

UNIVERSITÉ PARIS 8 – VINCENNES- SAINT-DENIS

UFR ARTS, PHILOSOPHIE, ESTHÉTIQUE

Département de Musique

Master Arts, mention Musique

Spécialité : Musicologie, création et société

**VERS UNE NOUVELLE GESTUALITÉ INSTRUMENTALE:
L'EXPLOITATION DE DONNÉES VOCALES POUR UN
DÉVELOPPEMENT DE L'ÉCRITURE INSTRUMENTALE**

Mémoire de Master 2, présenté et soutenu par

Hadas PE'ERY

Le 11 septembre 2012

Réalisé sous la direction de Professeur Anne SEDES

Année universitaire 2011/2012

TABLE DE MATIÈRES

INTRODUCTION.....	1
I. LA VOIX EN TANT QUE MODELE.....	3
I.1 Le Contexte historique.....	3
I.2 Peter Ablinger – Phonorealism : une reproduction instrumentale de la voix...4	
I.2.1 Pensée.....	4
I.2.2 Méthodes.....	6
I.2.3 Critique	10
I.3 Jonathan Harvey – <i>Speakings</i> : un modèle vocal pour l’orchestration.....	13
I.3.1 Pensée.....	13
I.3.2 Méthodes.....	13
I.3.3 Critique	16
I.4 Résumé.....	17
II. LE TRAITEMENT NUMERIQUE DES ELEMENTS VOCAUX TRANSITOIRES....	19
II.1 Les sons transitoires.....	19
II.2 Méthodes d'analyse et de représentations.....	21
II.2.1 La Transformée de Fourier à court terme.....	21
II.2.2 La Transformée en ondelettes.....	23
II.2.3 Les Coefficients cepstraux en échelle de Mel.....	26
II.3 Applications informatiques.....	30
II.3.1 La Segmentation : <i>Sphinx IV</i> et <i>op.recognize</i>	31
II.3.1.1 Applications.....	35
II.3.2 La Représentation temps-fréquence avec Max/MSP.....	37
II.3.2.1 Applications.....	40
II.4 Résumé.....	44
III. REALISATIONS MUSICALES.....	45
III.1 <i>Auto-détermination</i> : transcription sur un texte caché.....	45
III.1.1 Le Projet et les choix artistiques.....	45
III.1.2 L'Écriture instrumentale.....	46
III.1.2.1 Approche phonétique-acoustique.....	46
III.1.2.2 Approche structurelle-globale.....	50
III.1.3 Résumé.....	53
III.2 <i>Leçons d'hébreu</i> pour ensemble et voix de femme.....	54
III.2.1 Le Projet.....	54

III.2.2	Le Choix de texte.....	54
III.2.3	La Forme.....	55
III.2.4	L'Écriture.....	57
III.2.4.1	Les Équivalents sonores.....	57
III.2.4.1.1	Les Consonnes fricatives.....	57
III.2.4.1.2	Les Consonnes occlusives.....	60
III.2.4.2	<i>Leçon A, Leçon G</i>	61
III.2.4.3	<i>Leçon E</i>	68
III.2.4.4	<i>Autres mouvements</i>	73
CONCLUSIONS.....		75
BIBLIOGRAPHIE.....		77
ANNEXE 1 : Lettre de Arnold Schoenberg : texte		
ANNEXE 2 : <i>l'Auto-détermination</i> pour flûte basse : partition		
ANNEXE 3 : <i>Leçons d'hébreu</i> : textes		
ANNEXE 4 : « Leçon A », « Leçon E » : partitions		
ANNEXE 5 : <i>Leçons d'hébreu</i> : partition (sur le CD ci-joint)		
ANNEXE 6 : Exemples sonores (sur le CD ci-joint)		
ANNEXE 7 : Explications de l'acoustique vocale et de la phonétique, extrait de notre travail du M1 (sur le CD ci-joint)		

INTRODUCTION

La composition d'une œuvre instrumentale cohérente entraîne l'emploi de plusieurs facultés de raisonnement, partant de la pensée la plus abstraite nécessaire pour la construction d'un discours en sons, et aboutissant aux notions les plus concrètes nécessaires pour l'achèvement d'une partition exécutable. Ainsi, lors de la réalisation d'une telle pièce, le compositeur met en œuvre divers processus d'agencement permettant de définir, de développer et de manipuler des éléments sonores et rhétoriques. L'application régulière desdits processus conduit à la codification, consciente ou non, de ce que nous pourrions qualifier de dialectique propre à la composition instrumentale. Cette dialectique guide maintes décisions prises dans le travail d'écriture instrumentale.

Or, nous posons la question : comment s'éloigner de cette logique si profondément enracinée ? Est-il possible de lui substituer une autre ? Dans ce travail nous abordons ces questions en proposant une approche de l'écriture instrumentale basée sur des éléments de l'expression vocale. Partant de l'hypothèse que la parole contient des structures intrinsèques purement sonores auxquelles l'auditeur est sensible, nous examinons comment il est possible d'accéder à ces structures, et comment celles-ci peuvent être formalisées et finalement exploitées dans une composition instrumentale.

Nous entreprenons cette tentative suite aux recherches menées dans notre travail de M1, dans lequel nous avons exploré la relation texte-musique du point de vue du traitement numérique en appliquant nos conclusions à la réalisation d'une étude pour voix et électronique en temps réel. Lors de la conception de cette étude, nous avons constaté que nos choix compositionnels étaient dictés par un raisonnement autre, adapté aux propriétés physiques-acoustiques de la production vocale et à la manière dont celles-ci se transforment en fonction du texte ; la logique qui s'y manifeste s'écarte considérablement de celle présente dans notre pensée musicale habituelle. Ainsi, à l'issue de ce travail, nous avons postulé que les structures dirigeant cette logique sont aussi valables pour la musique instrumentale que vocale.

Dans l'approche que nous proposons ici, nous tentons donc d'isoler des qualités sensibles de la voix parlée et de les transférer au langage instrumental afin d'engendrer des nouveaux éléments musicaux et une nouvelle dialectique au sein d'une œuvre instrumentale. Le but de ce travail n'est ni d'imiter la voix parlée à l'instrument, ni de transmettre un texte à travers le jeu instrumental, mais plutôt de dévier des pistes de pensée héritées en s'appuyant

sur des structures avec lesquelles l'auditeur est forcément familier et les exploitant dans un contexte autre. Cette approche soulève plusieurs questions concernant la perception : où se situe la frontière entre l'évocation d'une source et sa mise en évidence ? A quel point est-il nécessaire que l'auditeur soit conscient de ladite source ? Qu'est-ce qui distingue l'imitation de la composition s'appropriant des caractéristiques d'un langage existant ?

Ces questions sont abordées de manières différentes dans les trois sections de ce mémoire. Dans la première, nous analysons les œuvres des compositeurs contemporains ayant entrepris des tentatives semblables. Ici sont présentées leurs théories artistiques, les façons dont celles-ci sont modifiées par l'emploi des outils technologiques, et finalement les rendus musicaux qui en sont issus.

Dans la deuxième section, nous explorons divers outils de quantification des données acoustiques de la parole, dans le but de pouvoir les reproduire dans une certaine mesure à l'instrument. Cette section s'occupe également du rapport entre les propriétés mesurables par des outils numériques et celles pertinentes pour la perception humaine du son.

Dans la dernière section, nous présenterons deux réalisations musicales originales : une transcription pour flûte basse de l'étude vocale mentionnée ci-dessus, et une œuvre pour voix et ensemble instrumental. La transcription constitue une étude, une application intuitive des notions tirées des deux sections précédentes, traitant principalement de la traduction du son vocal en son instrumental. Dans l'œuvre ultime, il ne s'agit pas d'une démonstration de tous les éléments recherchés et développés dans ce travail : c'est plutôt une composition musicale indépendante cherchant à intégrer lesdits éléments, et surtout une logique compositionnelle provenant de l'expression vocale, à divers niveaux du travail de création.

I. LA VOIX EN TANT QUE MODÈLE

Dans cette première section de notre travail, nous examinons des œuvres de compositeurs contemporains ayant employé la voix comme modèle pour l'écriture instrumentale, en s'appuyant sur des nouvelles technologies pour le transfert des propriétés vocales aux instruments¹. Avant d'aborder lesdites œuvres, nous souhaitons exposer brièvement certains événements historiques qui ont conduit à l'état actuel des recherches.

I.1 CONTEXTE HISTORIQUE

L'emploi des éléments vocaux dans la musique instrumentale peut déjà être observée dans la musique ancienne. Bien que les techniques appliquées soient issues des exigences pratiques et soumises aux contraintes artistiques de l'époque, nous trouvons certaines similitudes entre la pensée instrumentale dans la musique ancienne et celle développée au fil de notre travail.

Dans nombreuses compositions polyphoniques de la Renaissance et du début du Baroque écrites pour voix et instruments, les instruments servent principalement à doubler les parties vocales. En interprétant des telles compositions, les instrumentistes imitent souvent la prosodie de la voix : les ponctuations sont placées en concordance avec les limites des mots, les phrases sont sculptées afin de ressembler à l'entretien des paroles, et les notes, dont l'organisation rythmique est une reproduction des syllabes du texte, sont articulées d'une manière correspondant aux qualités phonétiques (Aud. 1)². Nous constatons que cette pensée de l'articulation est souvent présente dans l'interprétation des œuvres purement instrumentales du baroque tardif, malgré le développement d'une écriture instrumentale indépendante.

Avec l'évolution des formes instrumentales, d'une écriture instrumentale idiomatique et de la fabrication des instruments à partir du 17^{ème} siècle, ainsi qu'avec la mise en valeur de la virtuosité d'interprétation dans l'époque romantique, les dits liens entre la musique vocale

1. Dans le cadre de ce travail nous ne présentons que deux pièces dont les analyses sont particulièrement pertinentes vis-à-vis du développement des idées compositionnelles exposées par la suite. Néanmoins, nous souhaitons mentionner certaines œuvres de la littérature contemporaine dans lesquelles la relation entre la voix parlée et le jeu instrumental est explorée d'une autre manière, et dont l'influence sur l'écriture vocale et instrumentale d'aujourd'hui est incontestable. Parmi lesdites œuvres nous pouvons citer : G. GRISEY, *Les Chants de l'amour*; B. FURRER, *Invocation III*; C. MALHERBE, *Locus*; F. BASCHET, *La Muette*.

2. Les exemples sonores se trouvant sur le CD ci-joint sont indiqués dans le texte par l'abréviation « Aud. ».

et instrumentale furent largement déçous.

Le vingtième siècle voit la réémergence d'un raisonnement selon laquelle des qualités vocales sont transférables aux autres média. Ce développement est largement attribué à Pierre Schaeffer et aux principes exposés dans son *Traité des objets musicaux*. En réévaluant le phénomène perceptif, Schaeffer propose une nouvelle écoute – « l'écoute réduite » – qui se concentre sur les caractéristiques acoustiques intrinsèques d'un son en faisant abstraction de sa causalité. Ainsi, en altérant les critères selon lesquels le son est qualifié, l'approche de Schaeffer permet d'embrouiller les frontières rigides entre son vocal, son instrumental et tout autre bruit, comme l'illustre l'exemple suivant extrait de son livre *Solfège de l'objet sonore* (Aud. 2).

Étant fortement influencé par les avancées dans le domaine de la linguistique et de la grammaire générative des années cinquante et soixante, Schaeffer fait souvent allusion aux éléments phonétiques dans son œuvre. Un des termes fondamentaux de la typo-morphologie Schaefferienne par exemple, le couple « articulation/appui », est décrit d'une manière associant l'articulation avec des consonnes et l'appui avec des voyelles. A travers de maintes telles références qui se trouvent dans ses écrits nous observons les germes d'une pensée qui permettra, quelques décennies plus tard, de développer les technologies et de songer des idées musicales discutées dans la suite de ce travail.³

I.2.1 PETER ABLINGER — LE PHONOREALISME : UNE REPRODUCTION INSTRUMENTALE DE LA VOIX

En premier lieu nous abordons une série des œuvres musicales écrite par le compositeur Autrichien Peter Ablinger et intitulée *Quadraturen*.

I.2.1 PENSEE

Les pièces formant la série *Quadraturen* se basent sur la notion du « phonoréalisme », terme inventé par le compositeur pour décrire l'adaptation des techniques de la peinture photoréaliste à la composition musicale. Ainsi, afin de pouvoir mieux comprendre les

3. CHION, Michel. *Guide des objets sonores*. p. 113-120.

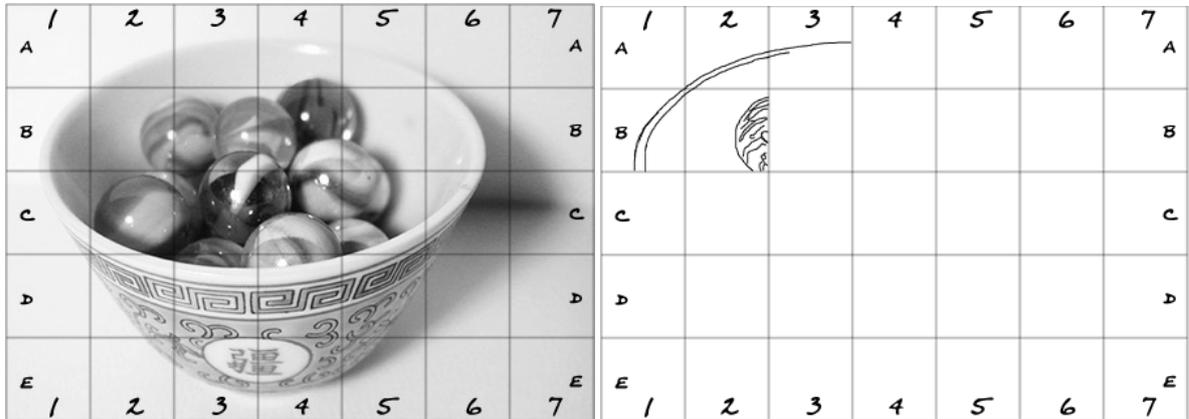
méthodes et la volonté de Ablinger, il nous faut d'abord aborder quelques principes fondamentaux du photoréalisme sur lesquels il s'est appuyé lors de la composition de *Quadraten*.

Dans le rendu photoréaliste, le processus de création commence avec la prise d'une photo. L'image représentée dans cette photo est ensuite agrandie et transférée à une toile mécaniquement de façon que l'image peinte ne puisse pas être distinguée de l'image originale. Ledit transfert d'une photo à une toile peinte est réalisé en employant une technique semblable à celle utilisée pour l'affiche d'une image numérique : la reconstruction de l'image à partir des échantillons graphiques s'appelant *pixels*. Le pixel prend la forme d'un carré minuscule dont la quantité (qui est proportionnelle à la taille) détermine la qualité de l'image. Une image composée d'un grand nombre de pixels aura une résolution élevée et proposera donc un rendu plus réaliste. Par contre, une représentation composée d'une trop faible quantité de pixels sera difficile à reconnaître à cause de sa pauvre résolution (Ex. 1).



Ex. 1 : images en découpée en quatre blocs ayant des résolutions différentes

D'une manière similaire, le peintre photoréaliste reproduit l'image photographique en la découpant en quadrillage et en peignant le contenu de chaque carré le plus précisément possible (Ex. 2). Comme avec l'image numérique, une reproduction fidèle nécessite l'emploi d'un quadrillage plus fin, ce qui signifie que l'image originale doit être agrandie considérablement, afin que le peintre puisse copier les détails y contenus malgré les limitations de l'œil humain. Ce processus engendre un phénomène perceptif remarquable : en regardant le toile de près, le spectateur voit une image abstraite et méconnaissable ; en s'en éloignant, le spectateur voit un rendu de plus en plus clair qui semble avoir été généré à l'aide d'un appareil photo.



Ex. 2 : technique de quadrillage employé dans la peinture photoréaliste

Ce phénomène est au cœur de la pensée compositionnelle de Ablinger dans *Quadraturen*. Ainsi, le compositeur a tâché d'adapter au monde sonore le processus présenté ci-dessous afin d'explorer la frontière entre « la structure abstraite et le passage soudain à la reconnaissance »⁴ manifeste dans la peinture photoréaliste. Pendant des années (1995-2004), Ablinger a développé son idée à l'aide des chercheurs au Studio Expérimental de Freiburg et à l'Institut de Musique Électronique (IEM) à Graz. Cette recherche a donné naissance aux plusieurs pièces tentant à reconstituer des enregistrements (appelés « phonographes » par Ablinger). Les pièces faisant partie de ce travail sont composées pour des divers instruments et révèlent des « résolutions » temporelles et fréquentielles différentes, partant des rendus suffisamment lents qu'ils puissent être joués par des musiciens et aboutissant aux rendus d'une résolution tellement élevée qu'ils donnent l'impression que l'instrument parle.

I.2.1 METHODES

Comme dans le processus photoréaliste qui commence avec une photographie, la création d'un œuvre phonoréaliste emploie l'enregistrement sonore comme source. Dans *Quadratten*, Ablinger utilise principalement des enregistrements du bruit ambiant des villes

4. ABLINGER, P. *Phonorealism, the Reproduction of « Phonographs » by Instruments*

ainsi que de la parole. Nous nous intéressons notamment dans le traitement de ce dernier matériel par la suite de nos recherches.

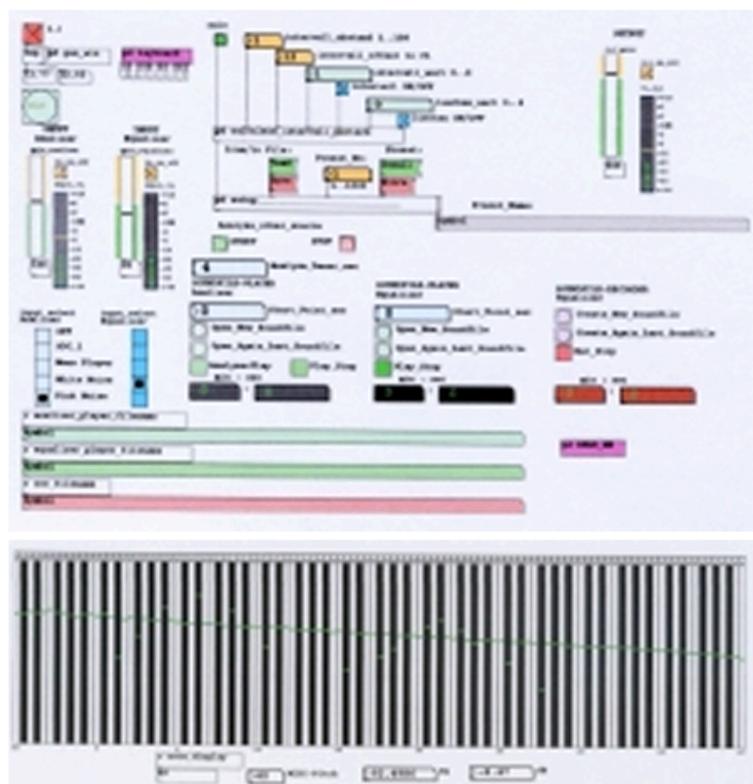
Les enregistrements étant obtenus, il s'agit d'appliquer la technique de quadrillage aux sons. Ceci est entrepris au moyen d'un *freeze*, traitement permettant de figer un instant du son afin d'observer et de manipuler ses propriétés, en l'occurrence le contenu fréquentiel. En employant de nombreux filtres, le son est ensuite analysé par case de fréquence, ce qui constitue un des axes du quadrillage. Le deuxième axe, celui du temps, est produit par l'enchaînement successif des dites « analyses statiques » générées par le freeze et par les filtres. Si dans le photoréalisme la taille et le nombre des pixels déterminent la résolution de l'image, dans le phonoréalisme elle est fonction de l'écartement des filtres d'une part et de la rapidité avec laquelle les analyses statiques se suivent de l'autre.

Le quadrillage temps-fréquence qu'y en résulte est géré par un logiciel codé en C et en Pure Data⁵ ayant deux fonctions : le regroupement des données produits par les filtres selon des échelles d'intervalles microtonaux équidistants, et la transformation ces informations en MIDI (Ex. 3). La première de ses fonctions sert à adapter les informations fréquentielles à une échelle correspondant à la perception humaine de la hauteur⁶ et à contrôler la résolution intervallique (microtonale, semitonale, etc.). La deuxième sert à traduire les hauteurs en notation musicale ou à les communiquer aux instruments électroniques et aux synthétiseurs. En outre des données fréquentielles qui ne constituent qu'un seul axe du quadrillage, la durée des analyses statique peut être gérées afin de varier la résolution temporelle.

Dans les premières pièces de *Quadraturen*, *Quadraturen Ia*, *Ib* et *Ic*, les enregistrements d'origine sont « interprétés » en bruit blanc. La troisième de ces pièces, intitulée « Sprache ist » et basée sur la voix du compositeur, reproduit l'enregistrement à une résolution constant de 250 ms (soit une double croche à un tempo de soixante à la noire), ce qui crée un effet très mécanique (Aud. 3). Par conséquent, c'est le quadrillage qui est mis en évidence plutôt que le contenu vocal. En l'écoutant nous remarquons que bien que la voix, et surtout les parole, sont méconnaissable, nous pouvons néanmoins distinguer un caractère clairement vocal.

5. Programmé par Thomas Musil de l'IEM de Graz.

6. La perception de la hauteur est logarithmique par rapport à la fréquence. Voir la section II.2.3,



Ex. 3 : patch Pure Data employant un banc de filtres

La prochaine tentative de Ablinger que nous discutons, *Quadraturen II*, prend la forme d'une installation / sculpture pour six haut-parleurs recréant en bruit blanc les paroles des discours de Lenin, Hitler, Gandhi, M.L. King, J.F. Kennedy et Castro. Comme les sons correspondants à chaque orateur sont fusionnés, une texture sonore est formée dans laquelle les voix, et surtout les paroles, ne peuvent pas être identifiées (Aud. 4). Mais l'installation de Ablinger est disposée d'une telle manière qu'en se plaçant dans des endroit précis marqués dans la salle, l'auditeur entende un des personnages plus clairement que les autres et puisse même reconnaître certains éléments timbriques de sa voix. Par ailleurs, un des noms de six personnages est noté sur des plaques fixées à chacun des enceintes-sculptures. Ainsi, en s'appuyant sur ses oreilles ainsi que sur les informations affichées, l'auditeur arrive à identifier la source de ce qu'il entend de temps à autre. Ce jeu perceptif s'approche au but initial de Ablinger en composant *Quadraturen*.

Ensuite, Ablinger a entrepris la composition des pièces instrumentales dérivées des enregistrements vocaux en faisant emploi de cette même technique de quadrillage. Ayant déjà écrit quelques études instrumentales en utilisant les textes des *Quadraturen II*, Ablinger a trouvé ces études inadéquates car la résolution temporelle a été limitée par les capacités

motrices des interprètes.⁷ Or, Ablinger a souhaité générer une pièce d'une résolution temporelle suffisamment haute pour que les textes lus dans les enregistrements d'origine soient compréhensibles. Selon ces recherches, une résolution qui serait acceptable à cette fin correspond à seize notes par seconde, en prenant compte du fait que chaque note devrait être jouée à une nuance différente. Inutile de dire que ce niveau de précision est impossible pour un interprète humaine à atteindre, même en étant doté d'une maîtrise technique invraisemblable. Par conséquent, Ablinger a demandé de l'ingénieur Winfried Ritsch à fabriquer un piano mécanique capable de jouer à cette vitesse.

Ainsi équipé d'un instrument capable de jouer seize notes par seconde aux nuances différentes, Ablinger a entamé la réalisation d'un groupe des pièces pour ledit piano mécanique, *Quadraturen III*. Ces pièces reproduisent plusieurs textes, dont la déclaration du Tribunal Environnemental International rédigé par le Dalai Lama, des prières lues par les parents du compositeur et une lettre écrite par A. Schöneberg à son éditeur entre autres. Les résultats de ce travail produisent encore un phénomène perceptif intéressant : le piano mécanique ne donne pas un rendu immédiatement tangible de la parole ; par contre, en lisant le texte d'origine, l'auditeur est capable de le reconnaître dans l'interprétation du piano (Aud. 5, annexe 1).

A l'issue de cette exploration exhaustive, Ablinger et ses collaborateurs ont utilisé les technologies développées lors de ce travail afin de construire une « machine » permettant de transformer tout enregistrement en partition orchestrale (Ex. 4). Le but de cette dernière partie de *Quadraturen* a été d'accéder aux niveaux plus abstraits de la création phonoréaliste, c'est-à-dire de s'éloigner des aspects imitatifs inhérents qui sont mis en évidence dans les autres pièces de la série. Grâce à son caractère sonore tellement connoté et à sa richesse de timbres, l'orchestre symphonique engendre un univers musical abstrus, même en étant employé en tant qu'instrument de reproduction. Cependant, les enregistrements originaux sur lesquels la composition orchestrale s'appuie peuvent être identifiées par moment. Ainsi, cette dernière partie du travail tente à créer « une région limitrophe entre un lieu d'une réalité purement sonore et celui d'une réminiscence musicale ».⁸

Les intentions de Ablinger dans cette dernière pièce de la série, *Quadraturen V (Musik)*, deviennent plus évidentes en écoutant un extrait de la dernière pièce (Aud. 6) ainsi

7. ABLINGER, P. Lecture in *Impuls Academy*.

8. ABLINGER, P. *Quadraturen*.

que ce même extrait superposé sur l'enregistrement original (Aud. 7), en l'occurrence l'hymne national de la République démocratique allemande « Auferstanden aus Ruinen » composé par Hanns Eisler en 1949. Nous constatons ici l'effet perceptif voulu par Ablinger : en écoutant ce premier extrait, nous pourrions le considérer en tant qu'une composition musicale d'une nature abstraite ; en écoutant les extraits superposés, nous remarquons que les sons orchestraux portent un sens autre. Ainsi, pour l'auditeur familier avec l'enregistrement d'origine, l'expérience d'écoute bascule entre la découverte d'une nouvelle musique et le souvenir d'une ancienne.

I.2.3 CRITIQUE

La pensée derrière l'œuvre de Ablinger est fascinante, ainsi que les tentatives entreprises par le compositeur afin d'engendrer une expérience d'écoute si déstabilisante. Sa volonté de créer une musique dérivée si étroitement de la parole sans que ce matériel initial soit mis en évidence, s'approche de notre but qui consiste à établir une écriture instrumentale indépendante, néanmoins fortement liée à l'expression vocale.

Cependant, deux aspects principaux de la réalisation des idées du compositeur nous paraissent peu satisfaisants : la dépendance à des données fréquentielles vocales pour la reproduction de la parole à l'instrument, et le rôle de l'imitation dans le processus compositionnel. Le premier des ces aspects indique un problème conceptuel de base concernant l'analyse du signal vocal. Ablinger et son équipe de chercheurs ont développé des outils numériques dont la fonction principale est l'extraction des hauteurs présentes dans des moments précis des enregistrements vocaux. Or, il est bien connu que la reconnaissance de la voix par la perception humaine dépend de diverses informations, notamment des données temporelles-dynamiques comme l'articulation et l'évolution des formants. Ces informations sont largement ignorées dans le procédé employé par Ablinger. Nous constatons par ailleurs que les efforts les plus réussis du compositeur de récréer la parole (*Quadraturen III*, pour piano mécanique) sont ceux dont l'interprétation consiste en un maximum des données dynamiques : l'exécution de seize notes par seconde, chacune ayant une nuance différente, commence à créer la notion d'un profil dynamique semblable à celui produit par l'articulation vocale. Ainsi nous trouvons inconsistant, même futile, l'entreprise d'une initiative aussi élaborée et investie en s'appuyant sur des informations d'une utilité limitée.

Fl. 1
 Fl. 2
 Kl. 1
 Kl. 2
 Kl. 3
 Kl. 4
 Hob. 1
 Hob. 2
 Hob. 3
 Hob. 4
 Trp. 1
 Trp. 2
 Pnc. 1
 Pnc. 2
 VI. I 192
 VI. I 194
 VI. I 196
 VI. I 198
 VI. I 1970
 VI. I 1912
 VI. II 192
 VI. II 194
 VI. II 196
 VI. II 198
 VI. II 1910
 VI. II 1912
 Va. 192
 Va. 194
 Va. 196
 Va. 198
 Va. 1910
 Vc. 192
 Vc. 194
 Vc. 196
 Vc. 198
 Cb. 192
 Cb. 194
 Cb. 196

Ex. 4 : extrait de la partition de *Quadratten V*

Quant au deuxième aspect, il semble que à force de son attachement dogmatique au processus d'imitation, l'approche de Ablinger manque d'une certaine musicalité. Certes, ce processus sert comme un point de départ très utile ; cependant, il devrait par la suite donner naissance à une pensée musicale indépendante. Bien que le but de Ablinger dans le dernier stade du travail était de générer une musique suffisamment éloignée des prototypes originaux pour que sa source ne puisse plus être identifiée, ce but a été atteint en soumettant le résultat du processus d'imitation à une série d'autres procédés mécaniques. Nous ne trouvons pas que la composition issue de cette démarche soit réussie en tant qu'elle, sans l'intérêt ajouté par le jeu perceptif décrit auparavant.

Ces observations nous conduisent à la conclusion qu'il nous serait avantageux pour la suite de notre travail d'extraire et de formaliser des propriétés vocales perceptibles, afin que nous puissions les employer dans divers contextes musicaux en faisant abstraction des sources vocales d'origine.

I.2.1 JONATHAN HARVEY — *SPEAKINGS* : UNE REPRODUCTION INSTRUMENTALE DE LA VOIX

Dans cette section nous exposerons certains aspects du travail du compositeur anglais Jonathan Harvey dans sa pièce *Speakings* pour orchestre et électronique, développé à l'IRCAM et créée lors du festival Agora en 2010.

I.3.1 PENSEE

Le but artistique de Harvey en composant *Speakings* fut de faire « parler » un orchestre en mettant en relief des structures non-sémantiques présentes dans la voix. Ainsi, Harvey a cherché à traiter diverses sonorités vocales dont le sens ne dépend pas de la parole mais se transmet plutôt par des qualités dynamiques et spectrales, comme les cris et les babillements d'un bébé, le papotage ou bien le chant d'un mantra. En abordant la réalisation desdites sonorités à travers l'écriture orchestrale et la synthèse électronique, Harvey a désiré évoquer chez l'auditeur la notion d'un orchestre apprenant d'abord à parler, puis employant une langue forte expressive.

Étant un bouddhiste dévoué, Harvey s'est inspirée de cette spiritualité afin de concevoir une forme musicale suivant l'évolution d'un être humain, en partant de son incarnation et aboutissant à son illumination. Cette métamorphose est agencée en trois sections principales correspondant aux trois mouvements de *Speakings* : un premier stade d'apprentissage de la parole, dans lequel la capacité d'exprimer des mots cohérents n'est pas encore établie ; un stade plus avancé dans la maîtrise de la langue, caractérisé par des sons de bavardage et de bredouillage ; un stade final, celui de l'illumination, consistant en des sons issus du chant de *mantra*. Ces trois formes d'expression vocale ont été analysées et formalisées par Gilbert Nouno à l'IRCAM afin qu'elles puissent être employées en tant que modèles acoustiques pour l'écriture instrumentale entreprise par Harvey par la suite.⁹

I.3.2 METHODES

Plusieurs moyens techniques ont été développés ou adaptés aux besoins du compositeur pour la réalisation de cette œuvre. Dans la première étape du travail de

9. HARVEY, J. Entretien informel.

modélisation le logiciel *Melodyne*¹⁰ a été employé afin de quantifier le contenu mélodique et harmonique des signaux vocaux. *Melodyne* a été choisi pour cette tâche grâce à sa fonction de segmentation syllabique, ce qui permet de raffiner l'analyse spectrale de base en traitant des unités correspondant déjà à la perception humaine. Ensuite, un algorithme de détection d'attaques et de suivi des partiels¹¹ a été employé afin d'analyser le contenu harmonique du signal vocal et d'extraire les partiels. Les éléments de cette analyse ayant une présence importante ont été soumis à un processus de quantification rythmique en OpenMusic.¹²

Au terme de ce processus, une partition symbolique représentant les composants mélodiques et harmoniques ainsi que leur évolutions temporelles a été générée (Ex. 5). Mais bien que cette partition révèle des données utiles pour la modélisation vocale, elle ignore des propriétés essentielles des signaux vocaux, dont les structures timbrales et formantiques qui permettent à l'auditeur d'identifier des sons vocaux, notamment des sons de voyelles.

The image displays a musical score for the piece 'A Glass of Chios'. It is divided into two systems. The first system, labeled 'vocal' and 'partiels', shows a vocal line on a single staff and a piano accompaniment on a grand staff (treble and bass clefs). The second system continues the piano accompaniment. The score includes various musical notations such as notes, rests, and dynamic markings, illustrating the melodic and harmonic structure of the piece.

Ex. 5 : extrait de la partition mélodique-harmonique

Ainsi dans l'étape suivante du travail de modélisation, les éléments manquants dans ladite partition ont été traités avec le logiciel *Orchidée*. Développé à l'IRCAM, *Orchidée* est un outil d'orchestration basé sur Matlab qui communique avec des environnements de

10. Logiciel d'édition sonore développé par *Celemony*.

11. Développé par Axel Roebel à l'IRCAM.

12. Langue de programmation visuelle basée sur Common Lisp.

programmation y compris Max/MSP, au moyen des messages en OSC¹³. En s'appuyant sur une base de données constituée des sons instrumentaux ainsi que sur des algorithmes d'orchestration efficaces, *Orchidée* propose des possibilités d'orchestration pour un son donné en tenant compte de ses composants spectraux même les plus fins. L'application de cet outil présente deux avantages principaux : il produit des partitions orchestrales d'une manière considérablement plus rapide que l'orchestration manuelle ; il suggère des combinaisons timbrales non-intuitives auxquelles le compositeur n'aurait pas forcément pensé à cause de sa conditionnement par l'étude d'orchestration classique.



Ex. 6 : Mantra employé par Harvey dans la dernière section de *Speakings*.

En composant *Speakings*, Harvey a fait emploi de *Orchidée* et de la partition harmonique-mélodique exposée antérieurement afin de réaliser le troisième mouvement de l'œuvre. Cette section ultime consiste en un choral étendu composé à partir d'un mantra Bouddhiste des trois syllabes contenant principalement des voyelles: « oh ah hum » (Ex. 6). Le compositeur a souhaité construire un passage dans lequel le timbre orchestral évolue graduellement lors de vingt-six répétitions de ces trois sons, en prenant l'enregistrement dudit mantra chanté par lui même comme modèle acoustique. Au début de ce processus, les trois syllabes sont interprétées par un effectif instrumental réduit d'une couleur orchestrale légèrement sombre et dont les attributs spectraux divergent de ceux donnés par le modèle. Mais au fil de ce passage, le son devient de plus en plus fort et l'orchestration de plus en plus brillante en s'approchant du modèle original (Aud. 8). Cette croissance orchestrale se base sur les combinaisons instrumentales fournies par *Orchidée*, dont l'emploi a été adapté aux intentions du compositeur par Nouno . (Ex. 7).¹⁴

13. Open Sound Control : protocol permettant à communiqué entre systèmes numériques à l'aide des messages.

14. NOUNO, G. « Making an Orchestra Speak » in *Sound and Music Computing Conference*, 2009. p. 1-5.

The image shows a musical score for an orchestra, spanning measures 12 to 14. The score is arranged in a standard orchestral layout with staves for Flute (Fl), Oboe (Ob), Clarinet (Cl), Horn (Hr), Trumpet (Tp), Violin I (Va), Violin II (Va), Viola (Va), Cello (Vc), and Double Bass (Cb). Each staff contains musical notation with dynamic markings (pp, mf, ff) and performance instructions such as 'ascol+ord', 'harm+trgr', 'legno-tratto', 'norvib', and 'Sordina'. The score is divided into two systems, with measures 12-13 in the first system and measures 14-15 in the second system.

Ex. 7 : Possibilités d'orchestration proposées par *Orchidée*.

I.3.3 CRITIQUE

Comme dans le travail de Ablinger, nous trouvons que les idées de Harvey sont innovantes et ont un grand potentiel. De plus, sa pensée s'approche de la nôtre en tentant de traduire les structures inhérentes à l'expression vocale de telle manière qu'elles puissent être exploitées dans l'écriture instrumentale. Nous estimons néanmoins qu'il existe une disparité entre les idées du compositeur et leur réalisation. Cette disparité, nous croyons, est encore due à la trop grande importance attribuée à la notion de hauteur ainsi qu'à l'attachement excessif aux processus de formalisation.

Si dans *Quadraturen* les nouvelles technologies ont été employées principalement pour l'extraction des données fréquentielles, dans *Speakings* le traitement numérique a abordé aussi les éléments du timbre de la voix et leur traduction en timbres orchestraux. Cependant, les possibilités d'orchestration proposées par Orchidée se basent surtout sur les composantes fréquentielles du timbre. Ainsi, cette expansion du traitement numérique ayant pour but de mieux représenter le signal vocal ne nous semble pas aussi fructueuse qu'espéré.

De plus, nous trouvons que le rapport de la dernière section de *Speaking* à l'expression vocale est trop peu perceptible, et ce malgré les gros efforts faits afin de traduire des données vocales d'une manière fidèle. Il nous paraît qu'en s'occupant des démarches de quantification et de formalisation du modèle vocal d'une façon si approfondie, le rôle de la perception n'a pas été suffisamment développé. Il est intéressant par ailleurs de remarquer que l'approche théorique des systèmes produits ici un résultat musical éloigné du modèle, bien qu'une approche similaire chez Ablinger ait généré une musique fortement dépendante du modèle.

Ainsi, à ce stade de nos recherches nous arrivons aux conclusions similaires à celles tirées de l'analyse de *Quadraturen* : dans la suite de notre travail, nous proposons de nous focaliser sur les qualités temporelles-dynamiques de la voix, c'est-à-dire les articulations et les profils dynamiques de l'énoncé vocal, et ce en s'appuyant aussi bien sur la perception que sur les outils informatiques.

I.4 RESUME

Suite à cette exploration des oeuvres de Ablinger et de Harvey, nous avons pu approfondir nos connaissances sur la modélisation vocale et sur ses applications; Nous avons également pu comprendre la manière dont ces modèles sont perçus par l'auditeur. Les réflexions à l'origine des efforts des deux compositeurs nous offrent un riche aperçu des possibilités inhérentes à ce travail.

Étant néanmoins déçus par les réalisations et les résultats sonores, nous avons jugé qu'il nous serait peu utile d'adopter l'approche à la technologie et la modélisation, telle que les compositeurs y ont eu recours. Ainsi dans la suite de nos recherches, nous n'attribuerons qu'une importance secondaire à l'analyse des hauteurs en nous focalisant plutôt sur l'articulation vocale. Cet aspect de l'expression vocale, qui fut largement ignoré dans les œuvres analysées, contient de nombreuses informations que nous croyons importantes pour la perception de la voix.

En considérant l'articulation vocale du point de vue de la phonétique, il est question de l'agencement des consonnes. D'un point de vue acoustique, les sons consonnes sont beaucoup plus inharmonieux et instables, et ainsi riches en transitoires¹⁵. Par conséquent, la deuxième section de notre travail traitera de l'analyse et de la représentation numérique des sons transitoires, afin que ceux ci puissent être exploités pour la composition.

Par ailleurs, ayant déjà présenté notre avis sur le fait que l'attention donnée à la formalisation soit excessive et que celle donnée à la perception soit insuffisante, nous tâcherons dans la section suivante de notre travail de maintenir deux approches : une approche méthodique essentielle pour la recherche effectuée avec des outils numériques, et une approche empirique mettant en valeur la perception.

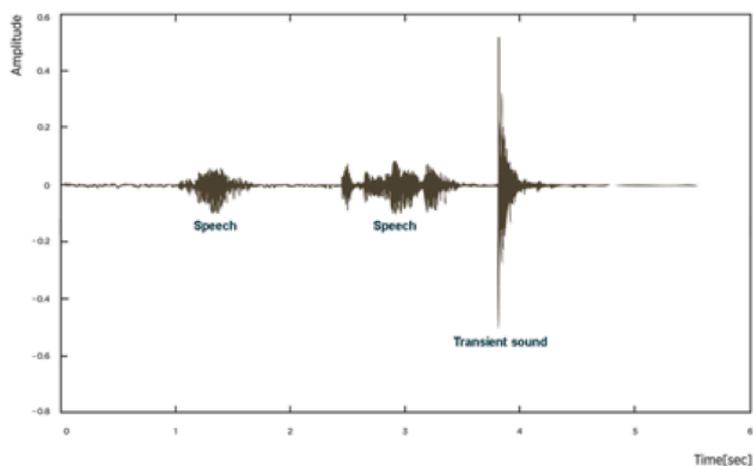
15. Des explications plus approfondies de la phonétique et de l'acoustique de la voix se trouvent dans l'annexe 6.

II. LE TRAITEMENT NUMERIQUE DES ELEMENTS VOCAUX TRANSITOIRES

II.1 LES SONS TRANSITOIRES

Partant des conclusions tirées de la première partie de notre travail, il nous paraît évident que nos recherches subséquentes dans le domaine du traitement du signal devraient se concentrer sur les sons transitoires. En conséquence, nous croyons important de discuter ici des principaux attributs de ces sons, ainsi que certains outils développés pour leur analyse et leur traitement.

Les transitoires sont caractérisés par des changements abrupts et souvent chaotiques d'amplitude, de phase et de fréquence: ils peuvent révéler des désinences très rapides suivant les régions d'*attaque*, ainsi que de brusques variations de hauteur ou de timbre (Ex. 8). Bien qu'ils soient d'une durée extrêmement courte, ils jouent un rôle fondamental dans l'identification d'un son vocal ou instrumental par la perception.¹⁶ Dans le jeu instrumental, des sections transitoires apparaissent principalement au début et à la fin du son (attaque percussive, engagement de l'archet, souffle, étouffement) et lors des changements de hauteur ou de timbre (portamento, vibrato); dans le chant, ils se trouvent dans l'énoncé de consonnes, dans les dites changements de hauteur, lors des transitions entre consonnes et voyelles et lors des transitions entre formants à l'intérieur d'une voyelle. (Ex.)

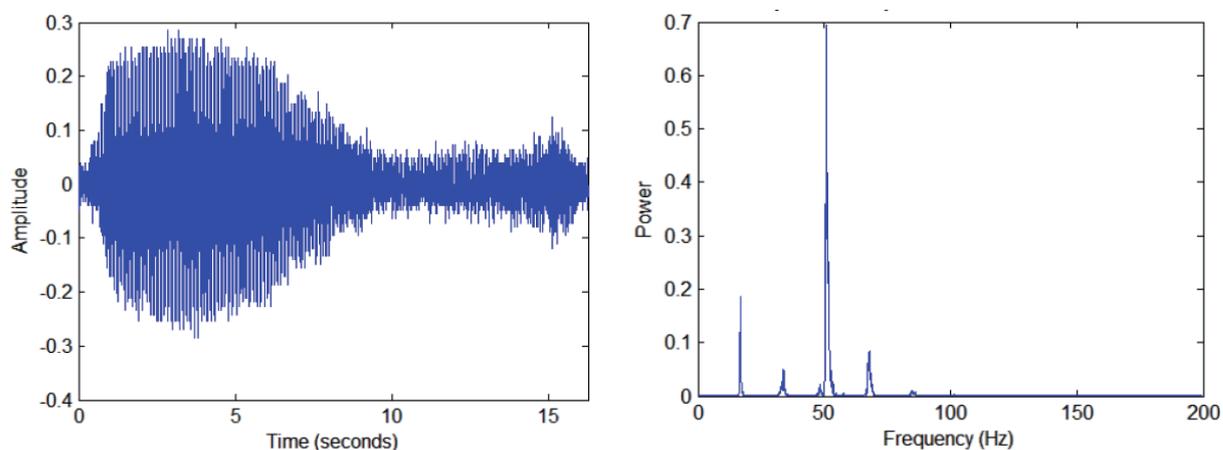


Ex. 8 : comparaison des signaux de parole stables (à gauche, au centre) au signal transitoire (à droite)¹⁷

16. LEVEAU, P. « Paramétrisation adaptée de transitoires pour la reconnaissance d'instruments de musique », p. 9-10.

17. KEIDSER, G. « Is the Hearing Instrument to Blame When Things Get Really Noisy ? », p. 3.

Puisque les transitoires contiennent des informations fréquentielles ainsi que des informations temporelles (amplitude, etc.) complexes, une analyse fructueuse nécessiterait une représentation à un niveau de haute résolution, ce dans le domaine du signal ainsi que dans le domaine du temps. Or, les formes standards de représentation du son ne fournissent des informations précises que dans l'un ou l'autre de ces domaines : la transformée de Fourier contient un maximum d'informations spectrales mais nous ne donne aucune indication sur l'écoulement du temps ; la forme d'onde présente des données temporelles mais ne révèle aucune information spectrale (Ex. 9)¹⁸. Cette imperfection est une expression du principe d'incertitude¹⁹, selon lequel il est fondamentalement impossible de déterminer simultanément la fréquence et le temps d'une manière précise.



Ex. 9 : à gauche, la représentation en forme d'onde d'un signal ; à droite, la représentation des composantes fréquentielles du même signal produites par la transformée de Fourier discrète.

Les progrès des dernières décennies dans le traitement du signal ont donné naissance aux méthodes d'analyse constituant des passerelles entre la représentation temporelle et fréquentielle du signal. Néanmoins, ces outils sont aussi limités par le principe d'incertitude : si nous traitons la fréquence à une haute résolution, nous perdrons des données essentielles définissant l'enveloppe dynamique et les attaques ; si nous faisons l'inverse, c'est-à-dire traiter

18. MATLAB documentation

19. Principe donné par l'inégalité : $\Delta t \cdot \Delta f \geq \frac{1}{4\pi}$. Énoncé par W. Heisenberg en 1927, cette inégalité indique qu'il existe une limite fondamentale à la précision avec laquelle certaines paires des propriétés physiques puissent être mesurés simultanément.

le spectre avec une faible résolution spectrale mais une haute résolution temporelle, notre capacité d'observer les transitions spectrales serait considérablement réduite.²⁰

Cependant, ces contraintes inhérentes à l'analyse des transitoires peuvent être contournées en employant des outils développés dans le domaine de la reconnaissance vocale et de la linguistique informatique. Ceux-ci se basent sur des méthodes de transformation plus adaptées aux propriétés uniques des transitoires et à la perception humaine des dites propriétés. Nous traitons ici de trois méthodes de transformation qui pourraient servir à la conception des gestes instrumentaux inspirés des structures acoustiques existantes dans l'expression vocale.

II.2 METHODES D'ANALYSE ET DE REPRESENTATION

II.2.1 LA TRANSFORMEE DE FOURIER A COURT TERME

La *transformée de Fourier à court terme* (TFCT), ou *transformée de Fourier à fenêtre glissante* est une transformation basée sur la transformée de Fourier dans laquelle la fonction de fenêtrage²¹ est d'une durée courte (typiquement entre 3ms et 750ms). Ainsi, en effectuant une transformée de Fourier, le contenu fréquentiel et la phase ne sont déterminés que localement pour une section limitée du signal mais la possibilité d'observer l'évolution temporelle du spectre est créée (Ex. 10). Afin d'analyser l'intégralité du signal, la fenêtre est glissée et la transformée est ré-calculée tout au long de l'axe du temps. Dans le cas d'une TFCT discrète, des fenêtres chevauchantes sont employées pour limiter les problèmes des artefacts spectraux qui peuvent se produire aux limites des sections transformées.

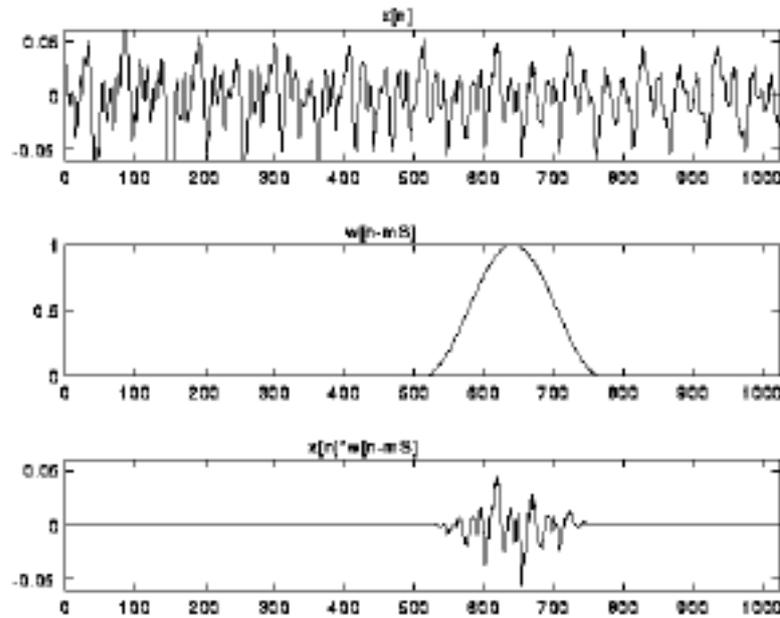
Grâce à la fonction de fenêtrage réduite, la TFCT propose un compromis effectif entre résolution en temps et en fréquence, autant convenable pour des signaux stables que pour des sons transitoires. Par conséquent, maintes logiciels d'analyse et de synthèse sonore comme les vocodeurs de phase, Audiosculpt²² et Spectral Modeling Synthesis²³

20. HAMNANE, S. « Traitement et représentation temps-fréquence des sons avec résolution adaptative », p. 14-15.

21. Fonction mathématique dont la valeur est nulle hors d'un intervalle déterminé

22. Logiciel d'affichage, analyse et traitement sonore développé à l'IRCAM et basé sur le moteur d'analyse/synthèse *Super Vocodeur de Phase*

23. Logiciel d'analyse, de transformation et de synthèse sonore développé par *Music Technology Group* à l'Université Pompeu Fabra



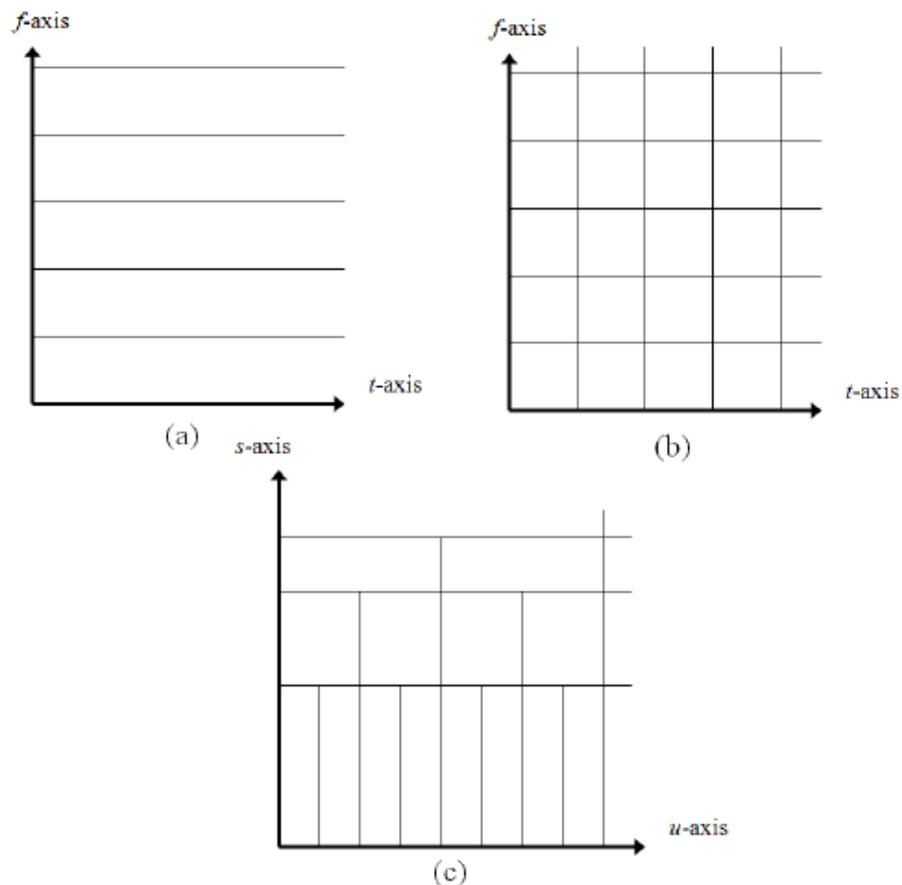
Ex. 10 : En haut, le signal d'origine ; au milieu, la fonction de fenêtrage ; en bas, le segment du signal sur lequel la transformée de Fourier est effectuée. La fonction de fenêtrage est glissée le long du signal afin d'effectuer la TFCT.²⁴

parmi d'autres se basent notamment sur cette forme de transformation. La TFCT possède néanmoins un gros désavantage : la taille de la fenêtre, dès qu'elle est établie, reste fixe à toutes les fréquences (Ex.11b). Or les signaux transitoires contiennent des fluctuations d'étendue variable auxquelles un abord plus flexible serait mieux adapté. Il serait souhaitable d'employer une fenêtre plus longue en traitant les fréquences graves, et une fenêtre plus courte en traitant les fréquences aiguës. Ainsi, la fenêtre constante de la TFCT pose des contraintes importantes dans le cadre de notre travail qui se focalise justement sur ces détails qui caractérisent un son transitoire et le distinguent d'un autre.²⁵

Cet inconvénient de la TFCT a provoqué la création de méthodes d'analyse multi-résolution dans lesquelles l'échelle de la fonction de fenêtrage est variable selon le contenu fréquentiel analysé (Ex. 11c). La *transformée en ondelettes* (discrète ou continue) est une méthode d'analyse multi-résolution privilégiée et très répandue parmi les systèmes commerciaux de reconnaissance et traitement vocal.

24. PEEVERS, A. « A Realtime Signal Analysis/Synthesis Tool Based on the Short Time Fourier Transform », p. 3.

25. HAMNANE, S. « Traitement et représentation temps-fréquence des sons avec résolution adaptative », p. 12-16.



Ex. 11: (a), le pavage temps-fréquence donné par une TF (b), le pavage temps-fréquence donné par une TFCT (c) le pavage temps-fréquence donné par une transformée en ondelettes²⁶.

II.2.2 LA TRANSFORMÉE EN ONDELETTES

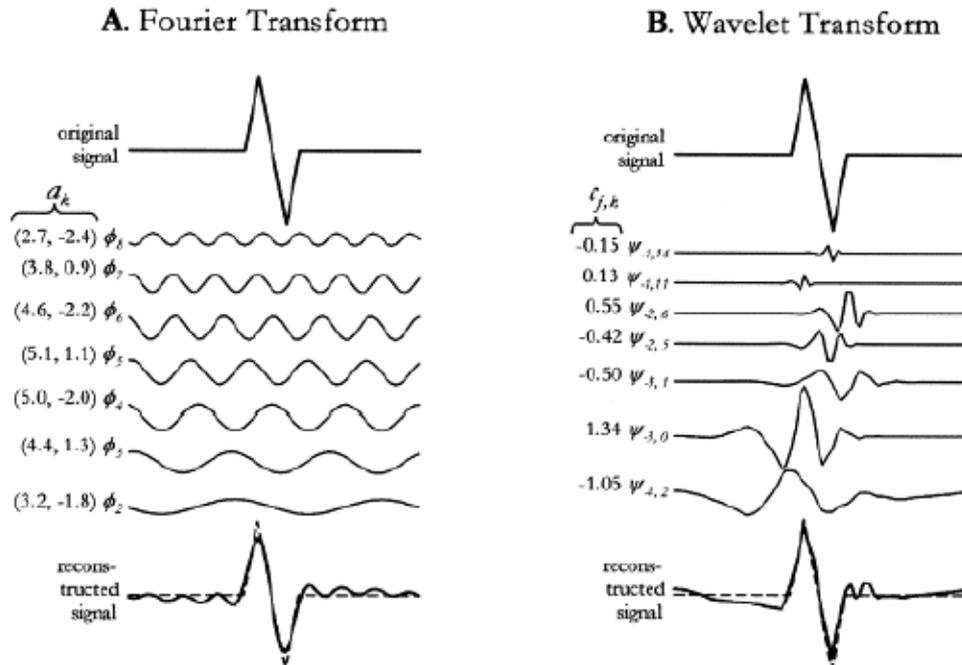
La transformée en ondelettes diffère des transformées de type Fourier par le fait qu'elle ne décompose pas le signal en éléments sinusoïdaux mais plutôt en fonctions de moyenne nulle s'appelant *ondelettes* (Ex. 12). Les ondelettes sont des oscillations extrêmement brèves, les formes les plus élémentaires trouvables dans un signal; ainsi elles sont souvent appelées « atomes temps-échelle »²⁷. Elles se groupent en *familles* composées des transformations temporelles d'un prototype s'appelant *l'ondelette mère*. L'ondelette mère joue un rôle comparable à celui d'une fonction de fenêtrage mais se distingue par sa forme d'une montée et d'une décroissance rapide qui est semblable aux fluctuations présentes dans

26. CHUN-LIN, L. « A Tutorial of the Wavelet Transform », p. 32.

27. RIOUL, O. « Analyse temps-fréquence et temps-échelle », p. 50.

des signaux transitoires. Par ailleurs, une *fonction d'échelle* est utilisée afin de contracter ou de dilater l'ondelette mère; ainsi un effet de « zooming » est produit autour des composantes du signal pour lesquelles une échelle plus ou moins grande serait mieux adaptée (Ex. 13).^{28 29}

30



Ex. 12 : la décomposition et la resynthèse d'un signal en employant des composantes sinusoïdales (A) ondelettes (B)³¹

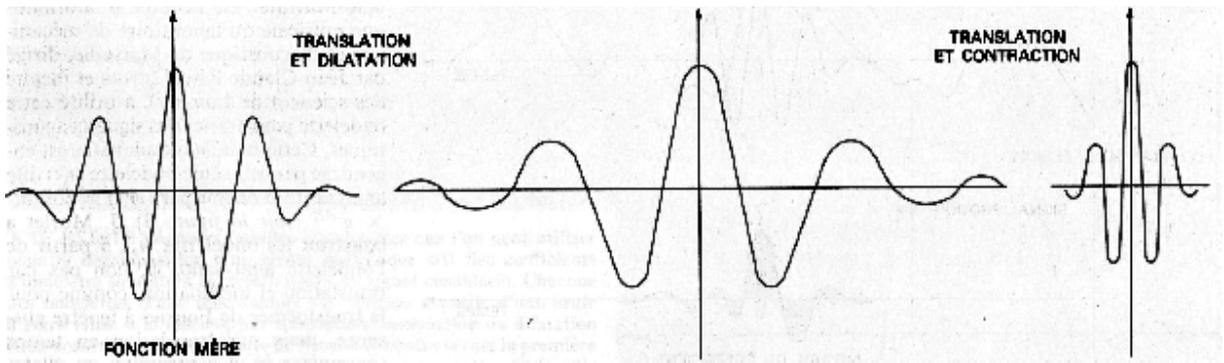
Ainsi la transformée en ondelettes permet de décomposer un signal dans des fonctions fondamentales; ces fonctions peuvent être dilatées (en traitant des fréquences graves) ou contractées (en traitant des fréquences aiguës) dans l'échelle temporelle afin d'obtenir une meilleure résolution. Chaque bande de fréquence du signal analysée est donc représentée à la résolution qui lui est la plus approprié.

28. MEYER, Y. « L'analyse par ondelettes », in *Pour la Science* Vol. 119, p. 28-37.

29. CHUN-LIN, L. « A Tutorial of the Wavelet Transform », p. 3-45.

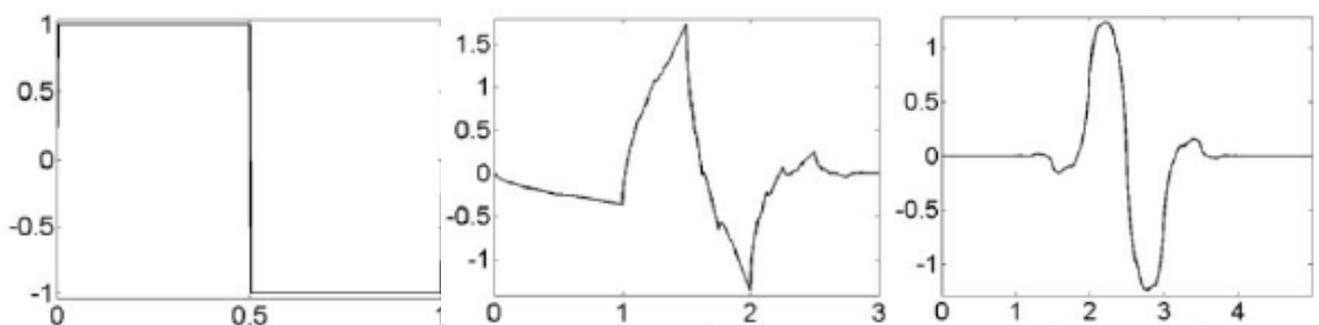
30. DAUBECHIES, I. *Ten Lectures on Wavelets*, p. 14-128.

31. LETELIER, J.C. « Spike Sorting Based on Discrete Wavelet Transform Coefficients », in *Journal of Neuroscience Methods* Vol. 101 No. 2, p. 96.



Ex. 13 : Modification d'échelle : la dilatation (centre) et contraction (à droite) d'une ondelette mère (à gauche)³²

Plusieurs familles d'ondelettes ont des applications dans le traitement des signaux numériques: les ondelettes discrètes orthogonales *Daubechies* et *Haar*, et bi-orthogonales *B-spline* (Ex. 14) par exemple, forment la base des algorithmes de compression et de débruitage des images (jpeg) et de la parole (systèmes d'amélioration vocale). Dans le domaine de la reconnaissance vocale, les ondelettes de la famille de Daubechies sont souvent employées pour identifier des phonèmes, pour suivre la segmentation de la parole et pour raffiner les algorithmes de détection activée par la voix. Nous pourrions donc imaginer que la transformée en ondelettes nous servira lorsque nous chercherons à isoler les phonèmes et à extraire leurs caractéristiques remarquables, d'une manière plus précise que les transformations de type Fourier.



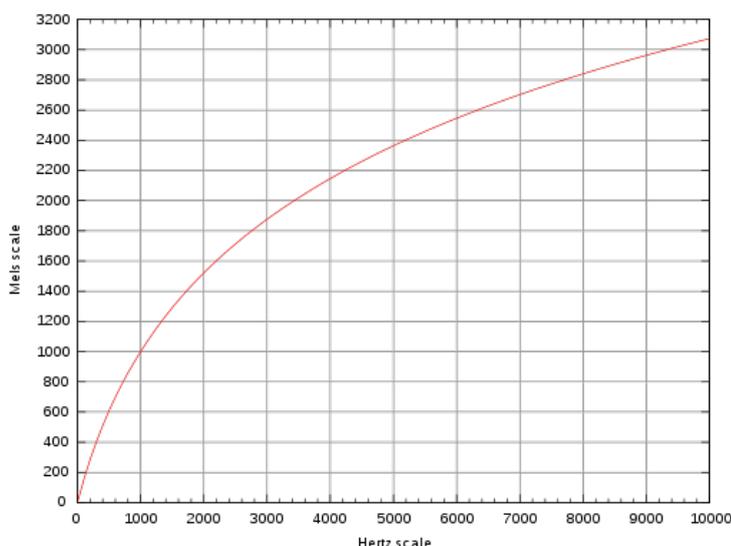
Ex. 14 : Fonctions d'ondelette mères : à gauche, *Haar* ; au centre, *Daubechies* ; à droite, *Biorthogonale*

32. MEYER, Y. « L'analyse par ondelettes », in *Pour la Science* Vol. 119, p. 29

II.2.3 LES COEFFICIENTS CEPSTRAUX EN ECHELLE DE MEL

Un grand nombre d'algorithmes de reconnaissance et d'amélioration vocale qui ne se basent ni sur la TFCT ni de la transformée en ondelettes font emploi de *coefficients cepstraux en échelle de fréquence Mel* (Mel frequency cepstral coefficients: MFCC). Afin d'expliquer cette méthode il nous faut tout d'abord définir les notions fondamentales auxquelles elle fait référence, à savoir l'échelle Mel d'une part, et le cepstre d'autre part.

Conçue par les acousticiens américains Stevens, Volkman et Newman, l'*échelle de Mel*³³ est une répartition de fréquences qui correspondent à la perception humaine, en tenant compte du fait que la hauteur perçue croît moins vite que la fréquence réelle à partir de 500 Hz environ³⁴ (Ex.). Des incréments fréquentiels de plus en plus grands sont ainsi présentés comme étant équidistants au delà de ce seuil. La représentation proposée par l'échelle de Mel est donc linéaire par rapport à l'audition et logarithmique par rapport à la fréquence.



Ex. 15 : L'échelle de Mel ³⁵

Le domaine de l'échelle est exprimé en *mels*, unités calculées selon la formule suivante:³⁶

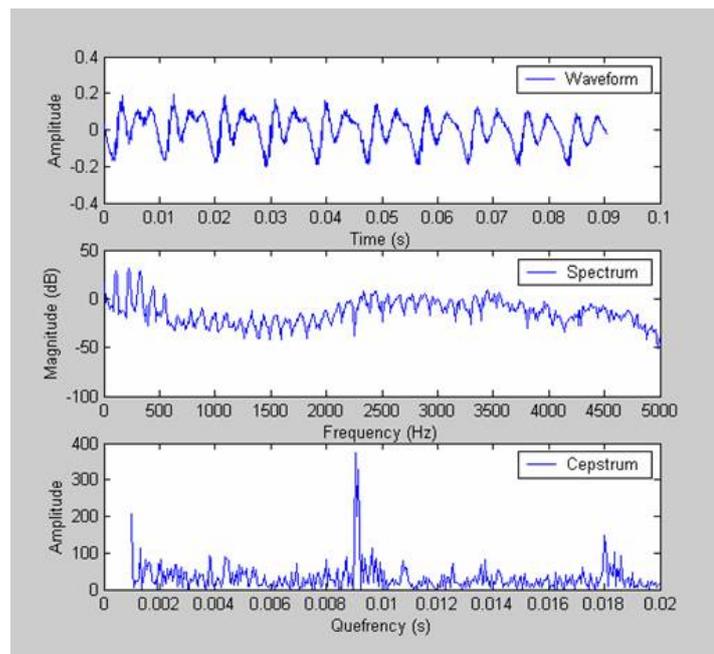
-
33. L'appellation « Mel » vient du mot « mélodie » afin d'indiquer que l'échelle soit basée sur des rapports de hauteurs.
 34. En vertu d'un principe général de la perception appelé « loi de Weber-Fechner », la perception des hauteurs est en première logarithmique, c'est-à-dire basée sur des rapports de fréquences. Il n'est pas question ici des hauteurs musicales, mais d'une échelle de hauteur plus fondamentale, à savoir la sensation « grave-aigu ».
 35. O'SHAUGHNESSY, D. *Speech Communication – Human and Machine*. p. 4

$$m = 2595 \log_{10} \left(1 + \frac{f}{700} \right) = 1127 \log_e \left(1 + \frac{f}{700} \right)$$

de telle façon qu'une mesure de 1000 mels corresponde à 1000 Hz, 40 dB au dessus du seuil de l'audition.

Le *cepstre* est la transformée de Fourier du logarithme du spectre d'un signal; sa courbe est fonction de la *quefrence*, un indice de temps (Ex. 16). Inventé à la base afin de faciliter la caractérisation des échos sismiques, le cepstre fournit des informations utiles sur le taux de change des différentes bandes spectrales. Plusieurs formes du cepstre existent : le *cepstre complexe*, le *cepstre réel*, le *cepstre de phase* et le *cepstre de puissance*. Ce dernier, qui est défini comme la magnitude carrée de la transformée du logarithme de la magnitude

carrée de la transformée d'un signal : $\left| F \left\{ \log \left(|F\{f(t)\}|^2 \right) \right\} \right|^2$ est souvent utilisé pour le traitement vocal.³⁷



Ex. 16 : Trois représentations d'un signal : au dessus, la forme d'onde ; au centre, le spectre, en bas, le cepstre.³⁸

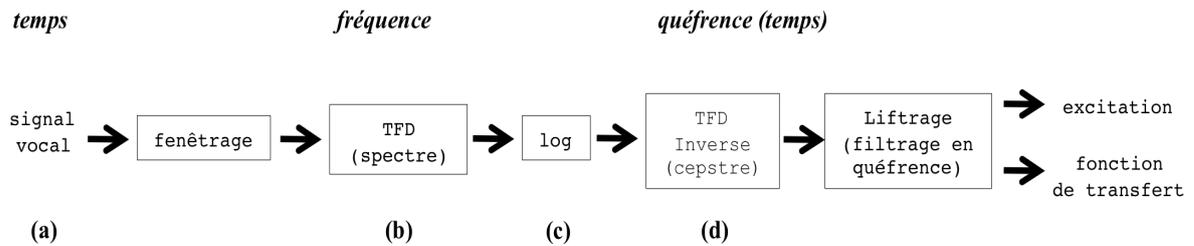
En tant que concept mathématique, le cepstre a une propriété particulièrement avantageuse pour l'analyse du signal vocal : la convolution de deux signaux peut être

36. *Ibidem.* p. 4

37. CHILDERS, D. G. « The Cepstrum : a Guide to Processing », in *Proceeding of the IEEE* Vol. 65 No. 10, p. 1428-1443

38. UCL dept of phonetics and linguistics

exprimée comme l'addition de leurs cepstres. Puisque la production vocale est représentée dans le domaine temporel par la convolution des signaux de la source glottique (excitation, non-voisée) et du conduit vocal (fonction de transfert, voisée), sa représentation dans le domaine des fréquences consiste en une multiplication des deux signaux, ce qui produit une combinaison linéaire dans le domaine des quéfrences (Ex. 17).³⁹



$$\begin{aligned}
 \text{a : } & S(n) = E(n) * \theta(n) \\
 \text{b : } & |S(\omega)| = |E(\omega)| \cdot |\theta(\omega)| \\
 \text{c : } & \log|S(\omega)| = \log|E(\omega)| + \log|\theta(\omega)| \\
 \text{d : } & F^{-1} \log|S(\omega)| = F^{-1} \log|E(\omega)| + F^{-1} \log|\theta(\omega)|
 \end{aligned}$$

Ex. 17 : en haut, la schéma d'opérations effectuées afin de séparer le signal de la source glottique du celui du conduit vocal ; en bas, les formules correspondantes à chaque étape.⁴⁰

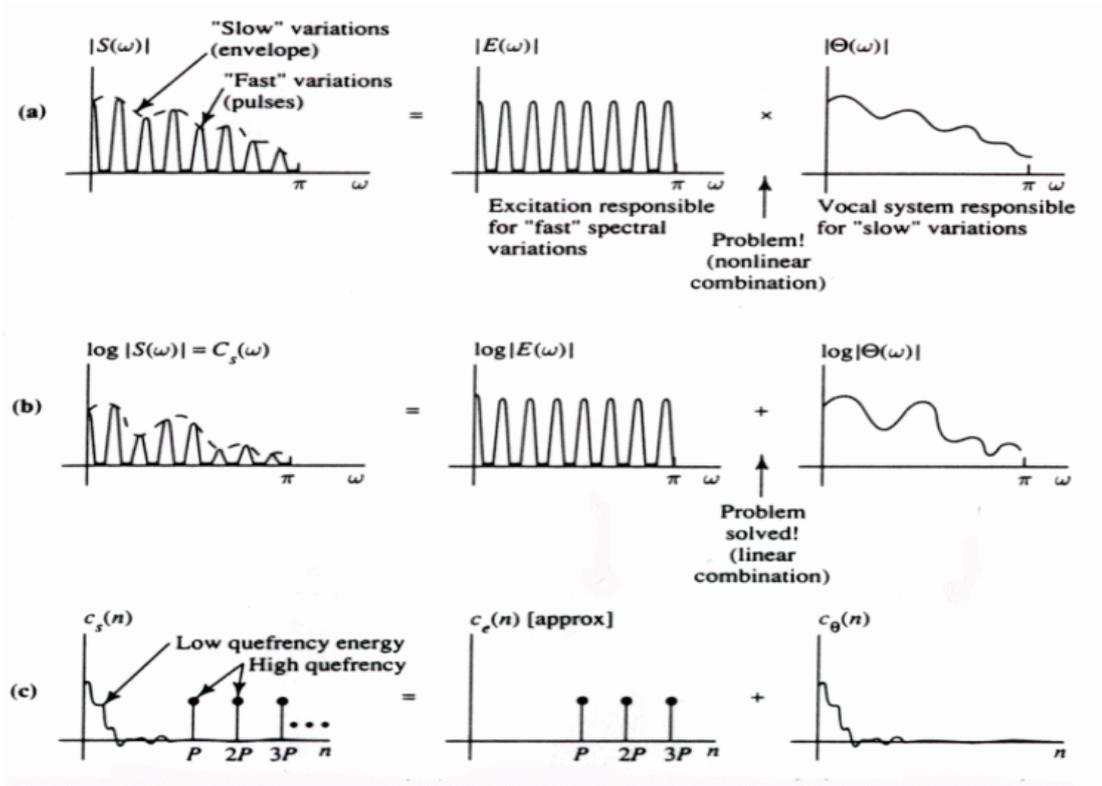
Ainsi sur la courbe du cepstre, les composants du conduit vocal qui varient lentement se trouvent concentrés dans des régions de basses quéfrences et les composants turbulents de la source glottique se trouvent dans des régions de hautes quéfrences. A partir de cette courbe, les transitoires peuvent être facilement séparés des parties stables du signal (Ex. 18)⁴¹.

Or nous pouvons aborder les méthodes de représentation temps-fréquence basées sur les MFCC, qui combinent plusieurs opérations liées aux termes discutés ci-dessus. Les coefficients cepstraux basés sur l'échelle de Mel d'un signal sont dérivés d'après le processus suivant, ou d'après des variations de celui-ci : d'abord, une transformée de Fourier est

39. VAN OSDOL, B. « Cepstrum », in *The Connexions Project*, p. 1-4.

40. *Ibidem*. p. 1-2.

41. CHILDERS, D. G. « The Cepstrum : a Guide to Processing », in *Proceeding of the IEEE* Vol. 65 No. 10, p. 1428-1443



Ex. 18 : représentation du signal vocal et ses composants : (a) spectre (b) logarithme du spectre (c) cepstre⁴²

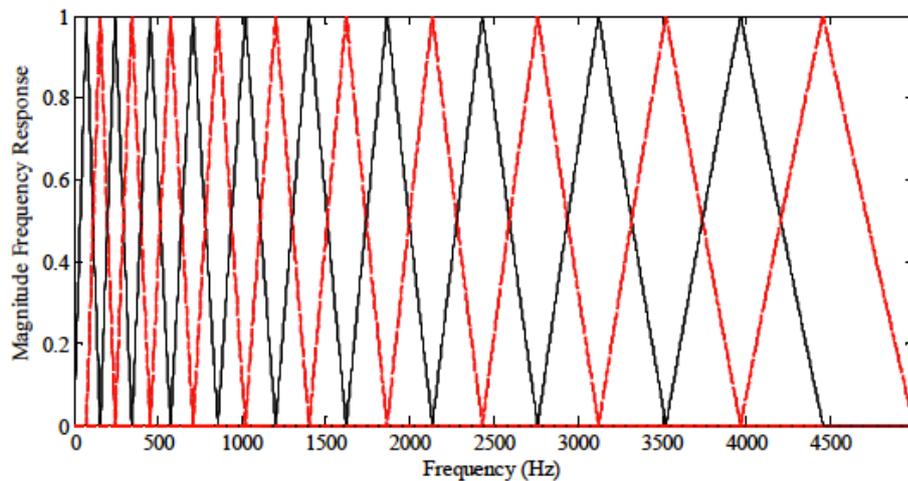
effectuée sur l'extrait du signal en question. Ensuite le spectre de puissance est mis à l'échelle de Mel au moyen d'un banc de filtres passe-bande triangulaires; les filtres sont conçus afin que leurs fréquences centrales correspondent à la répartition des bandes de fréquences produites par l'échelle de Mel, et se chevauchent d'une telle manière que leurs extrémités coïncident avec les fréquences centrales des filtres adjacents (Ex. 19). Agencé ainsi, le banc de filtres imite les bandes critiques⁴³. Ensuite, le logarithme du *spectre de mel* issu du dit filtrage est calculé et soumis à une transformée en cosinus discrète (TCD)⁴⁴. Les amplitudes du spectre qui en résulte sont les MFCC (Ex. 20).⁴⁵

42. JANG, R. *Audio Signal Processing and Recognition*, ch. 7-6.

43. Terme utilisé dans la psychoacoustique pour dénoter les filtres auditifs de l'oreille interne dont l'agencement suit des courbes de fréquence non-linéaires.

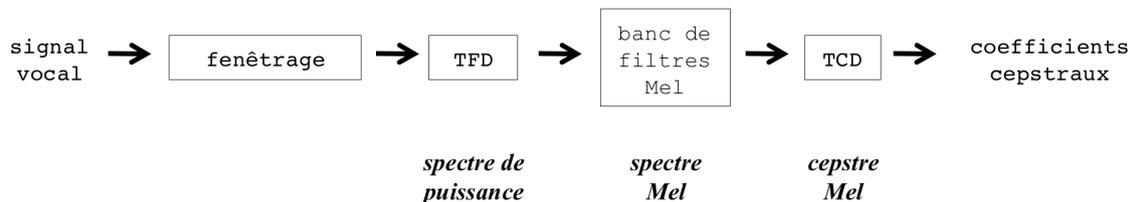
44. Transformation proche de la DFT qui décompose un signal en fonctions cosinoïdales uniquement. Elle produit donc des coefficients réels et non pas complexes, ce qui est particulièrement utile pour la compression du signal et de l'image.

45. SELTZER, M. « SPHINX III Signal Processing Front End Specification », p. 1-5.



Ex. 19 : Banc de filtres passe-bande triangulaires en échelle de Mel⁴⁶

Les MFCC jouent un rôle important dans le processus d'extraction de certaines caractéristiques sonores (feature extraction), ce qui constitue une étape importante des nombreux algorithmes de traitement vocal. Grâce à leur relation tangible avec la perception de la parole, les représentations temps-fréquence dérivées des MFCC pourraient avoir des applications utiles dans un travail artistique qui, malgré tout tentative de rationalisation, reste conditionné par la sensibilité humaine.



Ex. 20 : Processus d'extraction des coefficients cepstraux en échelle de Mel

II.3 APPLICATIONS INFORMATIQUES

En s'appuyant sur les informations exposées ci-dessus, nous avons identifié deux fonctions de l'analyse des sons transitoires qui pourraient être propices à notre travail artistique : la segmentation des mots en phonèmes et la représentation détaillée des informations spectrales et temporelles des dites phonèmes.

46 . RASETSHWANE, D. « Enhancement of Speech Intelligibility Using Speech Transients Extracted by a Wavelet Packet-Based Real-Time Algorithm », p. 36

II.3.1 LA SEGMENTATION : *SPHINX IV* ET *OP. RECOGNIZE*

La première fonction nous sert à isoler les parties d'un signal parlé qui nous intéressent en tant qu'unités sémiotiques dans le cadre d'un texte, ou bien en tant que phénomènes acoustiques. Puisque l'énoncé d'un mot contient plusieurs données auxquelles la sensibilité humaine est réceptive, dont le sens du mot, le caractère de la voix qui parle et les limites corporelles de la production vocale, une segmentation objective nous paraît difficile à réaliser en se fiant à l'audition uniquement. En revanche, la division d'un mot effectuée par un système indépendant des critères mentionnés ci-dessus pourrait nous proposer des possibilités sonores et conceptuelles auxquelles nous n'aurions pas accès autrement. Nous avons cherché donc un logiciel ou un algorithme pouvant servir à cette tâche.

Ayant parcouru les outils créés à cette fin, nous avons choisi d'employer le système de reconnaissance vocale à code source libre *Sphinx IV*, développé en java à l'université Américaine Carnegie Melon. Ce choix est dû principalement à l'existence d'une version de Sphinx portée sur Max/MSP, environnement auquel nous ferons appel dans la suite de nos recherches, d'une part car c'est un environnement ouvert et flexible qui permet d'organiser le travail de manière fluide, et d'autre part car il nous semble judicieux d'éviter d'éventuels problèmes de compatibilité ou de transfert de données entre systèmes.

Sphinx IV, comme la plupart des systèmes de reconnaissance vocale, l'analyse se fait en deux étapes : celle de l'extraction des caractéristiques, et celle de la modélisation probabiliste. L'extraction met en oeuvre un processus de dérivation des MFCC semblable à celui exposé dans la section précédente, en permettant à l'utilisateur de régler des paramètres d'échantillonnage, de fenêtrage et de filtrage afin d'obtenir la résolution la plus appropriée au signal traité. La deuxième étape s'appuie sur la prédiction probabiliste à partir des chaînes de Markov cachés⁴⁷ afin d'identifier les mots et les phonèmes contenus dans le signal. Ici, trois types de modèles de la parole sont employés : un modèle acoustique, un dictionnaire phonétique et un modèle de langue.⁴⁸

Le premier modèle contient les propriétés acoustiques des unités sous-phonétiques appelées *sénomes*⁴⁹. Le deuxième prend la forme d'un dictionnaire des phonèmes mettant en correspondance les mots et les phonèmes (Ex. 21).

47. Modèle statistique dans lequel le système modélisé est un processus stochastique de paramètres inconnus.

48. SELTZER, M. « SPHINX III Signal Processing Front End Specification », p. 1-5.

49. Composants élémentaires de la parole.

HELLO HH EH L OW
 GOODBYE G UH D B AY

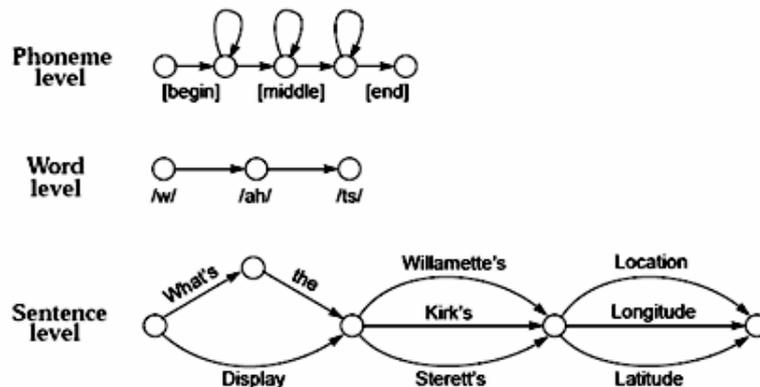
Ex. 21 : Extrait du dictionnaire phonétique

Cette mise en correspondance est peu précise à cause de sa capacité réduite à reconnaître des variations de prononciation; néanmoins, son emploi est pratique au sein d'un système reposant sur des couches de modélisation très nombreuses. Le troisième modèle définit des séquences de mots afin de restreindre le choix des mots possibles à la suite de mots déjà identifiés. Ce modèle est particulièrement efficace, mais est limité par le vocabulaire ainsi contenu. (Ex. 22).

(hello | goodbye) (I) (have) (lost | found) (my)
 (suitcase) * ;

Ex. 22 : Modèle de langue donnant les expressions régulières possibles : « Hello I have lost my suitcase » ou « Goodbye I have found my suitcase ».

Le processus de modélisation considère d'abord les unités les plus élémentaires de la parole, puis avance vers des agencements linguistiques d'une complexité croissante; il parcourt de cette manière plusieurs niveaux de perception de la parole (Ex. 23).



Ex. 23 : trois niveaux de modélisation probabiliste : phonème, mot, phrase.

A la suite de nos recherches, nous faisons emploi des dits procédés de reconnaissance proposés par *Sphinx* dans Max/MSP au moyen de l'objet externe (external) *op.recognize*⁵⁰ dans le dessein d'obtenir une segmentation en phonèmes d'un texte parlé. Les trois modèles élaborés ci-dessus sont utilisés sous la forme de fichiers : un fichier son pour le modèle acoustique et deux fichiers texte pour le dictionnaire phonétique et le modèle de langage. Nous illustrons l'application du système par l'exemple suivant, provenant de la bibliothèque de *Sphinx*.

Supposons que nous souhaitons isoler les phonèmes de la prononciation en anglais des chiffres « 1,9,8,4 », soit « one, nine, eight, four ». En tant que modèle acoustique, nous employons un enregistrement .mp3 fourni par la bibliothèque de *Sphinx* d'une voix d'homme lisant ces quatre mots (Aud. 9). En tant que dictionnaire phonétique, nous utilisons un fichier texte contenant des mots et les phonèmes qui les constituent. Ce fichier est organisé de la manière suivante: les mots entiers sont notés au début des lignes; les phonèmes, exprimées par une ou deux lettres grossièrement représentatives des règles de la prononciation anglaise, sont notées par la suite, séparées par une tabulation. Ainsi, les entrées du dictionnaire phonétique correspondant aux mots de notre modèle acoustique sont notées comme suit:

ONE	W AH N
NINE	N AY N
EIGHT	EY T
FOUR	F AO R

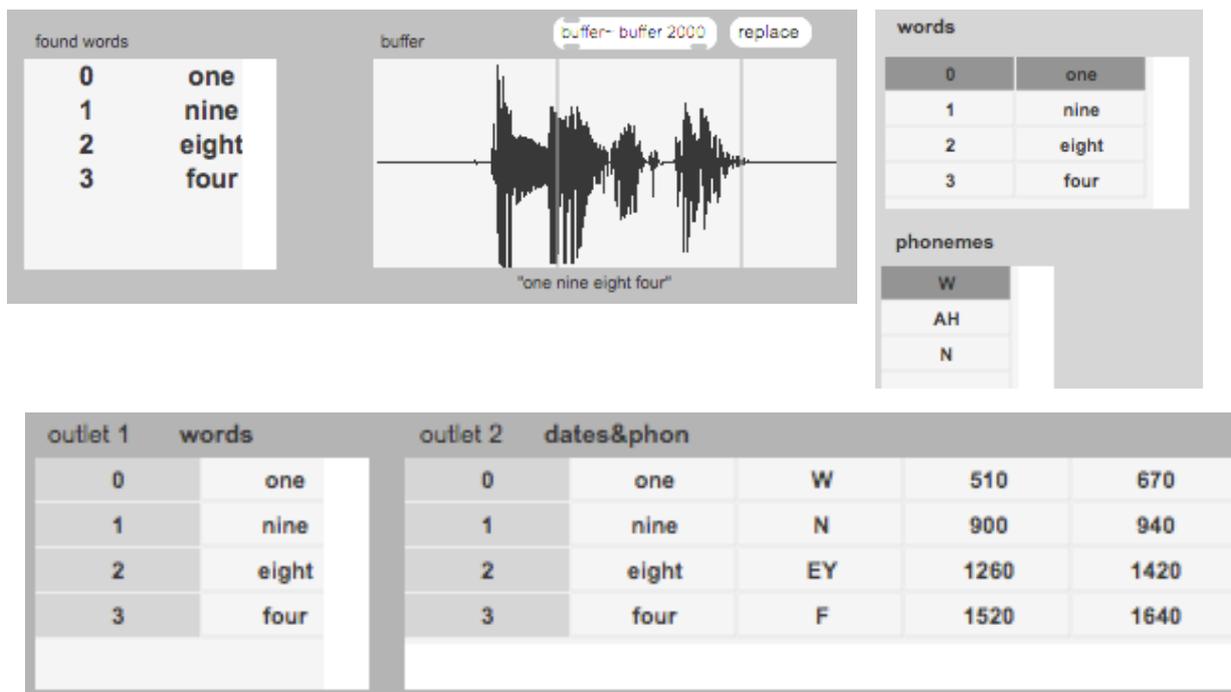
En tant que modèle de langue, nous employons un fichier texte .gram. Celui-ci contient un groupe de mots auquel appartiennent tous les mots du modèle acoustique parmi d'autres. Dans le cas de notre exemple, le fichier .gram exprime tous les chiffres de zéro à neuf :

```
public <numbers> = (zero | one | two | three | four |  
five | six | seven | eight | nine) * ;
```

Dans l'interface Max/MSP de *op.recognize*, nous sélectionnons les fichiers du modèle acoustique, du dictionnaire phonétique et du modèle de. Ensuite l'objet externe réalise la segmentation à partir des prédictions basées sur les modèles fournis. Les résultats sont présentés dans le sous-patch « phonemes & more »: une liste de tous les mots trouvés dans le fichier sont en ordre chronologique, accompagnée par les segments audio correspondant aux

50. Objet externe en java développé en 2007 par le compositeur de musique électronique Olivier Pasquet.

mots affichés (Ex. 24 Aud. 10). A partir de cette segmentation partielle, *op.recognize* nous procure une liste de phonèmes reconnus dans chaque mot ainsi que leurs données de localisation dans le buffer. En cliquant sur chaque mot affiché, nous obtenons également une liste de phonèmes permettant d'écouter les segments audio respectifs (Aud 11.).



Ex. 24 : La segmentation de la phrase « one nine eight four ».

Bien que cet exemple traite d'une phrase énoncée en anglais dont le contenu est peu pertinent à notre propos, nous trouvons les résultats générés par *op.recognize* dans cette simple illustration fructueux à un travail artistique. Considérons les segments audio dérivés du phonème N par exemple: pour le modèle acoustique « one, nine, eight, four », nous obtenons trois sons avec des enveloppes dynamiques et des évolutions fréquentielles distinctes (Aud. 12). Ainsi, nous ne pouvons pas traiter le symbole écrit N comme un élément structurel abstrait dont la signification est engendrée uniquement par la répétition et sa localisation dans le texte; il existe aussi des propriétés expressives propres à la parole qui sont transmises à travers la lettre N, indépendamment du sens sémiotique de celle-ci. L'emploi d'un outil informatique pour la réalisation de ce travail nous révèle donc une disparité fondamentale

entre la pensée formalisante abstraite inhérente à l'écriture classique et la pensée concrète répondant à l'observation sensible d'un phénomène trop complexe pour être ainsi systématisé.

II.3.1.1 APPLICATIONS

Partant de cet exemple, nous avons adapté *op.recognize* à notre propre emploi: nous avons récréé les trois modèles à partir d'un texte Hébreu extrait du cycle des poèmes « Leçons d'hébreu » de Hezy Leskly, texte utilisé dans la composition musicale entreprise dans la phase ultime de notre travail⁵¹. En premier lieu nous avons créé deux modèles acoustiques en enregistrant le texte lu par une voix d'homme et par une voix de femme (Aud. 13). Bien que la composition qui en résulte ait été conçue pour ensemble instrumental et voix de femme, il nous paraît astucieux dans ce premier stade de recherche de ne pas exclure les possibilités acoustiques unique à la voix d'homme qui pourraient éventuellement produire des effets étonnants.

Ensuite, nous avons construit un dictionnaire phonétique pour chaque vers du texte contenant exclusivement les mots appartenant au vers en question. Le premier vers du premier poème par exemple est représenté comme suit:

HAREGA	HH AH R EH G AH
ANI	AH N IH
LOMED	L AA M EH D
LIKHTOV	L IH K T AA V

Ici nous constatons une divergence importante de notre modèle idéal: le mode d'expression de phonèmes dans le code de *Sphinx* (employé également dans *op.recognize*) correspond à une approximation de la prononciation anglaise; or, l'hébreu comprend des sons qui n'existent pas dans la langue anglaise. Malgré ce défaut, nous avons décidé de poursuivre cette tentative: d'après notre expérience antérieure avec des systèmes informatiques, une marge d'erreur limitée produit par l'usage d'un outil à une fin à laquelle il n'est pas destiné engendre souvent une ouverture à l'innovation. Toutefois, nous tenons compte de l'éventualité de devoir compenser cette limitation dans une étape ultérieure de notre travail.

51. Le choix de texte ainsi que son sens et traitement sont élaborés dans la section

Finalement nous avons construit le modèle de langue. Comme dans le dictionnaire phonétique, nous avons limité ce modèle aux mots appartenant à un seul vers. Ainsi, le modèle pour le premier vers est donné par:

```
public <numbers> = ( HAREGA | ANI | LOMED |
LIKHTOV ) * ;
```

Ces modèles étant établis, nous avons réalisé la séparation des vers en mots. Les résultats donnés par *op.recognize* sont présentés ci-dessous:

Vers : HAREGA ANI LOMED LIKHTOV
Mots : HAREGA ANI ANI LIKHTOV (Ex. 25 ; Aud. 14)

outlet 1	words	outlet 2	dates&phon			
0	harega	0	harega	HH	60	320
1	ani	1	ani	AH	1700	1830
2	ani	2	ani	AH	2580	2630
3	likhtov	3	likhtov	L	3830	3890

Ex. 25 : La segmentation en mots de la phrase « Harega ani lomed likhtov ».

Vers : ILMALE HAMILA HAATUFA BEARIG SHAKHOR (Aud. 15)
Mots : HAMILA HAMILA HAATUFA BEARIG SHAKHOR
(Ex. 26 ; Aud. 16)

outlet 1	words	outlet 2	dates&phon			
0	hamila	0	hamila	HH	130	160
1	hamila	1	hamila	HH	800	1040
2	haatufa	2	haatufa	HH	1700	1770
3	bearig	3	bearig	B	2930	3040
4	shachor	4	shachor	SH	3900	3990

Ex. 26 : La segmentation en mots de la phrase « Ilmaleh hamila haatufa bearig shachor ».

Dans les deux exemples nous pouvons constater des erreurs dans l'identification des mots (soulignés au dessus) qui se sont reproduit indépendamment du genre de la voix employée dans le modèle acoustique. Mais en écoutant les fichiers son, nous remarquons que

la segmentation est réussie, bien que l'identification ne soit pas exacte. Ainsi nous avons pu avancer à l'étape suivante, la segmentation des mots en phonèmes avec *op.recognize*, dont quelques résultats sont donnés ci-dessous :

Mot : HAREGA
Phonèmes : HH AH R EH G AH (Aud. 17)

Mot : LIKHTOV
Phonèmes : L IH K T AA V (Aud. 18)

Mot : HAATUFA
Phonèmes : HH AH AH T UH F AH (Aud. 19)

Mot : BEARIG
Phonèmes : B EH AH R IH G (Aud. 20)

Mot : SHAKHOR
Phonèmes : SH AH K HH AA R (Aud. 21)

En écoutant les fichiers son ainsi générés, nous observons que parmi les phonèmes riches en sons transitoires, les consonnes fricatives⁵² (KH ou [x]; V ou [v]; F ou [f]; SH ou [ʃ]; Aud. 22) rendent les résultats sonores les plus fidèles. La segmentation des consonnes occlusives⁵³ (G [g]; T [t]; B [b]; Aud 23) par contre donne des résultats peu fiables : tantôt la consonne est rattaché à la voyelle précédente, tantôt à la voyelle suivante. Ce classement erroné est dû au fait que le relâchement du blocage d'air dans les consonnes occlusives engendre inévitablement un son voyelle plus ou moins court.

Ainsi, à l'issu de notre travail avec *op.recognize*, nous considérons que l'objet nous sera utile dans la suite de nos recherches mais que son emploi sera moins conséquent pour l'isolation des consonnes occlusives.

II.3.2 LA REPRESENTATION TEMPS-FREQUENCE AVEC MAX/MSP

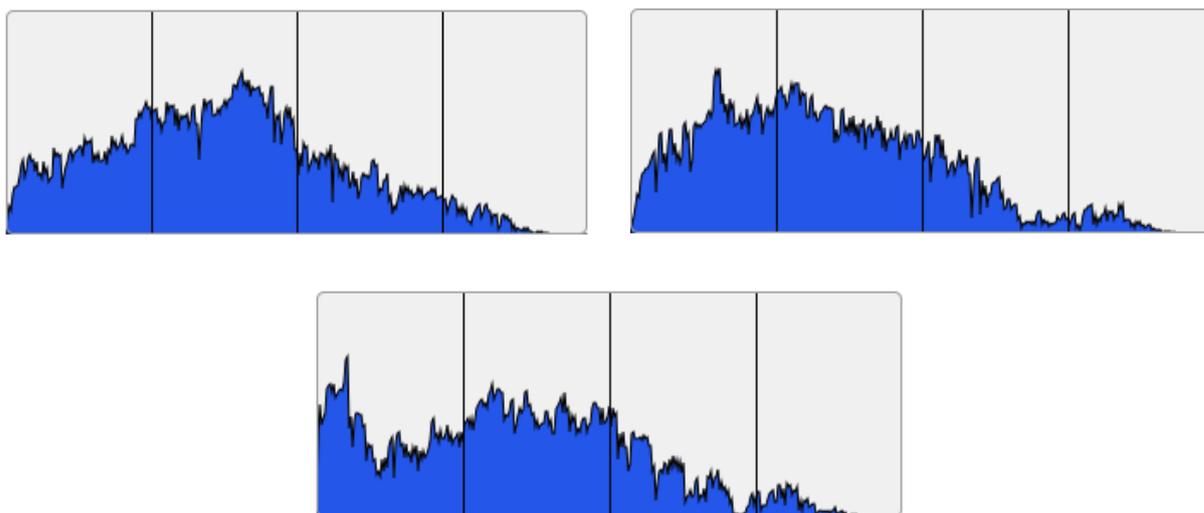
Ayant isolé des phonèmes qui nous intéressent, nous avons abordé l'analyse temps-fréquence des dits phonèmes, et ce dans le but de pouvoir les différencier visuellement. En employant un moyen de représentation dont la résolution est suffisamment élevée, nous

52. Consonne tenue produit par le resserrement du chenal expiratoire sans le fermer complètement. Voir annexe 6.

53. Consonne percussive produit par un blocage complète de l'écoulement d'air suivi par son relâchement soudain.

pourrions ensuite comparer des phonèmes avec des sons instrumentaux afin d'identifier les sons ayant des propriétés acoustiques semblables. Nous avons donc testé des outils s'appuyant sur les trois méthodes de représentation exposées antérieurement et applicable dans Max/MSP pour obtenir des représentations des phonèmes isolés avec *op.recognize*.

Le premier de ces outils est l'objet *spectroscope~* de Cycling '74, objet employant la transformée de Fourier à court terme afin d'afficher des informations spectrales et temporelles. En analysant les consonnes fricatives [s], [ʃ] et [x] (Ex. 27), nous avons trouvé la résolution insuffisante à nos fins : bien qu'une distinction entre ces trois phonèmes peut être observée, elle nous paraît trop subtile pour constituer une base fiable pour la suite de nos recherches. Sachant que la correspondance entre sons instrumentaux et phonèmes ne sera pas parfaite, il nous est important d'établir un maximum de différenciation lors de ce premier stade afin d'assurer que l'étape suivante de notre travail soit fructueuse.



Ex. 27 : représentations obtenues avec l'objet *spectroscope~*.

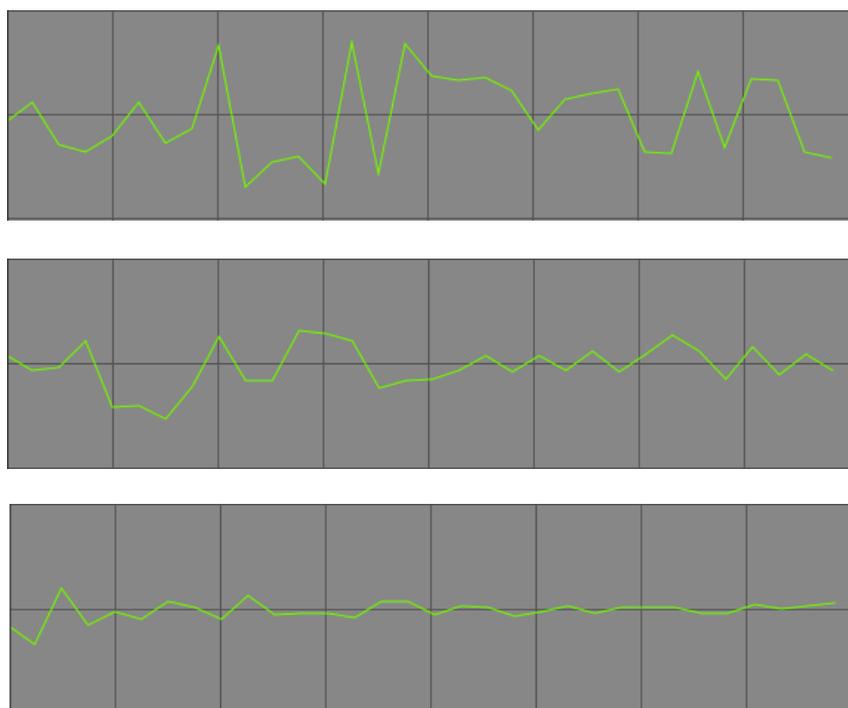
En haut à gauche, [s] ; en haut à droite, [ʃ] ; en bas, [x]

Nous avons par la suite utilisé l'objet *wavelet~* du CNMAT⁵⁴ afin d'obtenir une représentation dérivée de la transformée en ondelettes discrète. Mais l'emploi de cet outil nous a posé certaines contraintes : les représentations proposées par *wavelet~* varient considérablement en fonction des paramètres sélectionnés. Ces paramètres, avec la famille des ondelettes (en l'occurrence *Haar*, *Daubechies* ou *B-spline*) et la mesure des moments nuls⁵⁵,

54. *Center for New Music and Audio Technologies* à l'université de Californie Berkley.

55. Données importantes permettant à mesurer la régularité locale d'un signal.

doivent être modifiés pour chaque phonème afin d'obtenir la résolution idéale. Comme les rapports entre les dits paramètres et les sons traités sont non-linéaires et très complexes, leur maîtrise nécessite des connaissances mathématiques sur les ondelettes qui dépassent largement le cadre de notre travail. Il nous a fallu donc tâcher d'arriver aux résultats conséquents et stables de manière empirique. Les représentations ainsi obtenues pourraient nous servir dans la suite de notre travail, mais manifestent néanmoins un taux de différenciation plus faible et moins régulier que nous l'espérons. (Ex. 28)



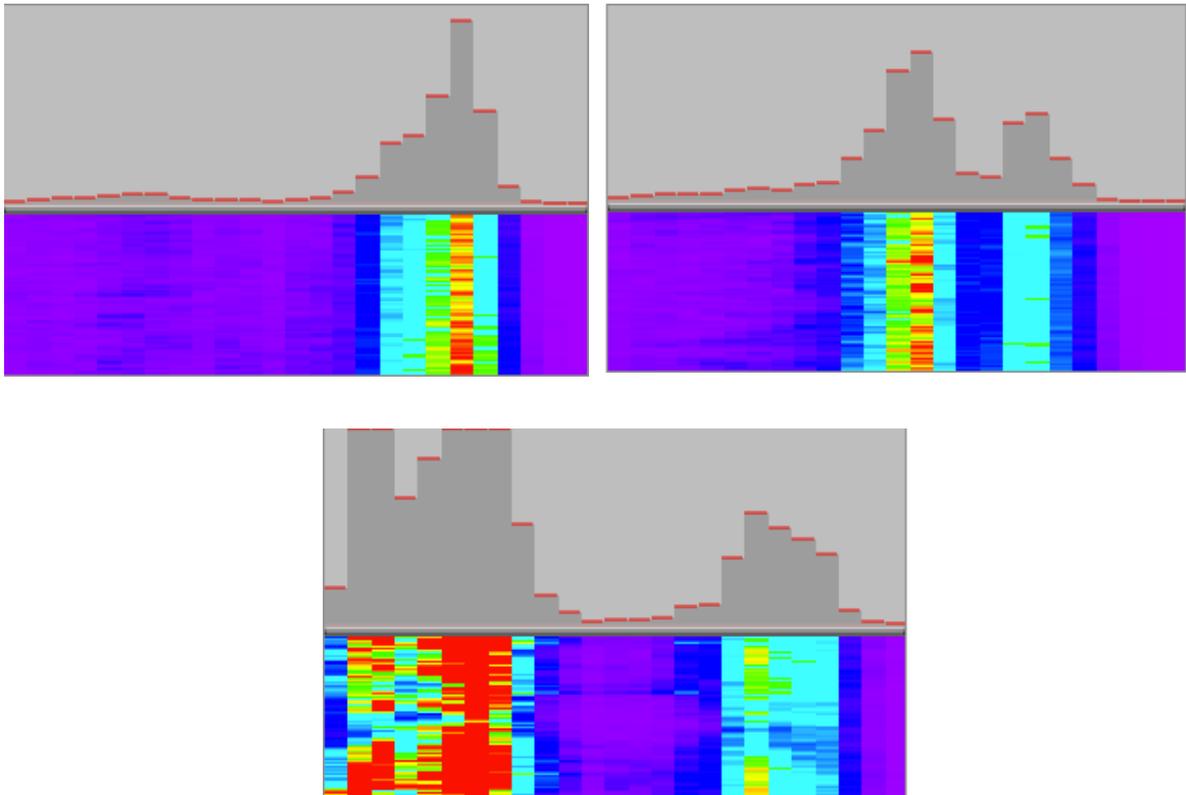
Ex. 28 : représentations obtenues avec l'objet *wavelet~*.

En haut, [s] ; au centre, [ʃ] ; en bas, [x]

Par conséquence, nous avons abordé la troisième méthode de représentation, soit les coefficients cepstraux en échelle de Mel. A cette fin, nous avons employé l'objet *Crossmel~* de la bibliothèque *Gabor*⁵⁶, qui produit le spectre Mel d'un son en le filtrant au moyen du banc de filtres décrit dans la section II.x.y. En utilisant cet outil pour l'analyse des trois phonèmes mentionnés au dessus, nous avons obtenu trois représentations assez distinctes (Ex. 29). Nous trouvons que ces représentations sont d'une résolution suffisamment haute et d'un niveau de différenciation suffisamment élevé pour que nous puissions les employer comme base de

56. Partie de la bibliothèque *FTMLib* développé à l'IRCAM.

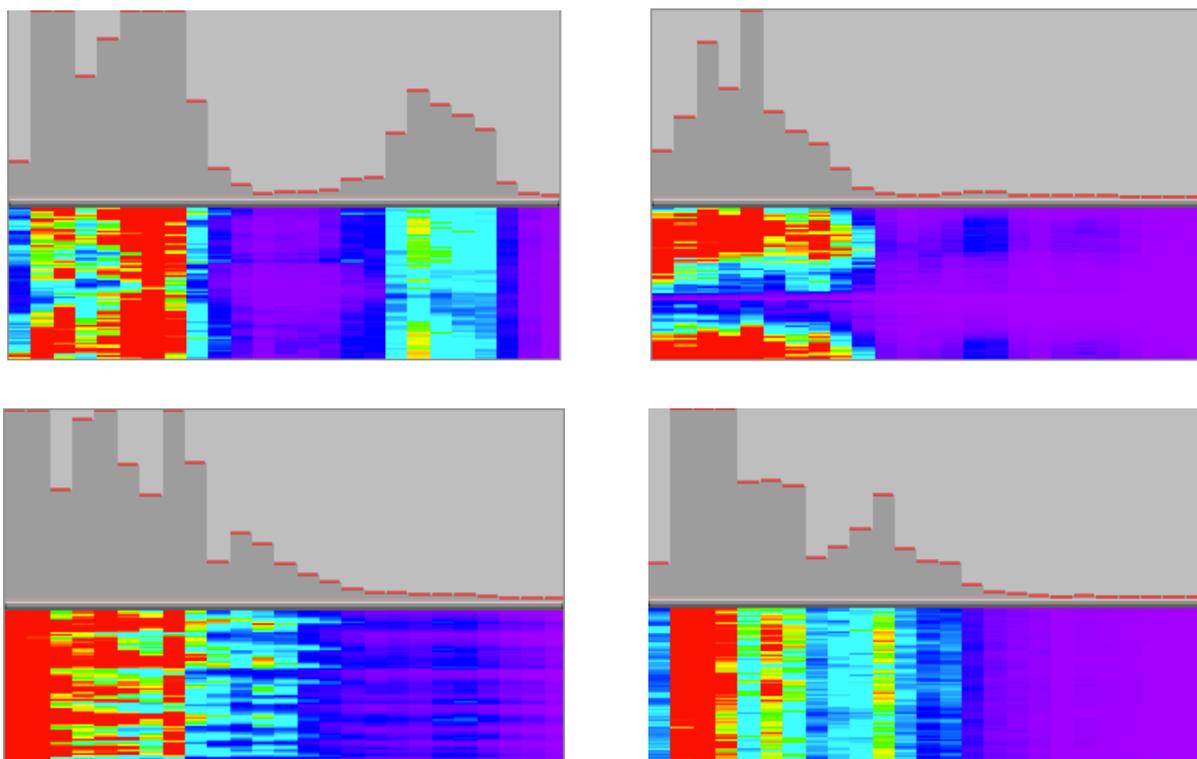
comparaison avec des sons instrumentaux. Ayant vérifié que les résultats ainsi générés soient stables et reproductibles, nous avons déterminé que cette méthode est satisfaisante à nos exigences.



Ex. 29 : représentations obtenues avec l'objet *crossmel~*.
En haut à gauche, [s] ; en haut à droite, [ʃ] ; en bas, [x]

II.3.2.1 APPLICATIONS

Ayant décidé que l'objet *crossmel~* soit le plus approprié à nos buts, nous l'avons utilisé afin d'analyser des sons instrumentaux et de les comparer aux phonèmes vocaux. Comme nous avons mentionné auparavant, nous entreprenons ce travail en sachant qu'une correspondance parfaite entre son instrumental et phonème vocal n'existe guère. Nous cherchons plutôt d'identifier des sons instrumentaux dont les spectres Mel partagent suffisamment des qualités quantifiables avec ceux les dits phonèmes pour qu'une association des deux soit envisageable. Certains résultats de cette tentative que nous avons trouvé fructueux sont affichés ci-dessous.



Ex. 30 : En haut à gauche, la consonne fricative [x] ; en haut à droite, *flutterzunge* à la flûte avec l'embouchure fermée ; en bas à droite, corde grave filée du piano éraflée avec l'ongle ; en bas à gauche, son écrasé au violon.

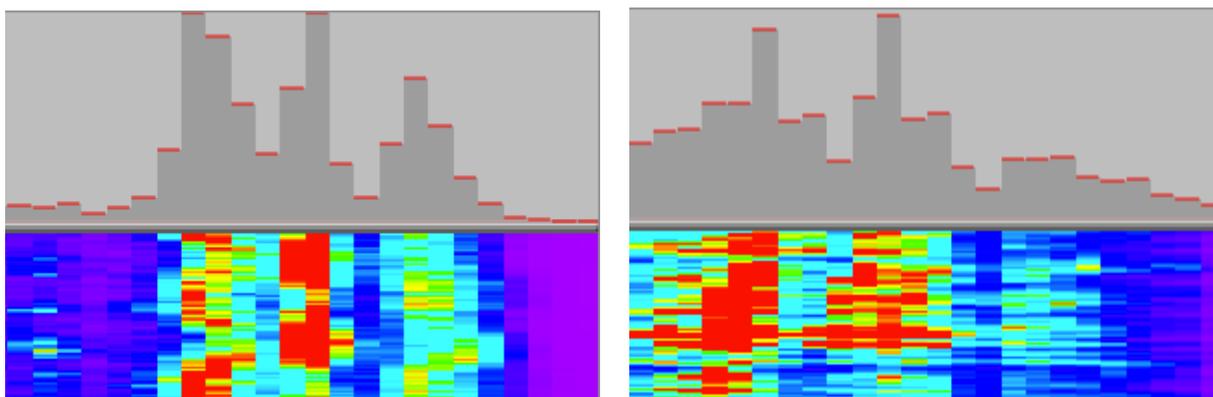
Ici nous avons comparé la consonne fricative [x] (notée KH dans le dictionnaire phonétique employé pour la segmentation) avec trois sons instrumentaux : le son *flutterzunge*⁵⁷ effectué en fermant l'embouchure avec la bouche ; le son produit par l'éraflure d'une corde grave filée du piano ; le son *écrasé*^{58 59} au violon (Aud. 24). En regardant les analyses ci-dessus, nous constatons que lesdits sons instrumentaux et la consonne [x] ont des propriétés communes : une composition fréquentielle concentrée dans les régions graves du spectre Mel ainsi qu'une texture granulaire-itérative indiquée par les bandes de couleur finement striées. Bien que les composants aigus qui se manifestent dans le spectre Mel de la consonne [x] (en haut à gauche) soient peu présents dans les sons de la flûte et du piano, et

57. Un coup de langue répété à une cadence très rapide produisant un effet de tremolo.

58. Son produit par l'application d'une pression excessive avec l'archet. L'archet n'avance que peu sur l'axe horizontal, et reste collé à la corde afin d'étouffer sa réverbération.

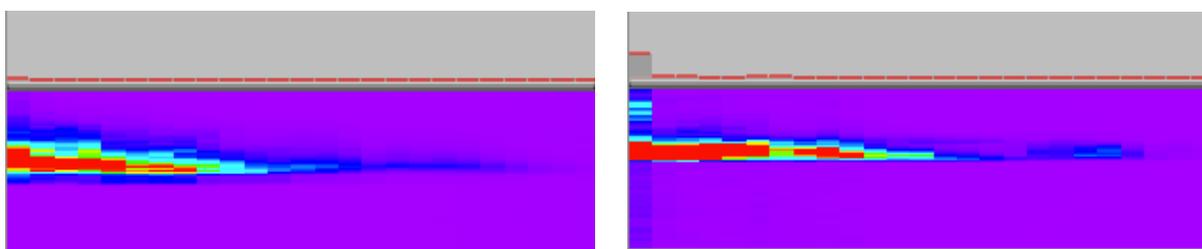
59. Nous traitons ici les sons du violon comme étant représentatifs des sons qui peuvent être obtenus avec d'autres instruments à cordes, notamment l'alto et le violoncelle.

manquent complètement dans le son du violon, nous trouvons qu'ils existe suffisamment de similitudes entre ces sons pour qu'ils soient mise en corrélation dans le cadre de notre travail musical.

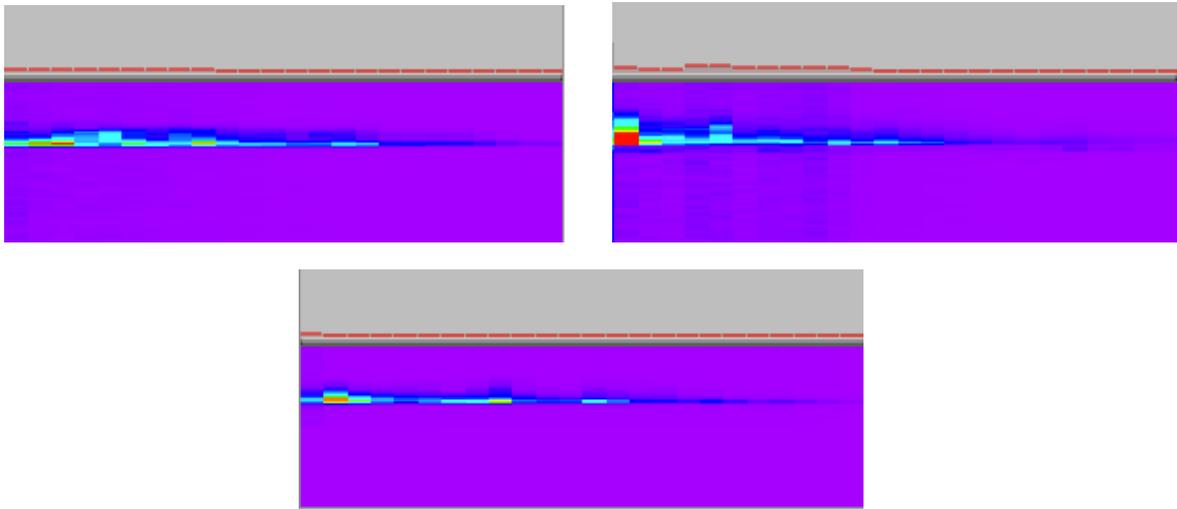


Ex 31 : à gauche, la consonne [j] énoncé en le filtrant avec le mouvement des lèvres ; à droite, trémolo au violon joué *flautando* sur la touche en modifiant la hauteurs avec la main gauche.

Dans cet exemple la consonne [j] (SH dans le dictionnaire phonétique) est comparée avec le son trémolo du violon joué *flautando* et haut sur la touche (Aud. 25). Par ailleurs, les deux sons subissent des variations de hauteur, effectuées par les lèvres dans le premier des cas et par le mouvement de la main gauche dans l'autre. Nous observons ici que des régions importantes des spectres Mel des deux sons se chevauchent et que les textures des deux représentations s'approchent.



Ex 32 : à gauche, la consonne occlusive voisée [d] ; à droite, *tongue ram* à la flûte.



Ex 33 : en haut à gauche, la consonne occlusive non-voisée [t] ; en haut à droite, *pizzicato* à la flûte ; en bas, *col legno battuto* au violon.

Dans les exemples 32 et 33 nous comparons les consonnes occlusives alvéolaires⁶⁰ [d] (D dans le dictionnaire phonétique) et [t] (T dans le dictionnaire phonétique) avec des sons instrumentaux (Aud. 26). Bien que les deux soient articulées de la même manière, nous constatons une différence tangible entre la première qui est voisée, et la deuxième qui est non-voisée. Nous trouvons ainsi que la consonne [t] soit semblable aux sons instrumentaux percussifs d'une résonance limitée, comme le *pizzicato*⁶¹ à la flûte et le *col legno battuto* avec les cordes étouffées au violon. En revanche, la consonne [d] ressemble aux sons percussifs plutôt résonant, comme le *tongue ram*⁶² à la flûte.

II.4 RESUME

Dans cette section de nos recherches, nous avons exploré trois méthodes servant à l'isolation des phonèmes au sein d'une phrase et puis, à l'analyse des leurs composants transitoires. Nous avons trouvé que l'objet externe Max/MSP *op.recognize* nous permet de

60. Consonnes articulées avec soit la pointe, soit la lame de la langue contre la crête alvéolaire.

61. Son produit par des poussées d'air générées avec la langue uniquement, sans que la pression d'air supplémentaire provienne des poumons.

62. Son produit en obturant l'embouchure avec la bouche et projetant violemment la langue.

segmenter des phrases et des mots de manière fiable. Les phonèmes isolés de cette façon peuvent en suite être analysés afin de produire une représentation temps-fréquence.

Pour cette deuxième tâche, nous estimons que l'objet *Crossmel~*, objet employant des filtres passe-band positionnés selon l'échelle de Mel, rend les représentations les plus convenables à nos fins. Les représentation des sons consonnes obtenues avec *Crossmel~* ont été comparées avec les représentations des sons instrumentaux obtenus de la même manière, dans le but d'identifier des sons instrumentaux ayant des qualités acoustiques communes avec lesdits sons vocaux. Bien que les résultats de cette démarche n'aient pas été entièrement conséquents, nous pourrions les employer afin de définir des sons instrumentaux analogues aux sons vocaux. De plus, les notions des propriétés acoustiques des sons vocaux ainsi que la possibilité de savoir de quelle manière elles sont perçues acquises lors de nos recherches, seront sûrement utiles dans la suite de notre travail.

Nous constatons par ailleurs que les deux outils jugés être les plus propices pour la réalisation d'une pièce musicale s'appuient sur l'échelle de Mel, et ainsi sur la perception. Cette observation ne nous étonne pas : notre compréhension de l'expression vocale est tout d'abord empirique, venant de nos expériences communicatives quotidiennes. Pour les outils cherchant à quantifier cette compréhension, la prise en compte de la sensibilité humaine est indispensable.

III. REALISATIONS MUSICALES

En s'appuyant sur les recherches et les analyses exposées dans les sections précédentes, nous avons entrepris la réalisation d'une composition musicale dont la syntaxe instrumentale est issue des caractéristiques intrinsèques de la vocalisation d'un texte. La première étape de ce travail prend la forme d'un exercice: la transcription pour flûte basse d'une pièce vocale. Cette étude nous a permis d'évaluer les résultats sonores des méthodes employées dans un cadre cohérent; les méthodes jugées réussies furent l'inspiration pour la composition principale.

Par la suite, nous avons envisagé une pièce pour ensemble instrumental et voix explorant cette forme de gestualité instrumentale d'une manière plus libre, tout en maintenant un lien, tantôt explicite, tantôt subtil, avec le texte et avec les qualités acoustiques-phonétiques de la voix. Ainsi, guidés par les notions acquises au fil de nos recherches, nous avons réalisé *Leçons d'hébreu*, œuvre consistant en six courts mouvements basés sur des poèmes de l'écrivain israélien Hezy Leskly.

III.1 *AUTO-DETERMINATION* : TRANSCRIPTION SUR UN TEXTE CACHE

III.1.1 LE PROJET ET LES CHOIX ARTISTIQUES

Dans une première tentative de façonner des combinaisons sonores instrumentales, deux notions nous ont paru fondamentales: la cohérence et la sensibilité musicale. Afin de générer des gestes satisfaisant ces notions, nous avons choisi d'employer comme socle la pièce *Auto-détermination* pour voix de femme et électronique en temps réel, composée dans le cadre de nos recherches antérieures.

Ce choix nous présente deux avantages principaux. Basée sur un texte caractérisé par l'agencement sonore des phonèmes⁶³, l'écriture vocale dans *Auto-détermination* isole des éléments articulatoires et les associe avec des éléments musicaux correspondants à leurs qualités acoustiques; ainsi, dans la traduction des sons vocaux en sons instrumentaux se manifesterait une logique propre, qui serait néanmoins distincte de l'écriture instrumentale. Parachevés en collaboration étroite avec la chanteuse Armelle Orioux, les critères musicaux

63. Le poème « Auto-détermination » de l'écrivain roumain Gherashim Luca, publié en 1953.

(comme les phrasés, les nuances et le discours global) firent l'objet d'une attention particulière de la part du compositeur et de l'interprète (Ex. 34); ainsi, l'intention musicale enracinée dans la pièce se transmettrait forcément, même à travers un autre instrument.

The image shows a musical score for a flute in G major. It begins with a box containing the number '2'. The tempo is marked 'a tempo' and the character '(accusatrice)'. The dynamics start at 'p' (piano) and progress through 'mp' (mezzo-piano), 'mf' (mezzo-forte), and 'f' (forte). The tempo marking 'accel. poco a poco fino a 4' is placed above the staff. The score consists of a single line of music with various note values, rests, and slurs. Below the staff, the lyrics are written: 'sa ma-nie de s-oi-oi sa man ière de oi - oi - oi - oi le soir d-d-de s' ass - oi - r le'.

Ex. 34 : extrait de *l'Auto-détermination* montrant le niveaux de précision des détails musicales.

Ayant déterminé la source de notre transcription, il s'est agi de désigner l'instrument pour lequel *Auto-détermination* serait transcrite. La flûte traversière en ut nous a semblé initialement bien adaptée à nos fins car la production sonore est semblable à celle de la voix chantée. Par ailleurs, l'appareil articulatoire reste libre, sans être entravé par l'anche ou l'embouchure, ce qui n'est pas le cas pour maint autres instruments à vent; par conséquent, la richesse d'articulation de la voix peut être préservée. Cependant, bien que les similarités entre flûte et voix permettent d'établir une relation plus conséquente entre leurs mondes sonores respectifs, elles empêchent de songer à une syntaxe appropriée à la flûte qui aura une existence indépendante dans une étape plus avancée de notre travail. Notre objectif étant de développer une écriture instrumentale novatrice et non pas d'imiter la voix, nous trouvons que la flûte en ut est un médium trop restreint.

Nous avons donc décidé d'employer la flûte basse afin de profiter des dits avantages de la flûte en ut, tout en ayant à notre disposition une palette sonore bien éloignée de celle de la voix de femme. La robustesse et la variété timbrale de sons percussifs, la richesse des harmoniques, la rondeur du registre médium-grave ainsi que la fragilité du registre aigu, forment une identité sonore unique qui peut être néanmoins manœuvrée afin de produire des sons manifestement liés à la voix.

III.1.2 L'ECRITURE INSTRUMENTALE

Notre travail de transcription se focalise dans une première étape sur la recherche des équivalents sonores entre flûte et voix. Nous déterminons deux approches qui nous seront utiles: une approche phonétique-acoustique, et une approche structurelle-globale. La première traite les phonèmes selon des critères largement objectifs, hors contexte de la pièce. Nous appliquons cette approche d'une manière plutôt libre, en faisant confiance à nos oreilles et à notre intuition pour le son, tout en s'appuyant sur les recherches menées dans la deuxième partie de notre travail. La deuxième approche tient compte de la fonction de chaque son au sein de l'œuvre entière. Nous l'employons afin de s'assurer que nos choix soient adaptés au discours musical de la pièce.

III.1.2.1 APPROCHE PHONETIQUE-ACOUSTIQUE

La consonne [d], consonne occlusive voisée, est le phonème le plus percussif d'*Auto-détermination*. Dans la version d'origine elle:

- est parlée (et non-pas chantée) avec la voix de poitrine et articulée d'une manière très percussive
- marque souvent l'accent d'une phrase
- est parfois répétée rapidement afin de briser la rigidité rythmique des cellules

Ainsi, un son équivalent à la flûte devrait:

- être bruyé, sans hauteur clairement distinguable, et d'une qualité percussive
- avoir une pesanteur très importante
- être facile à jouer plusieurs fois de suite, à grande vitesse

Le premier critère ci-dessus pourrait être satisfait par plusieurs modes de jeu: le claquement des clés, le pizzicato, le *tongue ram*, et nombreuses variations sur la note piquée mélangée avec du souffle. En considérant le deuxième critère, nous avons jugé le *tongue ram* le plus conforme à nos buts: ayant une force et une présence assez exceptionnel dans la gamme des sons typiquement produit par l'instrument, ce mode de jeu désignerait clairement les points d'appui. Par ailleurs, grâce à sa profondeur presque corporelle, le *tongue ram* ressemble au geste et au son de [d] annoncé en voix de poitrine. Notre choix cependant ne

satisfait point le troisième critère. Puisqu'il faut tourner la flûte et boucher l'embouchure afin de produire un *tongue ram*, ce son ne pourrait pas être rapidement réitéré, ni facilement suivi par tout son employant l'embouchure. Nous avons ainsi décidé d'utiliser les claquements des clés suite au *tongue ram*: bien qu'elles soient moins puissantes, elles évoquent l'effet vocal avec la précipitation voulue.

Le résultat de ce processus de traduction de [d] est donc la création d'un geste à la flûte composé d'un *tongue ram* suivi par un mouvement fougueux de claquement des clés (Ex 35 ; Aud. 27). Ceci constitue un objet sonore clairement aménagé pour la flûte basse d'une composition qui, bien qu'elle soit assez simple, ne nous aurait pas paru évidente sans le modèle vocal initial.



Ex. 35 : Geste dérivé de la consonne [d].

Ensuite nous traitons les consonnes fricatives non-voisées [s] (sifflantes) et [ʃ] (chuintantes), qui sont employées dans *Auto-détermination* de manière à mettre en relief leur caractère continu et bruité. Elles sont:

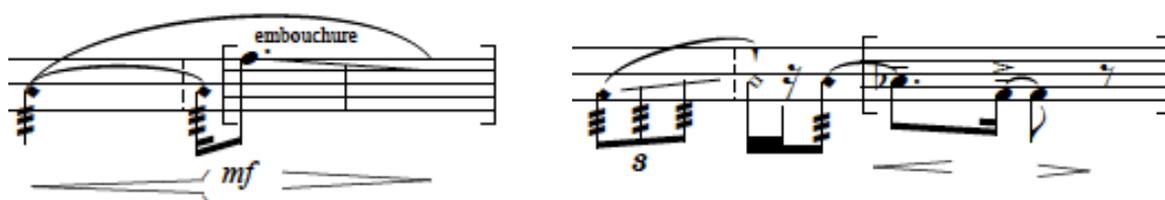
- tenues, afin que la turbulence du spectre soit audible
- peu résonantes, et exécutées avec un crescendo qui est pourtant peu perceptible, en exagérant leur nature inharmonique
- sans hauteur définie mais soumises à un effet de filtrage effectué avec les lèvres

Un son analogue à la flûte serait joué en:

- agitant la vibration habituelle de l'instrument
- étouffant la résonance du tuyau, d'une manière à atténuer l'effet d'un crescendo
- changeant les pics de résonance du son, sans qu'une hauteur précise soit distinguable

En considérant ces qualités, il nous a paru évident d'employer le *flatterzunge* à la flûte afin de rendre la vibration moins régulière et donc plus semblable au spectre turbulent produit par les fricatives non-voisées. Pour étouffer le son, nous avons choisi de bloquer l'embouchure intégralement, car elle est le point le plus résonant de l'instrument. Ainsi le son du *flatterzunge* semble émaner de l'intérieur de la flûte, sa capacité de projeter le son étant largement réduite. Ayant établi ce son peu résonant et peu harmonique, nous avons pu modifier le spectre en actionnant les clés, sans générer ni des hauteurs précises, ni une notion harmonique-mélodique.

Puisque dans la version d'origine les sifflantes et chuintantes sont souvent suivies par des voyelles chantées (que nous traduisons simplement comme des sons tenus à la flûte), le geste final généré par notre transcription est constitué d'un son inharmonique peu puissant, suivi par un son harmonique révélant la résonance de la flûte basse (Ex. 36 ; Aud 28). Ce passage extrême entre son bruité et hauteur pure, entre la gamme de nuances très réduite est le son plein, capte un procédé intrinsèque à la voix mais inhabituel à la flûte.



Ex. 36 : Geste dérivé des consonnes fricatives [ʃ] et [s].

Autre son marquant dans *Auto-détermination* est la combinaison [ɥi] (« oui »). Construit d'une semi-consonne suivie par une voyelle, cette syllabe subit un changement interne de fréquence et de formants qui peut être mise en évidence à travers l'écriture et l'exécution. Dans la version d'origine, [ɥi] est donc chanté:

- en déplaçant doucement la langue et changeant la forme de la bouche afin de produire une transition de formants
- à un niveau sonore élevé et avec une couleur peu utilisée ailleurs dans la pièce, à cause de l'écriture largement déclamatoire

à la flûte, nous cherchons donc un son harmonique qui pourrait être joué:

- en transformant les partiels d'une manière quelconque
- puissamment, avec une couleur très particulière⁶⁴

Ici encore, nous avons plusieurs possibilités. Dans le dessein de simuler l'effet vocal original, nous aurions pu par exemple demander au flûtiste de chanter une voyelle en jouant et de la changer doucement, tout en maintenant un niveau sonore élevé. Mais au vu de notre but initial, il nous a paru plus propice de s'inspirer de ce timbre vocal pour concevoir un son plus approprié à l'instrument. Nous avons donc opté pour un son diphonique légèrement instable qui produit en l'occurrence l'intervalle d'une neuvième majeure. La dite instabilité génère des fluctuations à l'intérieur du son: tantôt la note grave, tantôt la note aigüe sonne plus clairement, tantôt un équilibre est établi (Ex. 37 ; Aud. 29). Ceci provoque des variations dans les partiels renforcés lorsque le diphonique sonne. Par ailleurs, la présence des deux notes change la perception du registre, en ouvrant l'espace fréquentiel et en donnant l'impression d'ampleur, tout en gardant une qualité de son pure qui rappelle l'intensité du son vocal d'origine.



Ex. 37 : Geste dérivé du phonème [qi]

III.1.2.2 APPROCHE STRUCTURELLE-GLOBALE

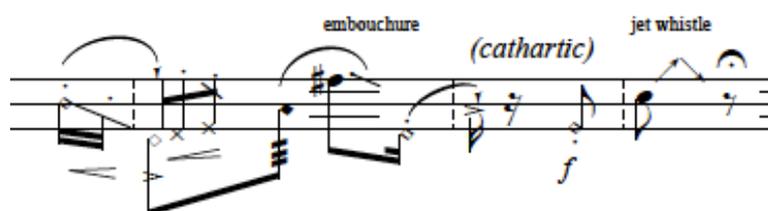
Dans cette section nous discutons des sons dont nous trouvons la fonctionnalité plus importante que la phonétique ou la production vocale. Dans les deux premières phrases du chiffre 11 d'*Auto-détermination*, par exemple, les paroles « sans chaise ! » et « sans chaise,

⁶⁴ Nous remarquons que ces critères sont plus figuratifs et moins précis que ceux utilisés auparavant. Ceci est dû à la singularité **dudit** son, qualité sensible que nous jugeons plus caractéristique que le serait le fait de se référer à ses propriétés plus objectives.

c'est ça ! », constituent le point culminant de la composition, leur force étant largement due au fait qu'elles soient diffusées sur six enceintes entourant le public plutôt que chantées par l'interprète. Il est inutile de préciser que cet effet spatial n'est pas facilement reproductible à la flûte. Pourtant, nous souhaitons créer un événement sonore d'un poids comparable, afin que l'agencement dramatique soit préservé. En analysant les propriétés des dites phrases qui nous permettent de les percevoir en tant que paroxysme, et qui pourrait être déplaçable au domaine instrumental, nous identifions:

- le niveau sonore, qui est bien plus fort qu'à tout autre moment de la pièce
- l'étrangeté d'une dimension sonore extérieure à celle créée par l'écriture vocale

En s'appuyant sur cette analyse, nous avons décidé d'employer le *jet whistle* à la flûte. En plus d'être jouable à une très forte dynamique, ce mode de jeu se déploie sur une gamme de fréquences assez large, ce qui crée un contraste évident avec les sons bruités et les hauteurs isolées dont notre transcription est constituée. Nous estimons que cette deuxième caractéristique du *jet whistle* produit un effet insolite semblable à celui généré par le changement de l'espace sonore dans la version originale (Ex. 38 ; Aud. 30).



Ex. 38 : Emploi du *jet whistle* à la fin de *l'Auto-détermination*

De façon similaire, les trois notes piquées répétées qui apparaissent à la fin des grands phrasés dans *Auto-détermination* (fin de chiffre 2, 5, 9) (Ex. 39), par exemple, ont comme rôle de mener les sections à leur terme. De plus, elles ont un caractère sonore facilement reconnaissable qui est pourtant indépendant du contenu phonétique du texte. Nous les percevons ainsi:

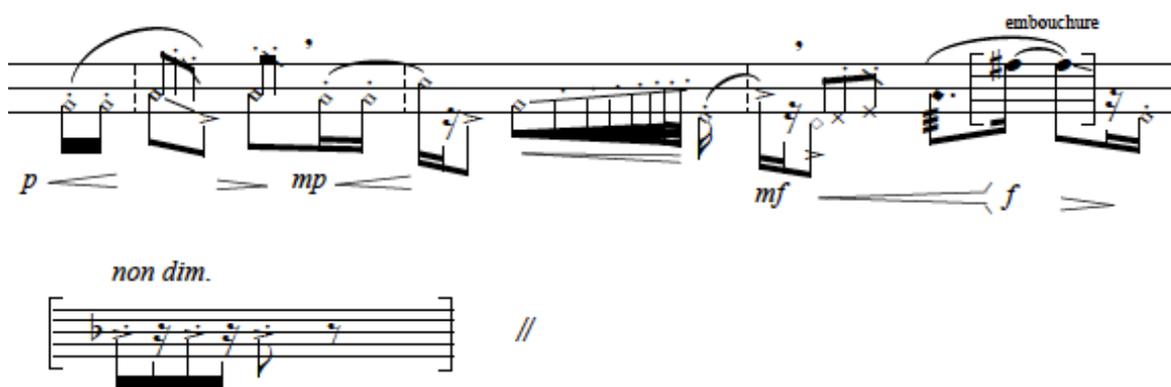
- percussives et bruitées, mais d'une présence moins importante que le [d] mentionné au dessus

- très articulées, avec un certain degré de raideur
- jouable à une variété de nuances



Ex. 39 : Les trois notes répétées

Le pizzicato ainsi que les notes très piquées mélangées avec du souffle pourraient bien répondre aux critères ci dessus; afin de choisir entre les deux, nous considérons l'intégralité de la composition. Ainsi nous constatons trois échelles de percussivité: l'échelle la plus élevée correspond au [d], la prochaine correspond aux terminaison des phrases, et la plus faible correspond aux consonnes [l], [m] et [n] qui apparaissent comme matériel élémentaire tout au long de la pièce. Ce point de vue global nous rend évident que le pizzicato, qui est d'un impact plus fort que le souffle, est plus accommodé aux trois notes répétées à la fin de phrase. Ceci étant établi, nous assignons les notes mélangées avec du souffle aux sons percussifs restant (Ex. 40 ; Aud. 31).



Ex. 40 : Emploi des diverses articulations percussives dans une seule phrase.

Au fil de la pièce, nous appliquons une logique similaire afin d'attribuer les pizzicati et les sons mélangés avec du souffle aux nombreux sons de la version d'origine, en tâchant de

préserver la structure métrique des phrases, d'éviter un trop fort didactisme, et de s'assurer que les enchaînements de modes de jeu ne soient pas excessivement compliqués à réaliser. Ainsi, au chiffre 2 par exemple, nous employons un pizzicato sur le premier temps de la troisième mesure pour dénoter un point d'appui intermédiaire d'une phrase étendue. Dans l'original, le mot « soir », dont nous ne considérons aucun phonème comme percussif, est utilisé dans ce même endroit. Néanmoins, nous jugeons ici la versification plus importante que l'adhérence aux règles établies au dessus, et la note en question trop courte pour exécuter les sons équivalents d'une manière convaincante.

III.1.3 RESUME

En examinant le résultat final de notre travail de transcription, nous jugeons certains éléments issus de nos objectifs initiaux réussis, d'autres infructueux.

Les procédés de recherche des équivalents plutôt intuitifs nous ont frappé dans leur efficacité: ils nous ont permis d'engendrer des gestes instrumentaux surprenants dont la logique est perceptible mais non pas manifeste. Nous trouvons l'agencement de sons fort dissemblable de celle présente dans notre écriture instrumentale habituelle: nous ne nous imaginons pas, par exemple, combiner plus que cinq modes de jeu distincts dans une seule phrase musicale, sans les avoir introduits d'une manière didactique auparavant. Par ailleurs, les gestes issus du passage entre voyelles et consonnes, écoutés hors de leur contexte original, nous semblent proposer une syntaxe insolite pour l'instrument. Ces aspects plutôt réussis peut être qualifiés comme les produits d'un *texte caché*: les choix musicaux sont largement déterminés par un texte qui n'est jamais exposé à l'auditeur et auquel aucune allusion n'est faite dans la pièce instrumentale finale.

En revanche, nous trouvons qu'au fil de la pièce transcrite pour flûte, l'intérêt des gestes diminue à une certaine mesure ce en raison du manque d'un discours global extérieur. Ainsi, nous croyons que l'emploi desdits gestes serait plus réussi au sein d'un contexte musical défini indépendamment du matériel sonore.

III.2 « LEÇONS D'HEBREU » POUR ENSEMBLE ET VOIX DE FEMME

III.2.1 LE PROJET

Ayant exploré des possibilités d'écritures instrumentales dérivées des éléments vocaux au moyen d'une transcription, nous avons souhaité réaliser une composition indépendante se servant d'une palette plus étendue des méthodes d'écriture. Si dans cette première tentative nous nous sommes concentrés sur la recherche de combinaisons sonores précises ici nous avons cherché à établir plusieurs échelles de syntaxe musicale s'inspirant de la parole, partant de la création des sons isolés en allant jusqu'à l'agencement de grandes sections. Si la transcription a été effectuée d'une manière plutôt intuitive, la nouvelle pièce a été composée en s'appuyant sur les recherches scientifiques antérieures ainsi que sur des notions physiques-corporelles de la production du son instrumental et vocal.

Bien que notre travail ait comme but le développement de gestes instrumentaux, nous avons trouvé important d'y intégrer la voix afin que la relation texte - expression vocale - jeu instrumental soit mise en relief. Toutefois, nous avons conçu une pièce dans laquelle voix et ensemble jouent des rôles variables de manière à éviter une hiérarchie voix-instrument qui se trouve assez souvent dans des pièces composées pour cette formation. Ainsi, la voix a été traitée tantôt comme un appareil déclamatoire se distinguant de l'ensemble, tantôt comme une couleur timbrale mise en fond, tantôt comme une présence scénique silencieuse.

III.2.2 LE CHOIX DE TEXTE

En composant *Auto-détermination* nous avons choisi un texte dont les éléments sonores sont plus remarquables et portent un sens plus important que l'éventuel contenu sémantique du texte. Effectivement, le poème de Gerashim Luca est un exemple représentatif du changement de paradigme subi par la littérature dans le courant du 20ème siècle selon lequel la signification abstraite qui transcende le texte est abandonnée en faveur d'une sensualité concrète issue du son des mots, de leur placement et de leur répétition. Lors de la composition de *Auto-détermination*, nous avons trouvé convenable cette mise en évidence du sonore qui nous a permis d'établir un langage vocal se basant sur la phonétique, sans nous occuper de la communication d'une idée métaphysique.

En abordant la composition d'une nouvelle pièce, il nous est apparu avantageux de faire emploi d'un texte dont le caractère sonore est bien défini mais qui possède aussi une dimension abstraite, sans laquelle notre composition ne serait qu'une mise en correspondance des phonèmes et des sons instrumentaux. Par conséquent aux recherches exposées dans les sections antérieures de ce travail, nous avons également souhaité utiliser un texte riche en sons transitoires. Nous avons ainsi décidé de chercher un texte en hébreu, cette langue consiste en une plus grande quantité de sons de consonnes et en une plus petite quantité de sons de voyelles que la plupart de langues occidentales. Par ailleurs, les consonnes les plus répandues dans l'usage de la langue hébraïque sont des fricatives non-voisées dont le spectre est caractérisé par une abondance des sons transitoires et dont la nature sonore est bien distinguable.

Après avoir examiné plusieurs textes en hébreu, nous avons décidé de nous focaliser sur l'œuvre de l'écrivain israélien Hezy Leskly. Le langage poétique de Leskly met en évidence les susdites particularités de la langue hébraïque sans que celles-ci soient en excès. Par ailleurs, la manière presque synesthésique dans laquelle il mélange des métaphores sensuelles tout en économisant le vocabulaire dote chaque poème d'une ambiance précise qui peut facilement évoquer des idées musicales.

Ayant parcouru l'œuvre du poète, nous avons sélectionné six poèmes du cycle *Leçons d'hébreu* : « Leçons A », « Leçons D », « Leçons E », « Leçons F », « Leçons G » et « Leçons H » (Annexe 3). Ce choix est dû à l'équilibre que nous y trouvons entre des éléments acoustiques inspirant l'écriture instrumentale, des éléments structurels se prêtant à la construction d'une forme conséquente, et des éléments saugrenus nous permettant de briser le didactisme inhérent à un travail fondée sur des recherches méthodiques. De plus, le propos du cycle auquel le titre fait allusion est lié à nos buts artistiques, soit l'exploration d'une langue parlée et sa réappropriation à travers des instruments musicaux.

III.2.3 LA FORME

Nous avons déterminé six poèmes du cycle mentionné ci-dessus qui pourraient s'enchaîner afin de former une composition musicale unifiée sans que chaque poème perde de son caractère propre. A cette fin, nous avons cherché à établir le rôle de chaque poème au sein de l'œuvre complète en partant de ses attributs individuels.

Notre conception formelle commence au troisième poème (« Leçon E »), que nous avons désignée comme extension de l'idée entreprise dans *Auto-détermination*, soit une « traduction » d'un texte en gestes instrumentaux. Cette désignation a été motivée par la répétition et le placement structurant des mots au sein du texte : nous avons donc estimé que les qualités rythmiques du poème ainsi que le sens du déroulement seront préservés bien que les paroles soient remplacées par des sons instrumentaux. Par ailleurs, nous trouvons que ce poème en particulier possède une signification plus profonde que le reste de texte choisi; nous souhaitons éviter la mise en musique d'un tel texte car nous le considérons inadéquat pour l'expression de son sens et préférons fournir un complément, une réflexion du texte exprimés à travers le médium instrumental.

Ayant ainsi établi que l'interprétation du troisième poème - qui constitue le mouvement central de la composition - sera purement instrumentale, nous avons considéré les textes précédents et subséquents. En observant la prosodie, le genre des consonnes employées, ainsi que les ambiances évoquées par chaque poème, nous avons développé certaines notions de ressemblance et de différence entre eux. Selon ces impressions, nous avons décidé de mettre en place des liens formels entre le deuxième et le quatrième poème (« Leçon D », « Leçon F »), ainsi qu'entre le premier et le cinquième poème (« Leçon A », « Leçon G »). Ces liens seraient mis en évidence principalement par la manière dont la voix est employée et par son rôle vis-à-vis l'ensemble, mais aussi par nos choix de tempo et d'orchestration.

Le résultat de cet agencement préalable est une forme presque symétrique encadrée par le premier et le cinquième mouvement, le troisième mouvement servant de noyau, le deuxième et le quatrième mouvement représentant des passages, le sixième (« Leçon H ») étant une coda (Ex. 41). A partir de ce schéma nous avons abordé l'écriture vocale et instrumentale.

	I.	II.	III.	IV.	V.	VI. (coda)
Tempo 90						
Tempo 45						
Tempo 60						
Libre						

Ex. 41 : La forme de Leçons d'hébreu.

III.2.4 L'ÉCRITURE

III.2.4.1 LES EQUIVALENTS SONORES

III.2.4.1.2 CONSONNES FRICATIVES

En s'appuyant sur les recherches menées dans le domaine du traitement du signal et exposées dans la deuxième section de notre travail, nous avons tenté de définir des possibilités d'équivalence sonore entre voix et instrument. Cette tentative ressemble à celle entreprise lors de la transcription de *Auto-détermination*, à l'exception du fait que nos efforts actuels se fondent sur les résultats du travail antérieur effectué avec des outils informatiques, ainsi que sur les notions en dérivées. Nous n'avons pas entamé ce travail dans le but de codifier l'écriture, ni afin de fixer des règles qui le gouverneraient. Nous avons plutôt souhaité construire une base de données auquel nous ferions appel tout au long du processus compositionnel, tout en réservant la possibilité de ne pas s'en servir quand une autre approche compositionnelle nous paraît plus propice. Comme dans la transcription de « Auto-détermination », nous instituons ces mesures de liberté afin d'éviter l'éventualité que la composition qui en résulte ne soit trop didactique.

Les sons instrumentaux analogues seront employés au fil de la pièce aux fins suivantes :

1. l'orchestration des phonèmes parlés ou chantés par la voix ; cette fonction peut être effectuée simultanément (effet de couleur ou d'accentuation), peu avant ou peu après (effet d'hétérophonie) l'énoncé des phonèmes par la chanteuse.
2. la préparation ou l'extension dans le temps d'un phonème parlé ou chanté par la voix.
3. le remplacement des phonèmes ou des paroles par des sons instrumentaux.

Nous avons commencé avec la consonne [x], consonne fricative vélaire non-voisée, qui est une des consonnes les plus répandues dans la langue hébraïque et dans « Leçons d'hébreu ». Les équivalents sonores ont été déterminés notamment grâce aux représentations générées par *Crossmel* ; nous avons également cherché des sons instrumentaux produits par des gestes physiques s'approchant des gestes de production vocale, en l'occurrence un frottement non-résonant provenant de la profondeur de l'instrument. Les résultats de cette enquête sont les suivants :

- instruments à cordes : son écrasé⁶⁵
- flûte : *flutterzunge* sans hauteur joué en fermant l'embouchure avec la bouche (voir III.1.2.1)
- clarinette : *flutterzunge*
- piano : éraflure d'une des cordes filées, trouvées dans la région la plus grave du piano, de telle manière que le frottement de l'ongle contre le filetage de la corde soit audible. La pédale est appuyée lorsque la corde est éraflée est relâchée à la fin du geste afin d'étouffer la résonance.

Nous avons ensuite abordé la consonne [ʃ], consonne fricative post-alvéolaire non-voisée dont l'emploi dans « Leçons d'hébreu » est fréquent. Ici aussi, nous nous sommes appuyés sur *Crossmel* ainsi que sur la notion gestuelle d'un frottement non-résonant plutôt superficiel (provenant des extrémités de l'instrument, ou ne suscitant qu'une réverbération des régions limitées de l'instrument). Nous avons obtenus les sons suivants :

- instruments à cordes : trémolo joué *flautando* haut sur la touche en étouffant les cordes avec la main gauche. La main gauche peut glisser le long de la corde afin de

65. Voir section II.3.2, p

produire l'effet du changement de hauteur sans que les hauteurs précises soient générées, en rappelant l'effet similaire produit par les lèvres de la chanteuse.

- flûte : *flatterzunge* joué uniquement avec du souffle. L'embouchure peut être serrée ou élargie en tournant la flûte légèrement afin de produire l'effet du changement de hauteur, en rappelant l'effet similaire produit par les lèvres de la chanteuse.

Ces sons pourraient également servir en tant qu'équivalent pour la consonne fricative alvéolaire non-voisée [s]; grâce à sa similitude à [ʃ] et son emploi réduit dans le texte de « Leçons d'hébreu » nous ne croyons pas que l'invention de sons supplémentaires soit méritée.

Par la suite nous avons abordé la consonne fricative labio-dentale non-voisée [f] (légèrement chuintante). Comme suggéré par la désignation entre parenthèses, le son et surtout le mode de production de [f] ressemblent à ceux de la consonne chuintante [ʃ]. Ainsi, en déterminant les sons instrumentaux analogues, nous n'avons pas considéré uniquement les analyses réalisées avec *Crossmel* ni les notions gestuelles associées avec la production de la consonne [f], mais aussi des propriétés permettant d'établir une différenciation entre les sons choisis pour ces deux consonnes. Les résultats sont comme suit :

- instruments à cordes : trait de l'archet sur le bois du chevalet, générant un son de « souffle »; trémolo joué *flautando*. Ce dernier peut être distingué de celui utilisé pour [ʃ] en variant le point de contact avec l'archet (ailleurs que sur la touche) ou le degré d'étouffement de la main gauche (plus ou moins de hauteur), ainsi qu'en préparant la corde avec une pince à linge.
- flûte : *flatterzunge* joué uniquement avec du souffle. Une différenciation peut être créée en variant la quantité de souffle et de hauteur employée.
- piano : frottement des cordes avec une brosse en fibre.

Ces mêmes sons peuvent être utilisés en traitant la consonne fricative labio-dentale voisée [v], et en favorisant la production des hauteurs afin que le caractère voisé de celle-ci soit réfléchi.

La dernière consonne fricative pour laquelle nous avons déterminé des équivalents est [h], consonne fricative glottale non-voisée. Ici notre travail de recherche a été assez simple : puisque le son engendré par [h] ressemble fortement à exhalaison, il nous semble possible de le reproduire avec les instruments à vent jouant uniquement du souffle. La production de ce

son à la flûte est évidente ; à la clarinette, elle nécessiterait éventuellement que l'embouchure soit enlevée (en fonction de l'instrumentiste).

A ce stade de notre travail, nous avons considéré que la quantité des équivalents sonores établie jusqu'ici suffirait pour la représentation des consonnes fricatives au sein de notre pièce. Il nous est apparu que la définition de sons supplémentaires engendrerait trop de possibilités sonores pour la création d'une syntaxe cohérente et perceptible.

III.2.4.1.2 CONSONNES OCCLUSIVES

Nous avons ensuite abordé la recherche des sons instrumentaux analogues aux consonnes occlusives utilisées en *Leçons d'hébreu*, soit [p], [b], [t], [d], [k] et [g]. En résumant les expériences recueillies lors du travail de segmentation en phonèmes (II.3.1) et de transcription (III.1.2.1), nous remarquons certains inconvénients posés par ces consonnes :

1. puisque leur contenu bruité est porté par des sons harmoniques dont les paramètres sont peu stables, la segmentation les isole rarement de manière conséquente. Ainsi, les résultats des analyses effectuées par la suite sont moins fiables que ceux obtenus pour les consonnes fricatives.
2. leurs densités spectrales et enveloppes dynamiques ne correspondent pas toujours au profil dynamique de la parole, qui est souvent gouverné par d'autres phénomènes linguistiques. Ainsi, dans maints cas nous ne pourrions pas respecter la répartition des accents dans un mot en étant fidèle aux qualités acoustiques des consonnes occlusives.

En conséquence de ces observations, nous avons décidé de construire une base de données généralisée des sons instrumentaux percussifs se rapprochant de ceux des consonnes occlusives, sans tâcher de désigner un son précis pour chaque consonne occlusive employée. Toutefois, nous avons pris note des sons qui correspondent particulièrement bien à une consonne ou une autre. Ainsi nous avons à notre disposition une liste des possibilités à partir desquelles nous pourrions choisir le son ou le geste le plus approprié à chaque cas musical : lorsque le contexte le permet, nous sélectionnons des sons en fonction des critères spectraux-dynamiques ; autrement, nous sélectionnons des sons selon la manière dont ils mettent en relief les points d'appui importants.

Nous avons donc obtenu les sons percussifs suivants :

- instruments à cordes : pizzicato et pizzicato Bartok en étouffant la corde avec la main gauche ; col legno battuto [t] ou jeté [k, t] en étouffant la corde avec la main gauche ; son écrasé très court [k]
- flûte : tongue ram ; pizz [t]
- clarinette : *slap tongue* [b, d, g]
- piano : son des cordes bloquées avec pâte-à-fixe [k, t] ; sons harmoniques produits en arrêtant la corde à un nœud avec la main ou avec de la gomme ; appui et relâchement rapide de la pédale [b, d, g]

Ainsi, munis de nombreuses possibilités développées dans cette partie de notre travail, nous abordons la composition de « Leçons d'hébreu ».

III.2.4.2 « LEÇON A », « LEÇON G »

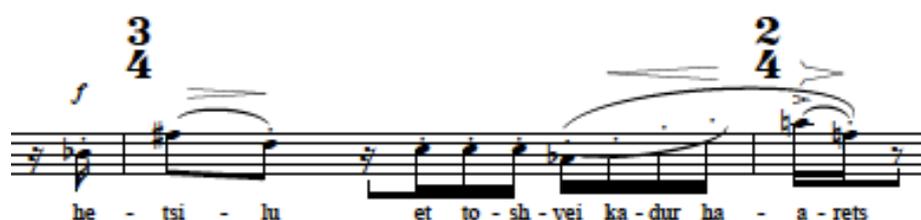
Comme nous l'avons mentionné antérieurement, nous traitons les premier et cinquième poèmes d'une manière similaire, en appliquant les mêmes principes d'écriture aux deux. Dans un premier stade, nous avons défini quelques attributs principaux de l'organisation musicale des ces deux mouvements :

1. l'écriture vocale est d'un style déclamatoire, ce qui permet de faire ressortir les consonnes et d'éviter un lyrisme excessif.
2. l'ensemble génère l'ambiance sonore et orchestre la partie vocale, en mettant en relief certains des équivalents sonores établis au dessus ; de plus, le jeu d'ensemble présente d'autres perspectives, ou même des commentaires sur le texte.

Ces choix sont dus au fait que ces deux textes :

1. sont plus longs et plus narratifs que les autres poèmes du cycle; ainsi, un abord syllabique déclamatoire correspond au caractère des textes et permet de transmettre l'intégralité sans interrompre la continuité rythmique inhérente à la prosodie.

2. contiennent une trop grande variété de consonnes pour que celles-ci puissent fournir une structure cohérente : nous craignons qu'un discours musical instrumental dérivé de tels textes soit trop hétérogène, voir incompréhensible pour l'auditeur. Ainsi, nous trouvons judicieux d'attribuer à l'ensemble un rôle lié au texte d'une manière moins concrète, et de n'utiliser des sons correspondants aux phonèmes précis que dans des moments isolés.



Ex. 42 : L'écriture vocale déclamatoire. « Leçon G »

En respectant la prosodie des poèmes et de l'hébreu, une écriture vocale déclamatoire implique des successions plutôt rapides des sons plutôt courts (Ex. 41). Nous avons décidé d'affermir cet effet en créant une texture instrumentale pointilliste composée des sons liés au contenu phonétique du texte (ex. 42 ; Aud. 33). Au sein de ces deux couches musicales, nous souhaitons introduire quelques combinaisons sonores instrumentales faisant appel aux équivalents sonores développés auparavant et ayant un rapport plus évident avec les phonèmes correspondants ; l'établissement de tels gestes serait plus réussi en traitant des phonèmes tenus et donc fricatifs, qui contrasteraient avec les sons éphémères formant la partie vocale et la texture instrumentale.

Ayant posé ces bases compositionnelles, nous avons examiné les poèmes plus en détail. Pour plus de concision, nous ne discuterons ici que le premier. Dans le texte reproduit ci-dessous, les sons que nous souhaitons exagérer dans la partie vocale et représenter dans la texture instrumentale sont mis en gras ; les phonèmes pour lesquels nous souhaitons construire des combinaisons sonores instrumentales plus précises sont soulignés.⁶⁶

66. Ce tri a été effectué d'une manière à ne pas contrarier la fluidité générale du poème; ainsi, certains phonèmes sont marqués aux endroits donnés dans le texte et non pas dans d'autres.

Ex. 43 : Texture instrumentale pointilliste. « Leçon A »

SHIUR ALEPH

Ha'rega ani lomed likhtov :⁶⁷

esh ve'khashmal, le'lo shgiot.

Ve'ha'rega okhezet esh be'shulkhan ha'ktiva

Ve'ha'khashmal ba'kheder kave.

Le'oro shel ha'shulkhan ha'bo'er

ani lomed leta'er

et ha'esh

ve'laga'at ba'khashmal.

LEÇON A

En ce moment j'apprends à écrire :

feu et électricité, sans fautes.

Et en ce moment le feu prend la table d'écriture

et l'électricité dans la salle s'éteint.

A la lumière de la table brûlante

j'apprends à décrire

le feu

et à toucher l'électricité.

67. Dans cette translittération, la consonne [j] est notée « sh » et la consonne [x] est notée « kh ».

À partir de cette classification des phonèmes du texte, nous avons composé des sons ou des assemblages des sons instrumentaux exprimant les phonèmes soulignés. Ce procédé a souvent nécessité une adaptation de l'écriture vocale afin que les phonèmes choisis soient énoncés, ou parfois répétés, sur une durée suffisante pour que le développement des sons instrumentaux soit audible. Ainsi, nous avons distordu la langue d'une certaine mesure en rallongeant des sons normalement éphémères ; ce choix compositionnel engendre une syntaxe vocale singulière qui met en relief les sons particuliers à l'hébreu d'une manière satirique correspondant à l'esprit manifeste dans les textes de Leskly. Nous croyons important néanmoins de signaler que lesdites distorsions ont été effectuées à l'intérieur d'un contexte fidèle à la prosodie et au caractère rythmique du texte, et donc ne compromettent pas l'établissement d'une relation texte-voix-instrument cohérente.

Par la suite nous illustrons ce travail avec quelques exemples.

The image shows a musical score for a vocal piece in 4/4 time. The score is divided into two systems. The first system shows the vocal line and piano accompaniment. The vocal line starts with a phrase marked *p* (piano) and ends with a note marked *sf* (sforzando). The piano accompaniment includes a *staccato* instruction. The second system shows the vocal line with lyrics: "h - a re - ga a -". The first syllable "h" is marked *mf* (mezzo-forte). The piano accompaniment includes *mp* (mezzo-piano) markings and performance instructions like *sul pont* and *pizz* (pizzicato). Red and blue circles highlight specific musical elements: a red circle around the first vocal phrase and a blue circle around a specific note in the bass line.

Ex. 44

Dans cet exemple, la première syllabe du texte, « Ha », est préparée par des sons du souffle à la flûte et à la clarinette basse, qui s'enchaînent avec l'énoncé de la consonne [h] à l'entrée de la voix. Dans la mesure suivante, la consonne occlusive [g] à la voix est imitée avec un léger décalage par un slap tongue à la clarinette et par un pizzicato Bartok au violoncelle. Au fil de la pièce, nous employons souvent de tels assemblages des sons percussifs non-synchronisés afin de prolonger l'effet des consonnes occlusives, de façon qu'elles puissent être perçues par l'auditeur en tant qu'unités sonores indépendantes.

The image displays a musical score for Example 45, consisting of three systems of staves. The first system features a flute part in 3/4 time, marked with a *p* dynamic. A box highlights a section in 5/4 time, with instructions: "vary pitch by turning flute", "turn outwards" (with an arrow pointing right), and "turn inwards" (with an arrow pointing right). The flute part shows a pitch contour that rises and then falls. The second system shows empty staves for the violin and piano. The third system shows a violin part in 3/4 time, marked with a *p* dynamic, with the syllable "e - sh" below it. A box highlights a section in 5/4 time, with instructions: "widen mouth" (with an arrow pointing right) and "narrow mouth" (with an arrow pointing right). The violin part shows a pitch contour that rises and then falls. The dynamic marking *f* is present in this section. The syllable "ve - ch" is written below the violin part. The piano part in the third system is marked with a *mp* dynamic and includes a note with a damper pedal symbol. A box highlights a section in 5/4 time, with the instruction "damp string, no distinct pitch". The piano part shows a pitch contour that rises and then falls, with dynamic markings *p* and *f*.

Ex. 45

Ici, la consonne fricative [ʃ] est prolongée dans la partie vocale en modifiant la hauteur au moyen d'un filtrage effectué par la bouche et les lèvres. Ce geste est reproduit à la flûte et au violon en utilisant les équivalents mentionnés ci-dessus. Comme dans l'exemple précédent, le léger décalage entre la voix et ces deux instruments étend le geste vocal. Par ailleurs, nous pouvons entendre dans la juxtaposition de ces trois sons ayant des profils dynamiques semblables mais temporellement différents la formation d'un objet sonore que nous pourrions réutiliser par la suite.

Ex. 46 :

The image displays a musical score for Example 46, consisting of three systems of staves. The first system features a vocal line and two instrumental staves. The vocal line is in 2/4 time, followed by a 4/4 section. The instrumental parts are in 4/4 time. The second system continues the vocal line with lyrics 'kti k-ti-k-ti-k-ti v-a' and includes dynamic markings like *mf* and *f*. The third system shows further instrumental and vocal parts with dynamic markings like *f* and *mf*. Red circles highlight specific passages in the vocal and instrumental lines, and a red oval highlights a section of the vocal line with the lyrics 'kti k-ti-k-ti-k-ti v-a'. The score includes various musical notations such as notes, rests, and dynamic markings.

Ici les consonnes occlusives [k] et [t] sont reproduites par le violon et le violoncelle jouants *col legno battuto*. Afin de différencier ces sons dérivés directement du texte et les sons similaires constituant la texture jouée par l'ensemble, ils nous a fallu modifier ces deux consonnes dans la partie vocale. Ainsi, nous avons décidé de répéter les phonèmes [k] et [t] plusieurs fois en accélérant lors de l'énoncé du mot « ktiva », ce qui génère un geste vocal tenu d'une qualité itérative et d'un profil dynamique précis. Ces phonèmes étant ainsi façonnés, nous avons pu les reproduire avec des jetés d'archet *col legno* semblables au profil itératif-accélérant du geste vocal. Puis, ayant trouvé que cet assemblage sonore manque d'ampleur, nous avons ajouté un

geste similaire à la flûte : l'itération rapide des sons mélangés avec du souffle et articulés en prononçant [k] et [t] avec la langue.

III.2.4.2 « LEÇON E »

Dans le mouvement central de *Leçons d'Hébreu*, nous avons développé l'idée de l'écriture instrumentale issue d'un texte caché que nous avons exposé dans *L'Auto-détermination* en l'appliquant à trois instruments de caractères sonores bien différenciés : le violon, la flûte et le piano. Par ailleurs, nous avons souhaité ici composer des gestes instrumentaux dérivés directement des paroles du poème, sans passer par le stade intermédiaire d'une composition vocale. Nous avons cru que cette décision favoriserait une plus grande souplesse dans l'invention des gestes.

Comme le texte de *L'Auto-détermination*, le troisième poème de *Leçons d'hébreu* consiste en un contenu phonétique limité dont certains phonèmes possèdent une fonction structurante, et ce grâce à leur répétition et leur positionnement. Mais en contraste avec « L'Auto-détermination » de Gerashim Luca, l'agencement phonétique manifesté dans « Léçon E » est subtil et secondaire. Ainsi nous estimons que la cohérence intrinsèque du texte sera révélée dans les gestes instrumentaux sans qu'elle dicte le déroulement musical aussi rigoureusement que dans *L'Auto-détermination*.

SHIUR HEH

Keshe'ha'mila tahafokh le'guf
ve'ha'guf yiftakh et piv
ve'yomar et ha'mila shmimena
notsar –
ekhabek et ha'guf ha'ze
ve'alín oto le'tsidi.

LEÇON E

Quand le mot deviendra corps
et le corps ouvrira sa bouche
et énoncera le mot duquel il fut
engendré –
j'êtreindrai ce corps
et le coucherai à mon coté.

En nous appuyant sur les conclusions tirées de la transcription de *l'Auto-détermination*, nous avons déterminé quelques critères qui nous guideront lors de la composition de ce mouvement :

1. bien qu'ils soient intéressants, les gestes issus de la traduction du son vocal en son instrumental ne suffisent pas en eux-mêmes pour la construction d'un discours musical autonome. Ainsi, il nous faut envisager une forme globale indépendamment du processus de traduction afin de cadrer lesdits gestes.
2. étant d'une haute complexité, les gestes doivent être répétés plusieurs fois afin qu'ils soient compris par l'auditeur.
3. pour cette même raison les gestes de chaque instrument doivent être exposés individuellement pour que l'auditeur puisse comprendre les notions syntactiques propres à chaque instrument.

En nous inspirant de la façon générative dont les paroles « guf » (corps) et « mila » (mot) sont exposés ainsi que de l'organisation inflexionnelle similaire des premiers vers, nous avons décidé d'employer une forme musicale imitative vaguement semblable à une exposition de fugue. Ce choix satisfait aussi aux premier et troisième critères établis ci-dessus : la forme imitative possède une cohérence assez évidente qui assurera la suite logique des événements musicaux ; l'introduction graduelle de chaque « voix » jouée par un instrument différent favorisera une écoute attentive aux premiers gestes exécutés par chaque instrument. Afin de satisfaire au deuxième critère nous avons décidé de répéter chacun des deux premiers vers trois fois. Ceci nous permet également de rallonger le poème, qui est trop court dans son état actuel pour servir à nos buts. Ainsi, le début du mouvement se déroulera comme suit (Ex.) :

1. Le violon joue les gestes correspondants au premier vers, trois fois.
2. Au terme de la troisième répétition, la flûte entre et joue ses propres gestes correspondants au premier vers, trois fois. Le violon passe au deuxième vers, qu'il joue aussi trois fois.
3. Au terme de la troisième répétition, le violon avance au troisième vers, la flûte au deuxième vers, et le piano entre en jouant ses propres gestes correspondant au premier vers. Ce vers n'est exécuté qu'une seule fois.

Ayant déterminé la forme, nous avons abordé la composition des gestes. A cette fin nous nous sommes appuyés fortement sur les équivalents sonores développés antérieurement, tout en essayant de respecter le phrasé de chaque vers. En entamant ce travail nous avons remarqué que notre base de données d'équivalents sonores était trop générale pour écrire un mouvement entier composé uniquement de tels gestes sonores. Par conséquent, nous avons trouvé nécessaire par moments de varier les équivalents déjà établis ou de les compléter avec d'autres sons. Les consonnes [f] et [v] par exemple, sont représentées par des sons semblables dans notre base de données à cause de leur nature acoustique presque identique ; cependant, chacune de ces consonnes a une fonction important au sein du texte. Ainsi, afin de préserver le rôle de chaque consonne, il nous a fallu chercher d'autres possibilités. Ce travail est illustré par les exemples suivants⁶⁸ :



Ex. 47

Dans cet exemple nous avons créé un geste pour violon dérivé de la parole « tahafokh ». D'abord, nous extrayons les consonnes: [t], [h], [f], [x]. Selon notre base de données, nous employons le son *col legno battuto* pour le [t] et le son écrasé pour le [x].

En traitant la consonne [f], il nous a paru évident que l'emploi de l'archet sur le bois du chevalet proposé par la base des données n'était pas convenable ici à cause du temps nécessaire pour soulever l'archet et le replacer sur le chevalet. Ceci engendre une déformation temporelle suffisamment importante pour que le geste ne soit plus entendu comme une seule unité. Or, l'autre son proposé par la base de données, le tremolo, est difficile à distinguer du son de [v] comme expliqué auparavant. Ainsi nous avons modifié le [f] en maintenant l'enveloppe dynamique, que nous trouvons caractéristique, et en changeant le timbre au moyen d'une pince à linge placée sur la corde du Sol du violon. Le son qui en résulte se distingue

⁶⁸ A cause de la pauvre qualité de l'interprétation de ce mouvement dans l'enregistrement, nous sommes dans l'impossibilité de fournir des exemples sonores correspondant aux extraits précis discutés ici. Par contre, l'enregistrement de la première partie du mouvement est inclus à la fin de cette section et l'enregistrement intégral est inclus en annexe.

du son que nous emploierons pour la consonne [v] et s'intègre facilement dans le geste cadré par le *col legno battuto* et le son écrasé.

Ensuite il s'est agi de trouver un son correspondant à la consonne [h] car nous n'avons pas déterminé un analogue pour cette consonne au violon lors de la construction de la base des données. Afin de récréer le son du souffle au violon⁶⁹, nous pouvons jouer avec l'archet sur le bois du chevalet ou sur le cordier, ou bien faire des mouvements verticaux avec l'archet le long de la touche. Cependant, ces trois possibilités présentent les mêmes problèmes temporels que nous avons rencontré en traitant le [f]. Ainsi, il nous a fallu définir encore un équivalent. Nous avons choisi le son harmonique car il partage certaines caractéristiques acoustiques avec le son des instruments à vent et car son exécution est faisable entre les sons correspondants à [t] et à [f].



Ex. 48

Ici nous avons composé un geste dérivé de la parole « guf » au violon. Comme cette parole a une fonction structurante, nous avons souhaité que le geste correspondant soit construit par des sons uniques qui ne s'emploient pas ailleurs. La juxtaposition d'un pizzicato à la main gauche pour la consonne [g] et du son de l'archet tiré sur le bois du chevalet pour la consonne [f] nous a paru être une combinaison réussite. Le pizzicato est un des sons que nous avons défini comme équivalent possible pour les consonnes occlusives voisées (y compris [g]). Exécuté sur la corde du Sol déjà préparé avec la pince à linge, il obtient en plus une couleur singulière. Par ailleurs, le pizzicato peut être joué avec la main gauche, ce qui laisse suffisamment de temps pour le déplacement de l'archet au chevalet avec la main droite. La différence entre les deux sons utilisés pour la consonne [f] aide à la distinction de la parole « guf ».

69. Du fait de la mauvaise qualité de l'interprétation de ce mouvement lors de l'enregistrement, nous sommes dans l'impossibilité de fournir des exemples sonores correspondant aux extraits précis discutés ici. Par contre l'enregistrement de la première partie du mouvement est inclus à la fin de cette section et l'enregistrement intégral est inclus en annexe.

Après avoir ainsi façonné trois phrases musicales correspondants aux trois premiers vers du poème, nous avons effectué de légères variations dans la structure d'accentuation des vers répétés, ce afin de mettre en relief les composants éphémères des gestes en les prolongeant. Le premier vers, par exemple, a été modifié comme suit (Ex. 49) :

Ex. 49

Vers 1a (originel) : kshe'ha'miLA tahafoKH le'Guf

Vers 1b : kshe'ha'MIla tahaFokh le'guF

Vers 1c : kshe'HA'mila TAhafokh le'guF

Ensuite nous avons recomposé les deux premières phrases pour la flûte en s'appuyant sur la partie du violon ainsi que sur le texte originel. Le rythme de cette partie a été ré-proportionné selon une division ternaire de la pulsation (Ex. 50), en contraste avec la division binaire de la partie du violon, dans le but de donner à chaque « voix » un caractère rythmique et de créer une texture non-synchron en les superposant.

Ex. 10 : La phrase du violon ré-écrite selon une division ternaire de la pulsation.

Finalemment, la première phrase du violon a été réécrite pour piano préparé en employant la même organisation rythmique : la texture formée par les trois instruments superposés ainsi que par le décalage entre les points de départ et les terminus de leurs phrases est d'une telle complexité que des variations supplémentaires ne sera pas perceptibles.

(Aud. 37)

III.2.4.2 AUTRES MOUVEMENTS

Nous souhaitons ici ajouter trois extraits provenant des autres mouvements de *Leçons d'hébreu* qui illustrent l'emploi des techniques exposées dans cette dernière section de notre travail (Aud. 38).

The musical score consists of four staves: Flute, Clarinet in B \flat , Piano, and Soprano. The piece is divided into three measures with time signatures $\frac{3}{4}$, $\frac{5}{4}$, and $\frac{4}{4}$. The tempo is marked as quarter note = 45. The Flute and Clarinet parts feature glissando markings and piano (pp) dynamics. The Soprano part includes performance instructions: "slide tongue towards soft pallet", "slide tongue towards teeth", and "as long as possible (leaving enough breath to finish the phrase)". The Soprano part also has a "pp" dynamic marking and a "1 - 1" marking below the first measure.

Ex. 51

3/4 4/4 3/4 3/4 3/4

pp p p mp sfz

air + pitch

pizz arco pizz

s sfz sfz

a t u f a be a

tr

sfz p

pizz I sfz

arco p

pizz I sfz

Ex. 51

CONCLUSIONS

L'évolution de la composition instrumentale, depuis la Renaissance jusqu'à aujourd'hui a vu le développement d'une dialectique autonome, dont la présence dans la pensée des compositeurs contemporains est évidente. Dans notre travail nous avons abordé la possibilité qu'une autre logique peut servir à la composition instrumentale ; ainsi, en employant l'expression vocale comme modèle, nous avons exploré une approche compositionnelle permettant d'établir une nouvelle syntaxe instrumentale. Nous résumerons ici les matières abordées dans nos recherches et les déductions effectuées.

Dans une première étape nous avons étudié des œuvres instrumentales et électroniques de P. Ablinger et J. Harvey dans lesquelles le discours et l'écriture se sont étroitement inspirées de modèles vocaux. Ce travail fut révélateur : notre conception du rôle du modèle et de ses exploitations possibles au sein d'une pièce musicale a été considérablement enrichie par la pensée de ces deux compositeurs. Par ailleurs, cette étude nous a aidé à identifier les approches de la modélisation qui ne sont pas appropriées à nos buts artistiques. Nous avons ainsi pu éviter des pistes peu fructueuses et démarrer nos recherches informatiques en ayant une idée claire des objectifs et des méthodes qui nous guideraient.

Suivant les conclusions tirées de ce stade préalable , nous avons abordé le traitement numérique des sons transitoires, afin que ceux-ci puissent servir de modèle pour l'écriture instrumentale. Ayant évalué les techniques de segmentation de la parole et d'analyse des transitoires, nous avons déterminé que les outils s'appuyant sur la perception humaine seraient les mieux adaptés. Ces outils nous ont permis d'observer le taux de similitude entre certains sons instrumentaux et certains phonèmes, ce qui constitue une première étape de l'établissement d'une syntaxe instrumentale issue des caractéristiques vocales.

Afin d'approfondir et adapter ce travail au domaine du traitement numérique, nous avons entrepris une dernière étude s'appuyant principalement sur les intuitions acquises lors des recherches menées jusqu'ici. Ainsi riches de nos expériences scientifiques d'une part et sensibles d'une autre, nous avons abordé la composition de *Leçons d'hébreu*. Le travail de composition de cette œuvre fut une synthèse des notions élaborées lors de nos recherches antérieures. Ici nous avons tâché de trouver un équilibre entre la formalisation et la perception, en changeant fréquemment notre angle d'approche, et ce afin de compenser les faiblesses de chacun. La composition résultant de ce processus nous paraît avoir satisfaite nos objectifs artistiques.

Au travers nos efforts de recherche et de création, nous avons démontré la possibilité d'une évolution de l'écriture instrumentale au moyen de structures manifestes dans la langue parlée. Nous croyons que cet abord à l'écriture instrumentale peut engendrer encore maintes possibilités insolites auxquelles le compositeur ne peut pas facilement accéder s'il se limite à la dialectique instrumentale. Nous pourrions imaginer par exemple des gestes instrumentaux dérivés de toute une gamme d'expressions vocales ; ces modèles pouvant être mis en œuvre de différentes manières, des plus abstraites au plus concrètes. Ainsi nous croyons que la poursuite de nos recherches dans des cadres variés peut faire naître des approches compositionnelles singulières, aux exploitations encore plus vastes.

BIBLIOGRAPHIE SÉLECTIONNÉE

ABLINGER, Peter. *Phonorealism, the Reproduction of « Phonographs » by Instruments*. <<http://ablinger.mur.at/phonorealism.html>>, [Consulté le 20 août, 2011].

ABLINGER, Peter. *Quadraturen*. <<http://ablinger.mur.at/docu11.html>> [Consulté le 20 août, 2011].

ABLINGER, Peter. Lecture donnée le 15 février 2011. *Impuls Academy 2011*, Graz, Autriche.

BROWMAN, Catherine P., **GOLDSTEIN**, Louis M. « Articulatory Gestures as Phonological Units » in *Phonology*, Vol. 6, No. 2, 1989, p. 201-251.

CHION, Michel. *Guide des objets sonores*. Paris, France : Editions Buchet/Chastel, 1983. p. 18-139.

CHUN-LIN, Liu. *A Tutorial of the Wavelet Transform*. DISP Lab, Graduate Institute of Communication Engineering, New Taipei University, Taiwan. En ligne : <<http://disp.ee.ntu.edu.tw/tutorial/WaveletTutorial.pdf>> [Consulté le 10 juillet, 2011].

DAUBECHIES, Ingrid. *Ten Lectures on Wavelets*. Philadelphia, Pennsylvania : SIAM Publications. 1992. p. 1-106.

DOLSON, Mark, **LAROCHE**, Jean. « New Phase-Vocoder Techniques for Pitch-Shifting, Harmonizing and Other Exotic Effects ». *IEEE Workshop on Applications of Signal Processing to Audio and Acoustics*, 1999, New Paltz, Etats-Unis (4 p.). Édition électronique : <<http://www.ee.columbia.edu/~dpwe/papers/LaroD99-pvoc.pdf>>, [Consulté le 3 janvier, 2011].

DUBNOV, Shlomo, **RODET**, Xavier. « Study of Spectro-Temporal Parameters in Musical Performance for Expressive Instrument Synthesis ». *Sound and Music Computing Conference* 2008, Berlin, Allemagne. (5 p.). Édition électronique : <www.ircam.fr/equipes/analyse-synthese/.../SMC98/SMC98dubnov.ps.gz>, [Consulté le 9 décembre, 2010].

HAMNANE, Sonia. *Traitement et représentation temps-fréquences des sons avec résolution adaptative*. Rapport de stage de Master ATIAM, Université de Pierre et Marie Curie, Paris, 2006. 60 p. Édition électronique : <<http://www.atiam.ircam.fr/Archives/Stages0506/SoniaHamnane.pdf>>, [Consulté le 19 juillet, 2012].

HARVEY, Jonathan. *Speakings, for orchestra and electronics*. London : Faber Music, 2008. (29 min.).

HARVEY, Jonathan. Entretien informel, réalisé le 7 avril 2008.

HASAN, Rashidul, **JAMIL**, Mustafa, **RABBANI**, Golam, **RAHMAN**, Saifu. « Speaker Identification Using Mel Frequency Cepstral Coefficients ». *3rd International Conference on Electrical and Computer Engineering* 2004, Dhaka, Bangladesh. (4 p.). Édition électronique : <https://www.assembla.com/spaces/strojno_ucenje.../p141omfccu.pdf>, [Consulté le 10 juillet, 2011].

LESKLY, Hezy. *Collected Poems*. Tel Aviv, Israel : Am Oved, 2009. p. 147-155

LETELIER, Juan Carlos, **WEBER**, Pamela P. « Spike Sorting Based on Discrete Wavelet Transform Coefficients » in *Journal of Neuroscience Methods*, Vol 101, No. 2, 2000, p. 93-106.

LEVEAU, Pierre. *Paramétrisation adaptée de transitoires pour la reconnaissance d'instruments de la musique*. Mémoire de stage de DEA ATIAM, Université de Pierre et Marie Curie, Paris, 2004. 58 p. Édition électronique : <<http://www.atiam.ircam.fr/Archives/Stages0304/leveau.pdf>>, [Consulté le 19 juillet, 2012].

MEYER, Yves, **JAFFARD**, Stéphane., **RIOUL**, Olivier. « L'analyse par ondelettes » in *Pour la Science*, Vol 119, 1987, p. 28-37.

NOUNO, Gilbert, **CONT**, Arshia, **CARPENTIER**, Grégoire. « Making an Orchestra Speak ». *Sound and Music Computing Conference 2009*, Porto, Portugal. (6 p.). Édition électronique : < articles.ircam.fr/textes/Nouno09a/index.pdf>, [Consulté le 10 décembre, 2010].

O'SHAUGHNESSY, Douglas. *Speech Communication – Human and Machine*. New York, New York : IEEE Press, 1950.

RASETSHWANE, Daniel. *Enhancement of Speech Intelligibility Using Speech Transients Extracted by a Wavelet Packet-Based Real-Time Algorithm*. Thèse. Swanson School of Engineering. Soutenu à la Pittsburgh University, Pittsburgh, États Unis, 2009. 168 p. Édition électronique : < d-scholarship.pitt.edu/8357/ >, [Consulté le 3 septembre, 2011].

RIOUL, Olivier. *Analyse temps-échelle et temps-fréquence*. Dept COMELEC GET/ENST.

Enligne :

< <http://perso.telecom-paristech.fr/~rioul/documents/200512AnalyseTF.pdf>> [Consulté le 10 juillet, 2012].

RÖBEL, Axel. « A New Approach to Transient Processing in the Phase Vocoder ». *The 6th International Conference on Digital Audio Effects*, 2003 London, Royaumes-Unis. (6 p.). Édition électronique :

<www.mp3-tech.org/programmer/docs/dafx32.pdf>, [Consulté le 3 janvier, 2011].

RODET, Xavier. « Synthesis and Processing of hte Singing Voice ». *The 1st IEEE Benelux Workshop on Model Based Processing and Coding of Audio*, 2002, Leuven, Belgique. (10 p.).

Édition électronique :

<www.esat.kuleuven.be/psi/spraak/seminars/mpca/papers/rodet.mpca02.ps>, [Consulté le 9 décembre, 2010].

TAYLOR, P. A., ISARD, S. D. « Automatic Diphone Segmentation Using Hidden Markov Models ». *Speech Science and Technology Conference*, 1990 Melbourne, Australie. (6 p.).

Édition électronique :

<www.assta.org/sst/SST-90/cache/SST-90-Chapter9-p8.pdf>, [Consulté le 23 août, 2011].

TAN, Bendeict, SEN, Deep. « The Use of Attack Transient Envelope in Instrument Recognition ». *Speech Science and Technology Conference*, 2006 Auckland, Nouvelle Zealand. (5 p.). Édition électronique :

<www.assta.org/sst/2006/sst2006-121.pdf>, [Consulté le 23 août, 2011].

ANNEXE 1

A letter from Schoenberg – texte

Mister:

You.... In spite of my protest,
you have published
Leibowitz' performance
of my Ode to Napoleon
with a woman's voice,
which I find
terrible.
(...behind the orchestra...)
I can only tell you now,
that you will
hear from me.
You will, I can tell you,
you will regret this act
severely.
I will
be busy to help you
to be ruined
by this
what I will do....
(Some of the instruments ... in small....)
You are not only a bugger ...
You are not only a man who disregards an artist's wishes,
his artistic beliefs,
you are also a man
who does not care
to keep a contract.
You know that you signed a contract,
according to which
you have
to account to me regularly.
You must have sold
quite
a number of records
of my Violin Phantasy,
of the Trio,
and other things which you...
but which you issued without my consent.
I tell you,
you will hear from me also about these things,
and I hope it will cost you very much money.

Yours...

AUTO-DÉTERMINATION

For Bass Flute

Hadas Pe'ery

♩ = 90

(serious and somewhat baffled)

1

p

p

mf

embouchure

embouchure rit

mp *mf* *p*

a tempo

accel. poco a poco fino a 4

2

p *mp* *mf* *f*

embouchure

non dim.

3 (slowly growing frenetic...)

Musical score for exercise 3, consisting of three staves of music. The first staff begins with a dynamic marking of *mf* and includes a triplet of eighth notes. The second staff features dynamic markings of *f* and *mp*, with the word "embouchure" written above the notes. The third staff starts with a dynamic marking of *f*, followed by *sf*, and includes the instruction "(indignant!)". It also contains two diagrams of fingerings for the embouchure, each with a vertical box containing the numbers 3, 4, 2, 3, 4. The score concludes with a double bar line (//).

4 Precipitated ♩ = 120
ad lib
(frenetic)

Musical score for exercise 4, consisting of three staves of music. The first staff begins with a dynamic marking of *f** and includes a triplet of eighth notes. The second staff features dynamic markings of *f* and *ff*, with the word "embouchure" written above the notes. The third staff starts with a dynamic marking of *f*, followed by *ff*, and includes two diagrams of fingerings for the embouchure, each with a vertical box containing the numbers 3, 4, 2, 3, 4. The score concludes with a double bar line (//).

5 ♩ = 45 - 50
(having calmed down)

p embouchure

poco rit.

//

6 a tempo

p embouchure

mp

//

7

p embouchure

mp *mf*

accel. poco a poco fino a 10

//

8

Musical score for measure 8. The notation is on a single staff. It begins with a *mf* dynamic marking. The first part of the measure consists of a triplet of eighth notes, followed by another triplet of eighth notes, and then a triplet of sixteenth notes. The dynamic then changes to *sf*. The second part of the measure features a triplet of eighth notes, followed by a triplet of sixteenth notes, and then a triplet of eighth notes. The dynamic remains *sf*. The measure concludes with a triplet of eighth notes and a final note. There are two vertical boxes containing the numbers 3 and 4, indicating fingerings for the triplets. The measure ends with a double bar line.

9

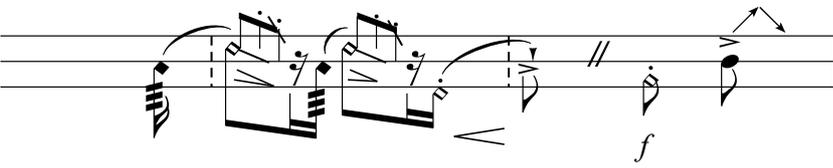
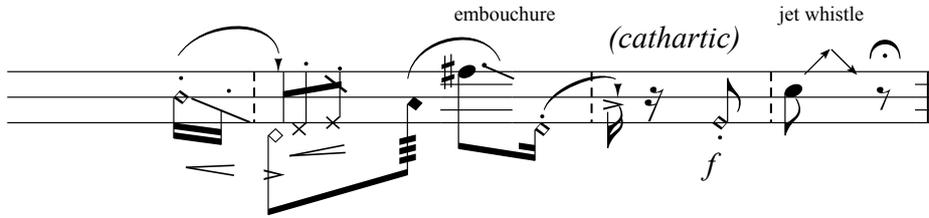
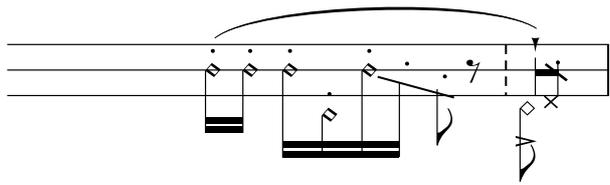
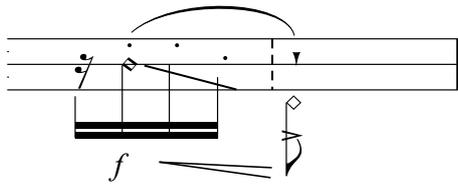
(slowly beginning to grow indignant again)

Musical score for measure 9. The notation is on a single staff. It begins with a *mf* dynamic marking. The first part of the measure consists of a triplet of eighth notes, followed by a triplet of eighth notes, and then a triplet of eighth notes. The dynamic then changes to *f*. The second part of the measure features a triplet of eighth notes, followed by a triplet of eighth notes, and then a triplet of eighth notes. The dynamic remains *f*. The measure concludes with a triplet of eighth notes and a final note. The instruction "embouchure" is written above the final note. The measure ends with a double bar line.

Musical score for measure 9 continuation. The notation is on a single staff. It begins with a *f* dynamic marking. The first part of the measure consists of a triplet of eighth notes, followed by a triplet of eighth notes, and then a triplet of eighth notes. The measure ends with a double bar line.

Musical score for measure 9 continuation. The notation is on a single staff. It begins with a *f* dynamic marking. The first part of the measure consists of a triplet of eighth notes, followed by a triplet of eighth notes, and then a triplet of eighth notes. The measure ends with a double bar line.

10 Tempo I



(simple and straightforward)



ANNEXE 3

SHIUREI IVRIT HEBREW LESSONS

by Hezi Leskali (1952-1994)

I. Shiur aleph

Harega ani lomed lichtov :
esh vechashmal, lelo shgiot.
Veharega ochezet esh beshulchan haktiva
vehchashmal bacheder
kave.
Le'oro shel hashulchan habo'er
ani lomed leta'er
et ha'esh
velaga'at bachashmal.

I. Lesson A

This very moment I am learning to write :
fire and electricity
without errors.
And this very moment, fire seizes the writing table
and the electricity in the room
goes out.
To the light of the burning table
I learn to describe
the fire
and to touch the electricity.

II. Shiur dalet

Ilmale hamila ha'atufa
be'arig shachor,
lo hayiti omer davar
al hamila hachatuma
bama'atafa ha'atuma.

II. Lesson D

Were it not for the word enveloped
in a black shawl,
I wouldn't say a thing
about the signed word
in the sealed envelope.

III. Shiur He

Keshe'hamila tahafoch leguf
vehaguf yiftach et piv
veyomar et hamila shmimena
notsar –
echabek et haguf haze
vealin oto letsidi.

III. Lesson E

When the word will turn to body
and the body will open its mouth
and speak the word from which it was
formed –
I will embrace this body
and lay it to sleep alongside me.

IV. Shiur Vav

Kisiti et hamila
besmicha
ve'amarti la :
layla
tov.
Vehatov haya ra
vehalayla haya tsiyur mufshat
al gabey hasmicha
vehamila lo teradem le'olam.

IV. Lesson F

I covered the word
with a blanket
and said :
good
night.
And the good was bad
and the night was an abstract drawing
on the blanket
and the word shall never fall asleep.

V. Shiur zayin

Milchey ha'haracha shel Madame Bonjour
hetsilu et kadur ha'arets
me'ilafon nitschi.
Kach to'enet Madame
Bonjour
bechol mesibot hacocktail
she'orechet shagrirut Israel
bepariz.
Ani lo yode'a tsarfatit
UMadame Bonjour eina yoda'at ivrit.
Lachen, muvan me'elav, ki bli ezrato shel meturgeman mevorach,
lo uchal lada'at
im hi shakranit
o doveret emet.

V. Lesson G

Madame Bonjour's smelling salts
saved the planet
from eternal unconsciousness.
So claims Madame
Bonjour
at all the cocktail parties
organized by the Israeli embassy
in Paris.
I don't know French
and Madame Bonjour doesn't know Hebrew.
Therefore, it is obvious, that without the help of a blessed translator,
I will never know
whether she is a liar
or whether she speaks the truth.

VI. Shiur chet

Keshe'afsik ledaber ivrit
ha'shulchan yomar : A-ba
Ve'ha'av yake ba'shulchan
be'egrofo
ve'yish'al : eyfo bni ?

VI. Lesson H

When I will stop speaking Hebrew,
the table will say : Pa-pa
and the father will strike the table with his fist
and ask : where is my son ?

I. SHIUR ALEPH שיעור א'

4/4 ♩ = 80 - 90

3/4

5/4

Flute: jet whistle, *f sf*, (embouchure), *ff*

Clarinet in B \flat : slap tongue, *f*

Piano: *f*, 8^{vb}

4/4 ♩ = 80 - 90

3/4

5/4

Soprano

Throughout this entire section, the strings are to be damped with the left hand so that no pitch is distinguishable. Only the harmonics (natural and artificial) should be played normally.

Violin: *f*, IV

Viola: *f*

Violoncello: *f*, 3, II

4

p

p *sf*

Sub

$\frac{4}{4}$ $\frac{2}{4}$

mf *f*

h - a re - ga a - ni lo - m e d li - ch

mp *f*

sul pont sul pont

mp *pizz* ϕ **3** *arco* *sul pont*

mp **3** *sf* **3** **3** **3**

4/4 **A** **2/4** **3/4** **5/4**

tongue ram

vary pitch by turning flute
 turn outwards → → → → →
 turn inwards → → → → →

sf *sf* *p* *f* *p*

f *sf* *sf*

f sf *8^{vb}* *8^{vb}* *8^{vb}*

4/4 **A** **2/4** **3/4** **5/4**

widen mouth → → → → →
ff
 narrow mouth → → → → →

p *mf*

e - sh ve - ch

⊕ damper string; no distinct pitch

repeat as much as necessary; go to m. 12 when cued by singer
mp *p* *f*
 sul pont

repeat as much as necessary; go to m. 12 when cued by singer
mp *p* *f*
 sul pont sul pont

f *sf* *mp* *mp* *mp* *mp*

pizz *arco*

f *sf* *mp* *mp* *mp* *mp*

f *mp* *mp* *mp* *mp* *mp*

11

p *f sf*

8^{vb} Ped.

cue players

f *p* *teacher-like*

a - sh m al le - lo sh - gi -

mf *f* *sul pont* *f*

sul pont *f*

4
4

C

3
4

21

sf

sf

mf

p

p

4
4

C

3
4

sf

ve.

mf

pizz arco pizz

mf

pizz arco

p

pizz arco pizz arco

sf

mf

p

pizz arco pizz arco

5/4

4/4

D

3/4

pp *f* *p* (embouchure) *gliss.* *sf*

3

5/4

4/4

D

3/4

p *sotto voce* *p* *gliss.* *p*

Le o - ro sh el ha shul ch an ha bo e r

sf(p) *gliss.* *pizz* *arco* *pizz* *arco* *pizz* *arco*

arco 3 *pizz* 3 *arco* 3 *pizz* 3 *arco* 3

30 **3/4** **E** **5/4** **4/4** **F** **2/4**

sf(p) *sf(p)*

sf(p) *pp*

3/4 **E** **5/4** **4/4** **F** **2/4**

p *p* whisper, no pitch till m. 38

a - ni lo - m e d le ta e r e - t ha

arco *pizz* *arco* *pizz* *arco* *pp* *arco* *pp*

2/4

4/4

3/4

a tempo

34

turn outwards →

turn inwards →

pp *mf* *f* *sf*

f *sf*

3

f *sf*

f *sf*

8th

8th 8th

2/4

4/4

3/4

a tempo

pp *f*

rit.

e - sh

ve la - ga - a - t ba cha-sh mal

f *sf*

pizz

pp *mf* *f*

3

3

pizz

3

f *sf*

rit.

gliss.

f

III. SHIUR HEH שיעור ה'

2
4

3
4

4
4

♩ = 54

Alto Flute

Piano

2
4

3
4

4
4

♩ = 54

Soprano

Violin

Viola

Violoncello

with clothes pin*

c.l. bat

behind pin**

III

l.h. pizz + on bridge

mf

p < *mf*

sf

f

*) place miniature clothes pin on G string, very close to the bridge.

***) place bow directly behind pin (between pin and fingerboard) so that the hair touches the pin as the tremolo is played.

4/4 2/4 3/4

5

4/4 2/4 3/4

mf p mf p mf sf f

bat. behind pin III pizz + ϕ jeté* on bridge

The musical score consists of a single melodic line on a treble clef staff. The first measure is in 4/4 time, the second in 2/4, and the third in 3/4. The dynamics are marked as *mf*, *p*, *mf*, *p*, *mf*, *sf*, and *f*. Technical markings include *bat.*, *behind pin*, *III*, *pizz +*, ϕ *jeté**, and *on bridge*. The score is divided into four measures corresponding to the time signatures.

*) begin with a jeté ; as the bow stops bouncing move towards the bridge until the bow is on the bridge only. This should be played as a single gesture.

4
4

2
4

9

mf k 3

4
4

2
4

c.l. jeté behind pin pizz + on bridge

mf *sf* *f*

2/4 **A** **3/4** **2/4**

12

pizz air closed emb 3 3 3

p *f* *sf* *f* *mf* *p*

k 3

2/4 **A** **3/4** **2/4**

pizz + on bridge flaut bat bat c.l. tratto

p *sf* *mf* *sf* *f*

2/4 3/4 2/4

15

pizz air cl emb tng ram cl emb

mf p f sf mf

2/4 3/4 2/4

pizz on bridge flaut bat jeté → tratto

pp sf mf sf f

2
4

18

mf

sf

mf

air

cl emb

tng ram

cl emb

f

B

2
4

B

pizz.

on bridge

flaut

bat.

bat. tratto

p

sf

f

mf

sf

f

21

air + sound overblow

3 3 3 3

p *sf* *f* *mf* *sf* *f* *mf*

brush

scratch coil

brush

mf *sf* *f* *mf*

3 4 2 4

3 4 2 4

sul pont

mf *mf* *mf*

3/4

3/4

24 air + sound ovrblw

3

p *sf* *f* *mf* *sf* *f* *f* *f* *sf*

brush

scratch

mf *f* *p* *f*

8^{va} Ped. *sf* *f* *p* *f* 8^{va} Ped.

3/4

3/4

sul pont

8^{va}

bwn pin and bridge

p *f* *p* *f*

3/4 **C**

4/4

2/4

27

ff

sf
ff

3/4 **C**

4/4

2/4

f sf

p

tsa

3

remove clothes pin

D

3
4

2
4

31

p

R.H.

L.H.

p

Ped.

D

3
4

2
4

unvoiced

be m

p sf *< p*

p

pizz

arco

pp

ppp

p

< p

pp

ANNEXE 6 :

EXEMPLES AUDIO

1. Extrait de « Vergibt mir Jesus mein Sünden »
2. « Quatre exemples de chaînes sonores », extrait du troisième disque du *Solfège de l'objet sonore* de Pierre Schaeffer.
3. Ablinger, *Quadraturen I*, « *Sprache ist* »
4. Ablinger, extrait de *Quadraturen II*, « *Raum der Erkenntnis / Vertreibung* »
5. Ablinger, *Quadraturen III*, « *A Letter from Schoenberg* »
6. Ablinger, extrait de *Quadraturen V*, « *Musik* »
7. *Ibidem.* superposée sur un extrait de *Auferstanden aus Ruinen* de Hanns Eisler
8. Harvey, extrait de *Speakings*
9. modèle acoustique : « one, nine, eight, four »
10. segmentation du modèle acoustique en mots
11. segmentation des mots en phonèmes
12. trois exemples du phonème « N » extrait de la segmentation en phonèmes
13. deux lectures du premier vers de *Leçons d'hébreu* : voix d'homme et voix de femme
14. segmentation du modèle acoustique en mots
15. modèle acoustique : « ilmale hamila haatufa bearig shakhor »
16. segmentation du modèle acoustique en mots
17. le mot « harega » et sa segmentation en phonèmes
18. le mot « likhtov » et sa segmentation en phonèmes
19. le mot « haatufa » et sa segmentation en phonèmes
20. le mot « bearig » et sa segmentation en phonèmes
21. le mot « shakhor » et sa segmentation en phonèmes
22. consonnes fricatives isolées par *op.recognize*
23. consonnes occlusives isolées par *op.recognize*
24. la consonne [x], *flutterzunge* à la flûte, éraflure d'une corde filée du piano, son écrasé au violon
25. la consonne [ʃ], tremolo au violon
26. la consonne [d], *tongue ram* à la flûte ; la consonne [t], pizzicato à la flûte, *col legno battuto* au violon
27. geste de la flûte dérivé de [d]

28. geste de la flûte dérivé de [j] et de [x]
29. geste de la flûte dérivé de [qi]
30. phrase terminant avec le *jet whistle*
31. phrase employant des divers sons percussifs de la flûte
32. la première page de *l'Auto-détermination*, d'abord chanté, puis joué à la flûte basse
33. texture pointilliste
34. Leçon A, premier exemple
35. Leçon A, deuxième exemple
36. Leçon A, troisième exemple
37. Leçon E, première partie
38. trois extraits de *Leçons d'hébreu* illustrant des idées similaires

Leçons d'hébreu – l'œuvre intégrale

39. I. Leçon A
40. II. Leçon D
41. III. Leçon E
42. IV. Leçon F
43. V. Leçon G
44. VI. Leçon H