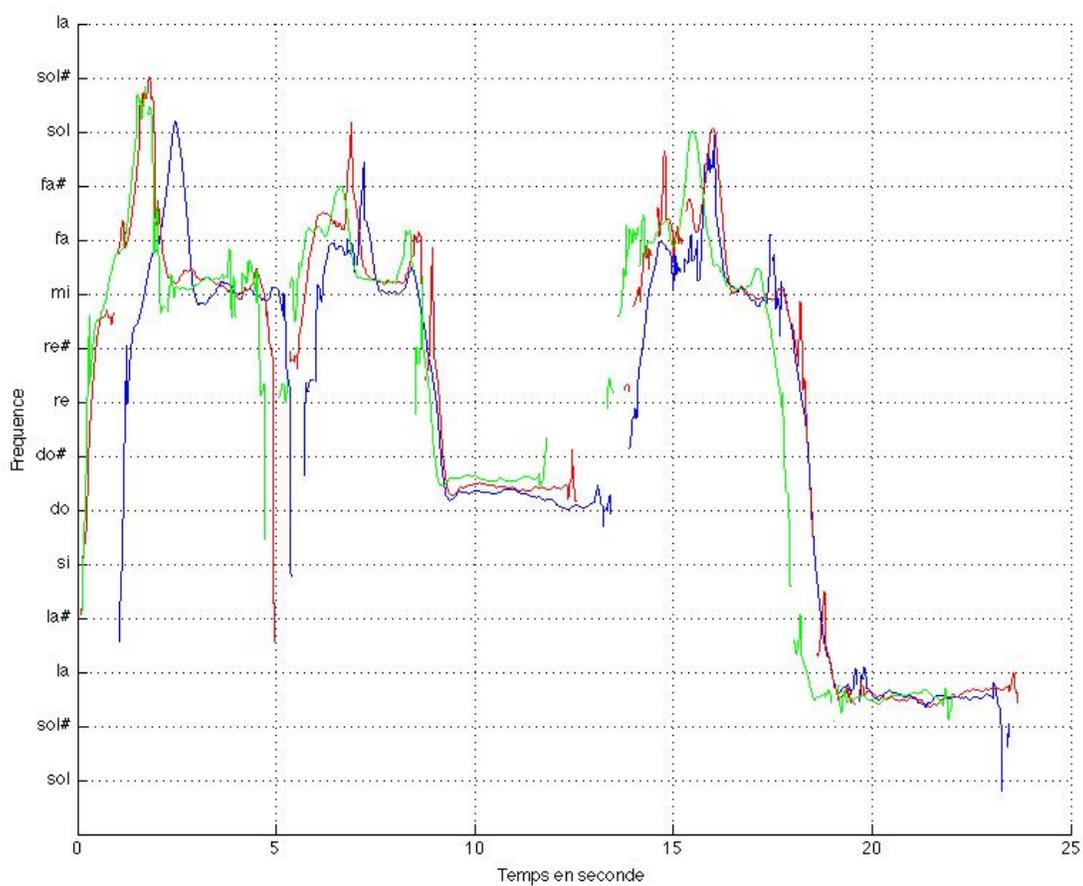


FIG. 16 – Affichage multiple.

En bleu, vert et rouge : Trois strophes de l'Aria del trainieri.

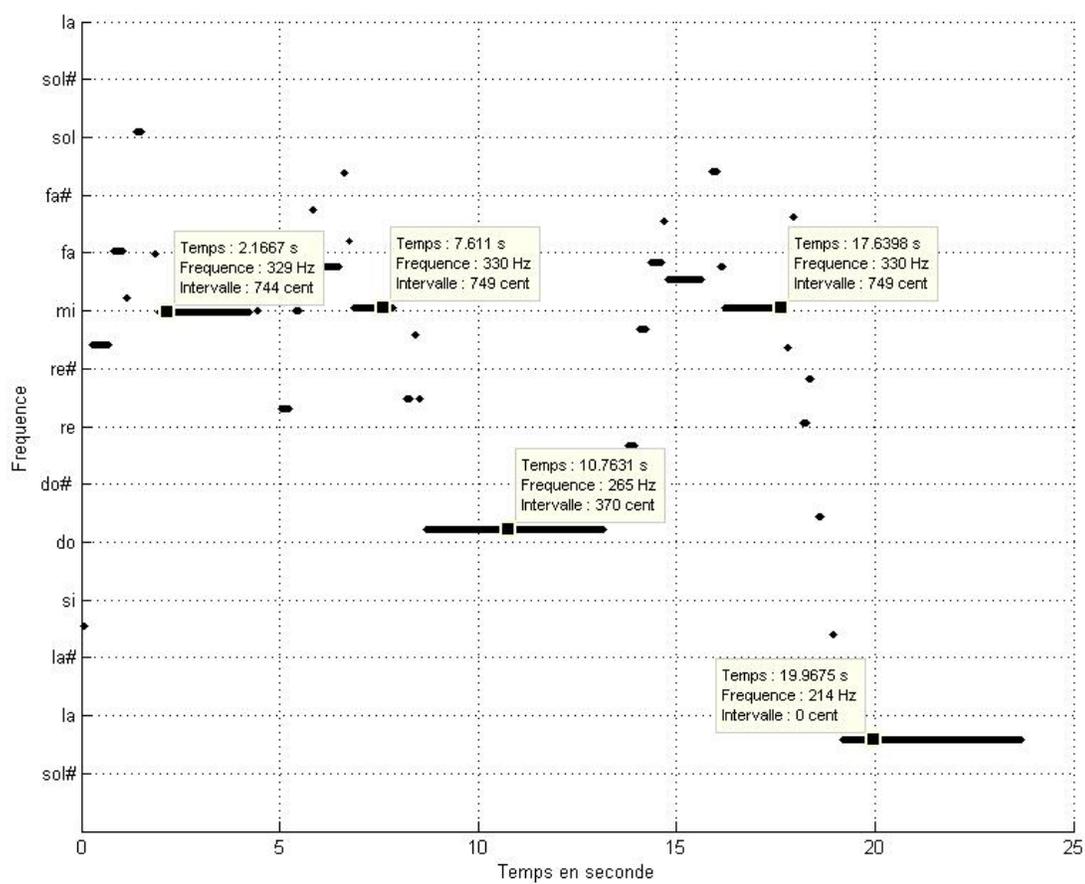


FIG. 17 – Affichage des intervalles utilisés.

Après sélection d'une note de référence, la sélection des autres notes entraîne l'affichage de leurs fréquences, ainsi que les intervalles en cents les séparant de la note principale.

4.3 Les interfaces graphiques

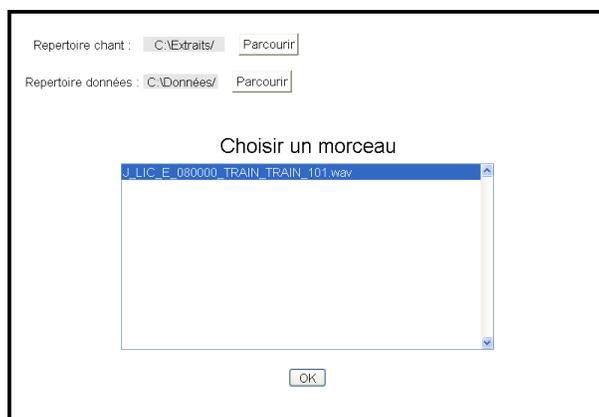


FIG. 18 – Fenêtre de démarrage.

Fenêtre permettant de choisir le ou les morceaux à analyser. Les deux répertoires de travail sont paramétrables : le répertoire où se trouvent les extraits et le répertoire où sont chargées et sauvegardées les données comme la fréquence fondamentale.



FIG. 19 – Fenêtre principale.

A gauche, un lecteur audio permet de lire le morceau.

Au centre, les étapes de traitement déjà réalisées s'affichent en bleu, et les boutons radios permettent de choisir une étape. Selon l'étape choisie, et les traitements déjà réalisés, différents boutons s'affichent sur la droite, ce qui permet de réaliser les actions correspondantes: "Afficher", "Modifier", "Calculer", "Jouer", etc.

4.4 Comparaison avec le logiciel Melodyne.

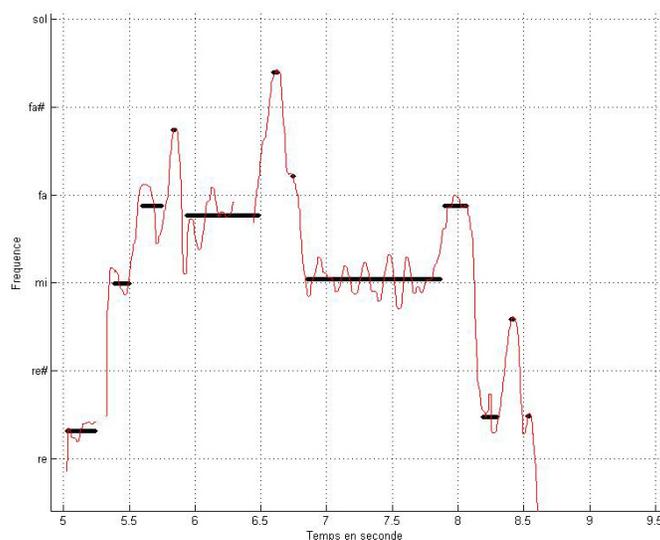


FIG. 20 – Notes reconnues par notre application.

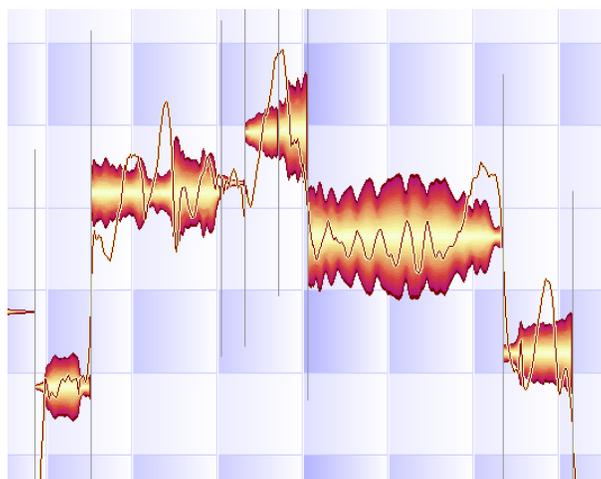


FIG. 21 – Notes reconnues par le logiciel Melodyne.

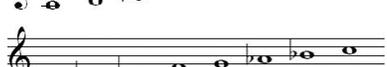
Comme nous pouvons le constater en comparant ces deux figures, l'application que nous avons réalisée permet d'obtenir des résultats, en terme d'analyse des notes chantées, plus précis que le logiciel Melodyne. Par exemple, le trille présent à la huitième seconde n'est pas détecté par ce logiciel. Ce type de détection contribue à une analyse plus précise du geste musical, et aide à comparer différents chanteurs.

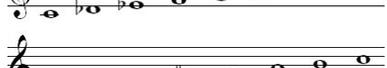
4.5 Exemples de modes musicaux

Nous présentons dans cette partie quelques-uns des modes musicaux les plus utilisés. Commençons par les modes dits grecs ou issus de la gamme majeure. Les intervalles présents dans ces modes se retrouvent en effet à partir des différents degrés de la gamme majeure. Pour plus de lisibilité, nous présentons tous ces modes à partir du degré DO.

Ionien 

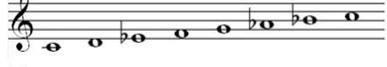
Dorien 

Phrygien 

Lydien 

Mixolydien 

Eolien 

Locrien 

Le mode mineur le plus courant dans la musique classique est le mode mineur naturel, qui correspond au mode Eolien présenté ci-dessus. Il se retrouve en effet à partir du degré LA de la gamme de DO majeur. Le mode mineur harmonique présente la particularité par rapport à ces autres modes de contenir un intervalle d'un ton et demi, entre LAb et SI.

Mineur harmonique 

En cas d'enchaînement des trois derniers degrés de la gamme (LAb-SI-DO), on peut lui préférer les mode mineurs mélodiques ascendants (LA-SI-DO) ou descendants (DO-SIb-LAb).

4.6 Manuel d'utilisation

Nous présentons ici des extraits du manuel d'utilisation de notre application.

Fenêtre de choix des chants

Au démarrage du programme s'ouvre la fenêtre de choix des chants. Dans cette fenêtre doivent être choisis le répertoire dans lequel se trouvent les chants à analyser et le répertoire où sont stockées et chargées les données. Le choix d'un morceau permet de passer à la fenêtre principale. Il est également possible de choisir plusieurs chants à la fois, et dans ce cas le programme n'ouvre pas la fenêtre principale, mais passe en mode "Affichage Multiple".

Fenêtre principale

Un lecteur permet d'écouter le morceau. Toutefois pour des morceaux assez longs, cette fonction peut ne pas être opérationnelle (*Erreur Matlab Audioplayer*).

Les boutons radios ("Signal", "Fondamentale", "Notes", "Tonique" ou "Intervalles") permettent de choisir le type d'information à afficher ou calculer. Quand les informations ont déjà été calculées, le type d'information apparaît en bleu.

Le bouton "Calculer" entraîne l'extraction des caractéristiques, les algorithmes utilisant des valeurs par défaut. Il est possible par la suite de modifier les paramètres des algorithmes, ou de changer les valeurs extraites, à l'aide du bouton "Modifier".

- Modification de l'analyse de la fréquence fondamentale

L'analyse de la fréquence fondamentale est effectuée à l'aide de l'algorithme Yin. Une utilisation plus adaptée des paramètres permet d'améliorer cette analyse :

- Pour trouver des fréquences plus basses : réduire "minf0".
- Pour trouver des fréquences plus hautes : augmenter "maxf0".
- Pour augmenter la résolution temporelle : augmenter "minf0", réduire "hop".
- Pour éviter les erreurs de sous-harmoniques : accroître le seuil de détection.

- Pour éviter les erreurs dues à la présence d'harmoniques: réduire le seuil de détection.

Différents filtres sont utilisables pour modifier les données extraites par le Yin. Le calcul par défaut utilise un filtrage par le taux de variation, un filtrage par l'énergie et un filtrage médian. Le filtre "Suppression par intervalle" peut être utile pour supprimer une partie indésirable d'une analyse.

- Modification de la segmentation de notes
 - Pour ne pas différencier les notes proches, par exemple un vibrato, augmenter le "Seuil en cents entre deux notes proches".
 - Pour trouver plus de notes, diminuer le "Seuil d'attaque".
 - Pour trouver moins de notes, augmenter le "Seuil d'attaque".

Affichage Multiple

Pour afficher plusieurs courbes de fréquence fondamentale sur le même graphique, et ainsi comparer plusieurs extraits, il suffit de sectionner plusieurs chants dans l'interface de choix de chants. Il peut être utile d'utiliser une fréquence fondamentale très filtrée (filtre médian) afin d'effectuer un affichage plus lisible. Un panneau de commande permet de déplacer les courbes. Les informations de fréquence fondamentale ne sont pas modifiées lors d'un déplacement de courbe.

Références

- AKKOÇ, C. (2002). «Non-deterministic scales used in traditional Turkish music». *Journal of New Music Research*, 31(4):285–293.
- AMBRAZEVIČIUS, I., R. et Wiśniewska (2008). «Chromaticisms or performance rules? Evidence from traditional singing». *Journal of Interdisciplinary Music Studies*, 2:19–31.
- AMBRAZEVIČIUS, R. (2004). «Scales in traditional solo singing». Dans *ICMPC8. Proceedings of the 8th International conference on music perception & cognition. August. 3–7*.
- AROM, S. (1991). «L' étude des échelles dans les musiques traditionnelles : une approche interactive». *Analyse musicale*, 23(23):21–24.
- AROM, S., N. FERNANDO, et P.F. MARANDOLA (2007). «An innovative method for the study of african musical scales: cognitive and technical aspects». Dans *Proceedings of the 4th Sound and Music Computing Conference*. 107–116.
- ASKENFELT, A. (1976). «Automatic notation of played music (status report)». *Dept. for Speech, Music, and Hearing: Quartely Progress and Status Report*, 17:1–11.
- ASSELIN, P.Y. (1984). *Musique et tempérament*. Éditions Costallat.
- BETTENS, O. (2001). «Intonation juste à la Renaissance: Ideal ou utopie?» *Esquisse d'un modèle fondé sur la théorie de Zarlino, Zarlino, retrieved*, 12.
- BOUGERET, G. (1989). «Correction du tempérament de l'orgue de Lorrain: Essai de généralisation». *Revue De Musicologie*, 75(1):5–24.
- BURNS, E.M. (1999). «Intervals, scales, and tuning». *The psychology of music*, 2:215–264.
- CAROLI, E. (2009). «La tarentule est vivante, elle n'est pas morte». *Cahiers d'études africaines*, 1:257–284.
- CARPENTIER, J., G. et Barthélemy (2005). «Weighted Scales». URL <http://articles.ircam.fr/textes/Carpentier05b/>.
- DE CHEVEIGNÉ, H., A. et Kawahara (2002). «YIN, a fundamental frequency estimator for speech and music». *The Journal of the Acoustical Society of America*, 111:1917.
- DE MARTINO, Ernesto (1966). *La terre du remords*. Gallimard.
- DEVANEY, D.P.W., J. et Ellis (2008). «An empirical approach to studying intonation tendencies in polyphonic vocal performances». *Journal of Interdisciplinary Music Studies*, 2(1-2):141–156.
- DURING, J. (2004). «Les intervalles musicaux en Asie intérieure à l'ère moderne». *Bibliothèque iranienne*, (58):217–236.
- ELLIS, A.J. (1885). «On the musical scales of various nations». *The journal of the society of arts*:485–527.

-
- FERNANDO, N. (2003). «Echelles et modes : vers une typologie des systèmes scalaires». Dans *Une encyclopédie pour le XXIe siècle*, tome 5, sous la direction de Jean-Jacques Nattiez. Actes sud / Cité de la musique, 945–979.
- FÜRNISS, S. (2006). «Aka Polyphony: Music, Theory, Back and Forth». *Analytical Studies in World Music*.
- GERVASI, F. (2008). «I repertori vocali di tradizione orale nelle recenti espressioni del folk-revival salentino. Nuove ipotesi di metodo a partire dalle problematiche di trascrizione del "livello immanente"». *Il nuovo in musica. Estetiche tecnologie linguaggi*.
- GERVASI, F. (2010a). «Pour une analyse stylistique de la voix chantée dans un contexte de réinvention de répertoires traditionnels». *A paraître dans les actes du colloque : L'analyse musicale aujourd'hui. Crise ou (r)évolution ?*
- GERVASI, F. (2010b). «Deux esthétiques pour un même chant. Esquisse d'une enquête d'ethnomusicologie italienne». *Jérôme Boivin et coll., Actes du 9e colloque étudiant du Département d'histoire:285–300*.
- GÓMEZ, J., E. et Bonada (2008). «Automatic melodic transcription of flamenco singing». Dans *Proceedings of the fourth Conference on Interdisciplinary Musicology, Thessaloniki, Greece*.
- HOWARD, D.M. (2007). «Equal or non-equal temperament in a capella SATB singing». *Logopedics Phoniatrics Vocology*, 32(2):87–94.
- ISIKHAN, G., C. et Ozcan (2008). «A survey of melody extraction techniques for music information retrieval». Dans *Proceedings of the Fourth Conference on Interdisciplinary Musicology (CIM08)*.
- LE BAGOUSSE, Sarah. (2008). «Étude perceptive et acoustique du timbre de la voix chantée dans le contexte des répertoires de tradition orale». *Rapport de stage de Master ATIAM*.
- LOMAX, A. (1954). *Alan Lomax in Salento: le fotografie del 1954*. Kurumuny Edizioni, 2006.
- MARANDOLA, F. (2004). «The study of musical scales in Central Africa: the use of interactive experimental methods». *Computer Music Modeling and Retrieval:34–41*.
- MCNAB, R.J., L.A. SMITH, et I.H. WITTEN (1996). «Signal processing for melody transcription». *Australian computer science communications*, 18:301–307.
- PELLETIER-ORTIZ, S. (1991). «Les échelles musicales d'Afrique centrale : problématiques, hypothèses». *Analyse musicale*, (23):26–30.
- POWERS, J., H.S. et Cowdery (2004). «Mode, IV: Modal scales and traditional music». *Grove Music Online*.
- PROUTSKOVA, P. (2007). «Musical Memory of the World-Data Infrastructure in Ethnomusicological Archives». *Culture*, 100:4–000.
- ROSSIGNOL, S. (2000). *Segmentation et indexation des signaux sonores musicaux*. Thèse de doctorat, Université Paris 6.

-
- ROUGET, M., G. et Leiris (1980). *La musique et la transe. Esquisse d'une théorie générale des relations de la musique et de la possession*. Gallimard.
- RUSSO, W.F., F.A. et Thompson (2005). «The subjective size of melodic intervals over a two-octave range». *Psychonomic bulletin & review*, 12(6):1068.
- SCHELLENBERG, S.E., E.G. et Trehub (1999). «Culture-general and culture-specific factors in the discrimination of melodies». *Journal of Experimental Child Psychology*, 74(2):107–127.
- SEEGER, C. (1951). «An instantaneous music notator». *Journal of the International Folk Music Council*, 3:103–106.
- SUN, X. (2000). «A pitch determination algorithm based on subharmonic-to-harmonic ratio». *6th International Conference of Spoken Language Processing*, 4:676–679.
- SUN, X. (2002). «Pitch determination and voice quality analysis using subharmonic-to-harmonic ratio». *Proc. IEEE Intl Conf. on Acoustics, Speech, and Signal Processing (ICASSP)*, 1:333–336.
- SUNDBERG, J. (1979). «Perception of singing». *The psychology of music*, 20:001–048.
- TYPKE, R., F. WIERING, et R.C. VELTKAMP (2005). «A survey of music information retrieval systems». Dans *Proceedings of the International Conference on Music Information Retrieval*. Citeseer, 153–160.
- TZANETAKIS, G., A. KAPUR, W.A. SCHLOSS, et M. WRIGHT (2007). «Computational ethnomusicology». *Journal of interdisciplinary music studies*, 1(2):1–24.
- VOISIN, F. (1994). «Musical scales in central Africa and Java: Modeling by synthesis». *Leonardo Music Journal*, 4:85–90.
- WARRIER, R.J., C.M. et Zatorre (2002). «Influence of tonal context and timbral variation on perception of pitch». *Perception & psychophysics*, 64(2):198.
- XU, Y., J.T. GANDOURA, et A.L. FRANCIS (2005). «Effects of language experience and stimulus complexity on the categorical perception of». *J. Acoust. Soc. Am.*, 109:775–794.