



HAL
open science

La voix, un instrument de timbre

Nathalie Henrich Bernardoni, Michèle Castellengo

► **To cite this version:**

Nathalie Henrich Bernardoni, Michèle Castellengo. La voix, un instrument de timbre. Cahiers d'ethnomusicologie, 2021, 34, pp.17-36. <hal-03508088v2>

HAL Id: hal-03508088

<https://hal.science/hal-03508088v2>

Submitted on 11 Mar 2026

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire HAL, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.



Distributed under a Creative Commons CC BY-SA 4.0 - Attribution - ShareAlike - International License

La voix, un instrument de timbre. Comment joue-t-on du timbre vocal ?

Nathalie Henrich Bernardoni et Michèle Castellengo



Édition électronique

URL : <https://journals.openedition.org/ethnomusicologie/4283>

ISSN : 2235-7688

Éditeur

ADEM - Ateliers d'ethnomusicologie

Édition imprimée

Date de publication : 1 décembre 2021

Pagination : 17-36

ISBN : 978-2-88968-032-0

ISSN : 1662-372X

Référence électronique

Nathalie Henrich Bernardoni et Michèle Castellengo, « La voix, un instrument de timbre. Comment joue-t-on du timbre vocal ? », *Cahiers d'ethnomusicologie* [En ligne], 34 | 2021, mis en ligne le 01 décembre 2022, consulté le 05 décembre 2022. URL : <http://journals.openedition.org/ethnomusicologie/4283>

La voix, un instrument de timbre

Comment joue-t-on du timbre vocal ?

NATHALIE HENRICH BERNARDONI et MICHÈLE CASTELLENGO

La voix humaine peut produire et imiter toutes sortes de sons. S'appuyant sur les fonctions physiologiques (respiratoires, phonatoires, articulatoires) qui lui permettent de parler et de communiquer, le chanteur apprend à ajuster et à contrôler finement les diverses parties de son organe vocal pour modeler le timbre dans l'expression musicale. Nous disposons aujourd'hui de puissants moyens d'analyse pour étudier le fonctionnement de l'appareil vocal et les différents niveaux de son contrôle. Il est important néanmoins de prendre aussi en compte le fait que le timbre vocal s'inscrit fortement dans une culture au sein de laquelle priment les effets perceptifs et les interactions plus ou moins conscientes qui existent entre les chanteurs et les auditeurs en vue d'un résultat global à atteindre.

C'est dans ce contexte que nous analysons le lien entre geste vocal et timbre vocal en appuyant notre argumentaire sur diverses réalisations remarquables du timbre vocal mises en œuvre dans quelques chants traditionnels : les styles *teshka* et *leka* du chant féminin bulgare ; le *yodel* autrichien et le *tahrir* iranien ; trois styles de chants diphoniques : tuva, mongol, xhosa ; un chant polyphonique de Sardaigne et l'art vocalique des Aka de Centrafrique et des Bunun de Taïwan.

Nous abordons la complexité de la notion de timbre vocal et de sa perception en différenciant d'une part les techniques vocales du chant soliste qui s'appuient sur la pratique de gestes phonatoires et articulatoires spécifiques, et celles que l'on peut remarquer dans les polyphonies qui consistent soit en contraste de timbres entre les voix individuelles, soit en fusion de timbres différenciés en vue d'un effet global spécifique interagissant avec la sensation de hauteur.

Ces réflexions nous amènent aussi à questionner le croisement entre hauteur musicale, contenu vocalique et timbre.

La voix, un instrument de timbre

La voix humaine est par essence un instrument de timbre, c'est-à-dire un instrument offrant des possibilités infinies de variations de son timbre. Commençons par préciser ce que nous entendons par le mot « timbre », terme apparu au XVIII^e siècle dans l'*Encyclopédie*¹. Du point de vue acoustique, le timbre n'est pas un paramètre que l'on peut mesurer objectivement – comme la fréquence ou l'amplitude – mais un ensemble de propriétés des sons discriminées par les auditeurs, lesquelles correspondent à deux attitudes d'écoute.

La première attitude d'écoute, quasi instinctive, nous permet de reconnaître en un temps très bref ce qui a produit le son : la frappe d'une cloche ou d'une corde de piano, le son d'une voix humaine ou animale. C'est le timbre identitaire lié essentiellement au mode de production du son.

La deuxième attitude d'écoute s'établit lorsque nous entendons différentes productions d'un même type de source : voix ou instrument. Nous portons alors attention aux différences de qualité entre les sons et aux effets ressentis à leur écoute : c'est le timbre qualitatif et c'est généralement dans ce sens que les musiciens parlent du timbre, l'art d'un instrumentiste étant justement d'apprendre à varier à l'infini les qualités des sons que permet de produire une source instrumentale ou vocale (Castellengo 2015 : 287-296).

Remarquons tout d'abord que les recherches acoustiques sur le timbre concernent soit les instruments de musique, soit la voix parlée ou chantée. Alors en quoi la voix humaine tient-elle une position exceptionnelle dans l'étude du timbre, comme nous le signalions en introduction ? Parce qu'il s'agit d'un « instrument vivant », qui offre une plasticité exceptionnelle à tous les niveaux de la production du son : de la naissance du son (bruits de souffle ou sons harmoniques, entretenus ou percutés) jusqu'aux transformations spectrales extraordinaires que permettent les organes articulateurs de la langue, du voile du palais et de la bouche. Avec la voix humaine nous pouvons imiter une infinité de sons instrumentaux, animaux, des bruits de toutes sortes sans apport extérieur.

Face à une telle diversité, quels sont alors les indices identitaires d'une voix humaine ? Ce sont la reconnaissance dans le spectre de la trace sonore des résonances du conduit vocal par lesquelles nous percevons des voyelles et, lors d'une production chantée, la sensation instinctive d'un son harmonique vivant – un son qui bouge imperceptiblement sans être instable – qualités qu'il est bien difficile d'introduire dans une synthèse sonore. Quant aux indices qualitatifs, riches et complexes, nous les découvrons à l'œuvre dans les chants que nous analysons.

¹ Encyclopédie Diderot-D'Alembert, t. XV, p. 345. À l'entrée « SON, en Musique », l'auteur prend pour exemple la flûte et le hautbois dont il oppose les qualités de timbre et précise plus loin : « Que dirons-nous des différents timbres des

voix de même force et de même portée ? Chacun est juge de la variété prodigieuse qui s'y trouve ». Cette entrée est distincte de l'entrée *Tymbre* du t. XVI qui se rapporte à d'autres sens du timbre homonyme.

Le geste vocal dans sa complexité

Pour produire un son vocal, l'«instrumentiste-instrument» ajuste et coordonne un ensemble de gestes. La maîtrise de ce contrôle gestuel peut constituer l'apprentissage de toute une vie, ou être induite tout naturellement par l'imprégnation dans une culture vocale environnante.

La complexité du geste vocal provient des différents niveaux de son contrôle et de leurs interactions. La voix humaine est tout d'abord un instrument à vent, que l'instrumentiste vocal contrôle par l'intensité de son souffle. Selon la pression d'air expirée au moment du jeu vocal, le son sera plus ou moins fort, plus ou moins timbré. Cette pression d'air est une énergie aérodynamique qui est transmise et se transforme en énergie mécanique pour mettre en mouvement des structures dans le conduit vocal, depuis le larynx jusqu'aux lèvres, puis en énergie acoustique pour produire des sources sonores (Henrich 2012). La phonation est ainsi le résultat d'une interaction forte entre de l'air expiré de façon contrôlée et les parois mobiles que sont les plis laryngés (plis vocaux, plis vestibulaires, plis aryépiglottiques).

Si ces gestes sont invisibles pour nos yeux, leur signature acoustique est bien visible pour nos oreilles, qui les captent et les interprètent. Le mouvement très rapide et répété de mise en contact puis d'écartement des bords libres des plis vocaux se traduit acoustiquement par un signal harmonique, ayant de l'énergie acoustique répartie dans un ensemble de fréquences dites harmoniques, en relation les unes avec les autres : la fréquence fondamentale f_0 ou premier harmonique H_1 (Titze *et al.*, 2015) et tous ses multiples ($H_n = n \times f_0$). C'est la source glottique, en référence à cet espace entre les deux plis vocaux que constitue la glotte. À l'écoute d'un tel son, nous percevons une hauteur, dite hauteur tonale, qui est en lien avec la fréquence du premier harmonique² (Castellengo 2015).

D'autres plis laryngés que les plis vocaux peuvent également participer du jeu vocal. Dans de nombreuses cultures vocales à travers le monde, nous trouvons le rapprochement, la mise en contact et la vibration des plis vestibulaires³. Bien qu'étant des oscillateurs moins performants que les plis vocaux, ces deux plis laryngés peuvent, lorsqu'ils sont suffisamment rapprochés l'un de l'autre, se mettre à auto-osciller en interaction avec le flux d'air. Leur rapprochement et leur éventuelle mise en contact interagissent avec l'auto-oscillation des plis vocaux. Ainsi, un contact vestibulaire tous les deux cycles glottiques entraîne un doublement de la période du cycle glottique, puisqu'un cycle glottique sur deux est

² Ce qui, du point de vue temporel, correspond à la période du mouvement vibratoire.

³ Le terme «vestibulaire» provient de la nomenclature anatomique moderne. Ces plis peuvent également être désignés sous les termes de «bandes ventriculaires» ou «fausses cordes

vocales». Ce sont des plis situés au-dessus et à proximité des plis vocaux. Pour plus de détails sur leurs propriétés vibratoires et leurs usages dans la parole et le chant, voir la thèse de Lucie Bailly (2009) ou l'article de Bailly *et al.* (2014).

altéré. Cela se traduit directement sur le spectre harmonique qui présente alors un premier harmonique à l'octave inférieure ($f_0/2$) avec tous ses multiples.

La source glottique traverse différents espaces dans son parcours, ceux qui sont internes à notre instrument (le conduit vocal, composé de diverses cavités : supra-laryngée, pharyngée, buccale et éventuellement nasales) et l'espace extérieur jusqu'à nos oreilles. Tous vont laisser leur empreinte sur le son initial en modifiant la répartition d'énergie acoustique dans les harmoniques. Ces empreintes proviennent des propriétés de résonance des cavités. Le terme « formant » est parfois utilisé pour désigner ces zones spectrales renforcées. D'autres sources sonores vont éventuellement se rajouter à la source glottique, en particulier des bruits de friction lors de constriction, de plosion lors du relâchement d'une occlusion, de turbulence lors de la rencontre d'un obstacle comme les dents, les lèvres. Nous pouvons ajouter ici que notre perception sera également très sensible à la façon douce ou rapide dont le son débute et prend fin.

Il existe ainsi différentes façons de faire varier l'énergie acoustique dans les harmoniques et donc de jouer sur leur saillance perceptive. Le premier ajustement possible est global, contrôlé par l'énergie mise dans le souffle expiré. Plus la pression ou le débit d'air en entrée de l'instrument est élevé, plus les harmoniques sont nombreux et d'amplitudes élevées, et plus le niveau sonore est fort. Le contrôle global du gain peut être modulé par la résistance que le larynx oppose au souffle. Plus les plis vocaux sont contractés et fermement accolés, plus l'air doit « pousser fort » pour se frayer un passage. La phase de fermeture, selon qu'elle est douce ou abrupte, tend à modifier l'énergie des harmoniques de rang élevé : une remise en contact douce favorise l'énergie dans les premiers harmoniques de la voix, une remise en contact rapide favorisera l'énergie dans le second harmonique et, au-delà, au détriment du premier harmonique (Henrich 2001). La façon dont les plis vocaux s'accolent et auto-oscillent dépend des caractéristiques vibromécaniques des plis. Ainsi nous appelons mécanismes laryngés M1 et M2 des ajustements du vibrateur laryngé pour lesquels les plis vocaux sont épais (M1) ou minces (M2), reflétant par leur géométrie des couplages ou des dissociations au sein de la microstructure fibreuse et multicouche des plis (Roubeau *et al.* 2009). Une production vocale en mécanisme M2 se traduit systématiquement par une saillance du premier harmonique par rapport aux autres harmoniques, tandis qu'une production en M1 peut permettre de renforcer notablement le second harmonique.

Si les ajustements biomécaniques du vibrateur laryngé conditionnent le contenu harmonique de la source vocale, les gestes articulatoires comme le mouvement de la langue, l'ouverture de la bouche, l'élévation du *velum* ou l'abaissement laryngé vont également avoir une incidence majeure sur le spectre harmonique (Henrich Bernardoni 2014). Quand une cavité résonne, toute fréquence à proximité des fréquences de résonance est amplifiée, toute fréquence en dehors est atténuée. Dans le cas d'un signal vocal, les harmoniques dont les

fréquences sont situées à proximité des fréquences de résonance (ou zone formantique) sont renforcés par rapport à ceux qui s'en trouvent éloignés. Prenons un cas concret, celui d'un son vocal ayant pour hauteur f_{a_3} , ce qui revient à dire que la fréquence du premier harmonique vaut 440 Hz et celle du second 880 Hz. Si la fréquence d'une résonance R1 est à proximité de 440 Hz, par exemple $f_{R1}=400$ Hz, elle contribue à amplifier le premier harmonique. Si elle se situe à distance, par exemple $f_{R1}=600$ Hz, elle ne contribue pas grandement à l'amplification du premier harmonique, ni à celle du second. Néanmoins, si elle se situe à $f_{R1}=800$ Hz, elle contribue à renforcer le second harmonique, au détriment du premier.

À ce stade, il est temps d'introduire la notion de voyelle : notion linguistique complexe s'inscrivant entre production et perception. Les voyelles sont des sons que l'être humain produit pour s'exprimer verbalement en utilisant des gestes articulatoires précis, semblables d'un individu à l'autre par leurs lieux et modes d'articulation. Ce sont des composants de la parole, qui vont être perçus et décodés pour comprendre ce qui est dit. La perception d'une voyelle dépend du système phonologique de l'auditeur, c'est-à-dire de la langue qu'il a apprise. Si une voyelle ne figure pas dans son système phonologique, l'auditeur aura du mal à la percevoir comme différente d'une autre voyelle acoustiquement proche et aura également du mal à la produire. C'est par exemple le cas des voyelles non natives /y/ et /ø/ pour des italophones apprenants du français (Cornaz 2014). Les voyelles de la parole se projettent sur un espace à deux ou trois dimensions dont les dimensions principales sont les fréquences des deux premières résonances ou celles des deux premiers formants F1 et F2 qui les constituent. Cet espace acoustique a son pendant articulatoire : le degré d'ouverture de la bouche⁴ et le point d'articulation, avant ou arrière⁵ (fig. 1). Les caractéristiques acoustiques des voyelles, en particulier les fréquences des deux premiers formants F1 et F2 qui permettent de les discriminer, dépendent des gestes articulatoires et également, dans une moindre mesure, de la morphologie de la personne (enfant, adulte), du genre (femme, homme) et de sa famille linguistique.

Le chant combine la contrainte d'une mélodie, dont découlent les fréquences des harmoniques, à celle de l'expression d'un texte, avec des voyelles qui imposent des plages de variations possibles mais restreintes pour les fréquences de résonance. Comme nous allons le voir par la suite dans cet article, le chanteur a plusieurs stratégies à sa disposition : un maintien de la voyelle dans le grave de sa tessiture, un ajustement de la voyelle pour amplifier sélectivement un ou plusieurs harmoniques, une adaptation complète de la voyelle pour suivre la mélodie ou s'ajuster finement au sein d'une polyphonie.

4 Pendant articulatoire de l'axe acoustique selon F1. 5 Pendant articulatoire de l'axe acoustique selon F2.

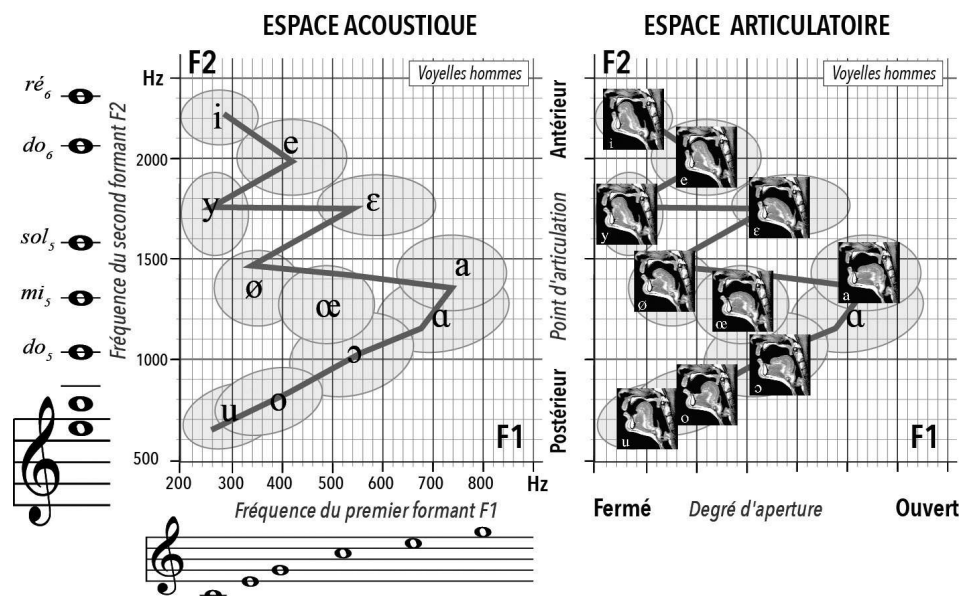


Fig. 1. Les voyelles orales du français, représentées par leur projection dans l'espace acoustique (deux premiers formants F1-F2) ou l'espace articulatoire (degré d'ouverture de la bouche – point d'articulation). D'après Castellengo 2015 : 452-457.

Jeu vocal et perception du timbre

La voix est toujours modifiée en vue d'un effet perceptif. La plupart des études sur le timbre de la voix chantée traitent du chant savant occidental de l'adulte, pour lequel l'effet perceptif recherché est celui d'une qualité globale homogène développée sur l'ensemble de la tessiture du chanteur, tout en respectant au mieux la reconnaissance des voyelles. L'étude des qualités vocales développées au sein de diverses cultures musicales du monde⁶ nous offre l'occasion de découvrir une grande diversité de création de timbres : le souhait d'une qualité générale du son, comme la clarté, la brillance ; la maîtrise d'une rupture mélodique par changement de mécanisme laryngé et le choix de voyelles qui en facilitent la rupture et contribuent à la saillance perceptive ; l'émergence d'une mélodie spectrale qui vient se superposer à la mélodie tonale, l'accompagnant ou l'éclipsant. La diversité des jeux du timbre vocal semble infinie.

⁶ Conscientes que l'étude des qualités vocales ne prend sens que dans une culture donnée, les auteures s'en tiennent ici à l'analyse acoustique

d'exemples choisis au titre de l'ingénieuse diversité des timbres.

Ajustements pour une qualité de timbre : cas d'un chant féminin de Bulgarie

Au sein de la diversité vocale développée dans les Balkans, nous nous arrêtons ici sur un style de chant féminin bulgare aux sonorités fascinantes et atypiques. Il s'agit à l'origine de chants monodiques ruraux et genrés, qui accompagnaient la vie quotidienne, depuis le travail dans les champs jusqu'aux fêtes qui rythmaient la vie des villages (Le Gonidec 2005). Ce style de chant met en avant la sonorité et la brillance d'un timbre vocal très singulier. Dans les chants traditionnels, les variations mélodiques portent sur une étendue d'une quarte (5 demi-tons) à une quinte (7 demi-tons), comprise dans une gamme de fréquences allant de si_2 ($f_0 = 247$ Hz) à si_3 ($f_0 = 494$ Hz). Il existe plusieurs dynamiques d'intensité possibles pour ce style de chant, comme la dynamique de forte intensité pour la qualité *teshka* (lourd) ou la dynamique plus douce pour la qualité *leka* (léger). Mais quelle que soit la dynamique, la caractéristique acoustique de ce timbre vocal est celle d'une prédominance du second harmonique par rapport aux autres harmoniques. Ce phénomène est illustré sur la figure 2.

Une étude mettant en œuvre une technique de mesure des résonances pendant la phonation a apporté un éclairage intéressant sur l'origine de cette caractéristique de timbre (Henrich *et al.* 2006). La figure 2 reprend ces résultats, en présentant la fréquence de la première résonance, R1, en fonction de la hauteur chantée pour toutes les voyelles du bulgare et toutes les hauteurs. Les voyelles pour lesquelles la fréquence de la première résonance est déjà de valeur assez élevée en parole, comme c'est le cas pour /a/, /o/, /œ/ et /e/, sont articulées de telle sorte que la première résonance vient systématiquement renforcer le second harmonique. Ce renforcement, que nous qualifions d'accord phono-résonantiel, est observé en *teshka* comme en *leka*, pour toutes les hauteurs chantées. La conséquence de cet accord phono-résonantiel choisi par la chanteuse pour renforcer la prédominance du second harmonique dans son timbre vocal est que la voyelle chantée va s'écarter de la voyelle parlée par le biais de l'adaptation de la première résonance à la note chantée. Nous pouvons observer cet éloignement dans l'espace acoustique $f_{R1} - f_{R2}$, tel que l'illustre la figure 2. Il est intéressant de noter qu'en langue bulgare, la seconde résonance a plus d'importance que la première dans la discrimination phonétique des voyelles. La chanteuse peut donc adapter la voyelle à la hauteur mélodique sans que cela altère grandement la compréhension du texte chanté.

Sonagramme du chant et analyses des résonances vocaliques

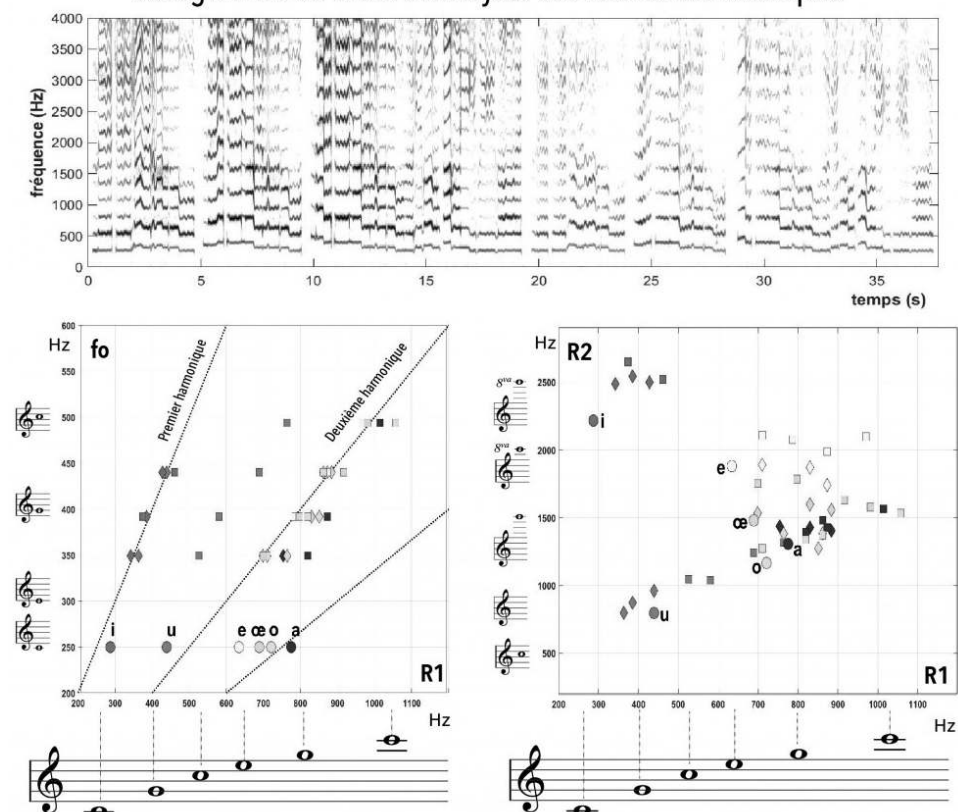


Fig. 2. *Partie supérieure* : analyse temps-fréquence d'un extrait chanté en *teshka* puis en *leka* par la chanteuse Mara Kiek (son n° 1). *Partie inférieure* : à gauche : évolution de la fréquence de la première résonance avec les notes chantées ; à droite : évolution des voyelles chantées dans l'espace vocalique R1-R2. Les marqueurs ronds correspondent aux voyelles parlées de la chanteuse, les carrés au style *teshka* et les losanges au style *leka*.



Chant bulgare, Maria Kiek. Source : Henrich, archives UNSW, Sydney Australie.

Ruptures mélodiques et ornements : le *yodel* et le *tahrir*

Si le chant bulgare féminin se caractérise par une grande stabilité dans l'expression d'un timbre vocal singulier qui est la signature de ces voix riches et sonores, de nombreuses expressions vocales à travers le monde jouent d'une toute autre façon du timbre vocal, par une recherche habile et répétée de la rupture mélodique et de la variation de timbre qui l'accompagne (Castellengo 1991 : 155-165).

C'est le cas de toutes les formes de *yodel*. Citons les pays germanophones alpins (Suisse, Autriche, Bavière), la musique country d'Australie et des Etats-Unis ou encore le *yodel* des pygmées de Centrafrique. Dans toutes les productions vocales yodlées, l'analyse acoustique permet de mettre en évidence,

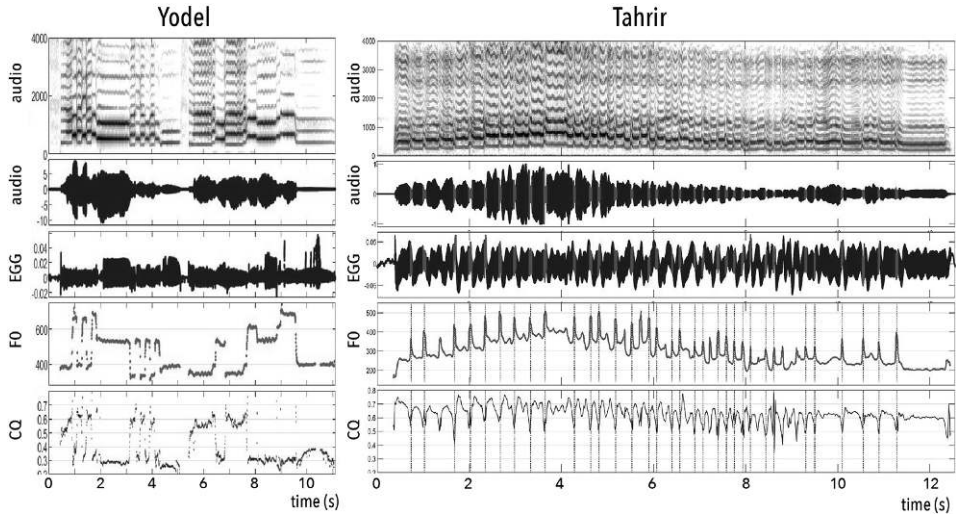


Fig. 3. A gauche, mélodie yodelée par une chanteuse suisse, Barbara Klossner (Lamesch 2010) (son n° 2) ; à droite, technique d'ornementation du *tahrir* iranien par Mohammad Reza Shâjariân (son n° 3). Analyses de haut en bas : sonagramme du signal audio, amplitude du signal audio, signal électroglottographique (EGG) représentant les variations de contact glottique au cours du temps, fréquence fondamentale, quotient de contact (CQ) dérivé de l'EGG représentant le temps de contact rapporté à la durée d'un cycle glottique (0 pas de contact pendant tout le cycle, 1 contact permanent).

Yodel, Barbara Klossner. Source : Lamesch, archives LAM, Paris, France.



Tahrir, Mohammad Reza Shâjariân. Source : Castellengo, archives LAM, Paris, France.



non seulement la rupture mélodique recherchée par le changement de mécanisme laryngé, mais aussi l'alternance d'amplitude entre les deux premiers harmoniques, provoquée par le changement de voyelle : voyelle ouverte chantée en mécanisme M1 (son grave avec H2 renforcé) et voyelle fermée chantée en mécanisme M2 (son aigu avec H1 renforcé) (Fürniss 1991 : 167-187).

La figure 3 (partie gauche) illustre ces ajustements laryngés à travers la mesure du contact glottique et leurs incidences sur le spectre harmonique.

L'art du *tahrir* persan joue aussi sur l'effet produit par un changement de mécanisme laryngé, mais d'une tout autre manière (Castellengo *et al.* 2020). La ligne mélodique reste en permanence en mécanisme M1 et l'emprunt au mécanisme M2, trop bref pour donner la sensation d'une « note », est perçu comme une brillante ornementation du chant : le *tekye*⁷. Dans les passages de virtuosité, on peut compter jusqu'à six *tekye* par seconde. L'homogénéité de timbre est renforcée par l'usage assez général de la voyelle /o/ qui renforce l'intensité de l'harmonique H2 sur toute la tessiture, comme illustré sur la figure 3 (partie droite).

⁷ *Tekeye*, nom donné à l'ornement : saut et bref emprunt au mécanisme laryngé M2.

Bien que jouant sur les mêmes paramètres, ces deux techniques de jeu de timbre s'opposent totalement du point de vue musical. L'une accentue les effets de rupture mélodique et vocalique, l'autre transforme l'effet de rupture en ornementation pour l'intégrer dans une esthétique de continuité mélodique et vocalique.

Emergence d'une mélodie spectrale en monophonie : le chant diphonique

Le chant harmonique – ou diphonique – est aujourd'hui bien connu. Il a fait l'objet de nombreuses recherches⁸ pour expliquer le fait intrigant qu'il est possible de chanter une mélodie alors que la fréquence fondamentale de la voix du chanteur – appelé diphonneur – reste constante. De nombreux artistes pratiquent ce chant « à l'occidentale », avec une voix quasi « pure » en ne sélectionnant qu'une succession d'harmoniques. Or, dans les pays d'origine des Mongols, des Bouriates ou des Touvas, le chercheur découvre une grande diversité de chants et de sonorités vocales⁹. Tous les diphonneurs sont capables de développer une voix puissante, riche en harmoniques et de tenir le souffle sur une longue durée¹⁰. Disposant d'une matière sonore riche, le diphonneur la sculpte en agissant sur les fréquences de résonance du conduit vocal et, en particulier, celles des voyelles qu'il peut contrôler finement pour cibler les notes de sa mélodie, les autres résonances contribuant aux qualités globales de sa voix.

La mélodie diphonique naît donc de l'accord entre un des harmoniques de la source et une des résonances du conduit vocal : c'est une mélodie spectrale (ou mélodie de timbre) qui exige de la part du diphonneur une grande précision dans l'ajustement du geste vocal. Bien que la source vocale soit unique, l'auditeur non averti peut avoir la sensation de deux sources sonores distinctes : un bourdon stable et une mélodie aiguë. Le bourdon peut être discret et la mélodie proche d'un sifflement : c'est la technique aiguë¹¹. Si le diphonneur utilise ses plis vestibulaires en couplage avec ses plis vocaux, le bourdon produit est grave et plus ou moins sonore¹². La vibration des plis vestibulaires altère la vibration glottique tous les deux cycles, induisant une période de la source vocale deux fois plus longue et un chant à l'octave basse (Bailly *et al.* 2010). C'est le phénomène du « doublement de période » ou *period-doubling*, terme parfois utilisé pour désigner ces « chants de gorge ».

⁸ P. ex. Zemp et Tran 1991.

⁹ Une étude approfondie des styles et techniques vocales d'un groupe de diphonneurs natifs de Mongolie est en cours (Curtet *et al.* 2021). Dans le cadre de cet article, nous faisons le choix de regrouper des productions diphonnées issues de cultures diverses.

¹⁰ Les diphonneurs Mongols s'entraînent face à un vent violent ou au pied d'une grande cascade (Curtet 2013).

¹¹ Technique désignée par les termes *khöömii*, *sygyt* ou *nariin* selon les cultures.

¹² Technique désignée par les termes *khargiraa* ou *budun* selon les cultures.

Pour diphoner, il faut habilement coordonner trois gestes complexes : ajuster les résonances du conduit vocal, sélectionner un groupe d'harmoniques aptes à générer une mélodie et choisir la fréquence fondamentale du son laryngé. Les résonances vocaliques – propres à un individu –, sont celles des deux formants F1 et F2 qui varient de façon coordonnée dans une zone assez bien définie, de 200 à 2000 Hz. La série des harmoniques est une donnée acoustique immuable. Du premier (H1) au sixième harmonique (H6), les intervalles entre harmoniques successifs sont respectivement l'octave, la quinte, la quarte, la tierce majeure et la tierce mineure : ceux d'une fanfare. C'est l'octave H6-H12 qui est utilisée comme plage mélodique par la plupart des diphoneurs, car elle renferme une succession d'intervalles proches d'une échelle pentatonique. Il s'agit d'intervalles « naturels » dits zarliniens : une quarte, constituée de deux intervalles proches d'une grande tierce mineure (H6-H7) et d'un grand ton (H7-H8), puis un ton majeur (H8-H9), un ton mineur (H9-H10), et une tierce mineure (H10-H12).

Compte tenu des deux contraintes précédentes, il n'est pas étonnant de constater que la fréquence fondamentale des voix diphoniques se situe, quel qu'en soit le style artistique, aux alentours de deux zones de fréquence : entre 170 et 210 Hz pour la technique aiguë, et entre 60 et 90 Hz pour la technique grave (Castellengo 2014)¹³.

La figure 4 permet de concrétiser musicalement ces deux solutions. Elle montre, dans sa partie gauche *a*), les deux groupes de voyelles disposés sur les portées musicales. Nous voyons *b*) que la technique aiguë ne peut s'appuyer que sur la seconde résonance (F2) alors que le chanteur de technique grave dispose des deux zones de résonance vocaliques. Le sonagramme *c*) d'un chant aigu montre que le chanteur sélectionne les harmoniques présents dans la zone de F2 qui concentre toute l'énergie ; F1 coïncide avec l'harmonique 2 de la voix et reste stable, ce qui explique l'absence de voyelle perceptible.

Deux réalisations très différentes d'une même mélodie populaire – *Artii-Sayir* – sont présentées sur la figure 5. Les formants F1 et F2 sont présents mais contrôlés différemment. En *a*), le chanteur choisit et sélectionne un groupe vocalique dont les formants F1 et F2 évoluent en parallèle, à l'octave l'un de l'autre. Il crée ainsi un timbre vocal clair sur un fondamental très grave (61,9 Hz, *si*₀) et la mélodie est très affirmée. En *b*), deux chanteurs à l'unisson ont choisi un parti contraire. Sur un fondamental de 74 Hz (*ré*₁), ils sélectionnent prioritairement des voyelles ouvertes de sorte que F2, constamment éloigné de F1 ne perturbe pas l'écoute de la mélodie. En chant diphonique, la présence constante des voyelles enrichit l'écoute des auditeurs qui peuvent, à tout moment, basculer de l'écoute d'une mélodie vocalique à celle d'une mélodie de « notes ».

¹³ Ces résultats ont été calculés sur 18 chanteurs d'origines diverses : 11 en style grave et 7 en style aigu.

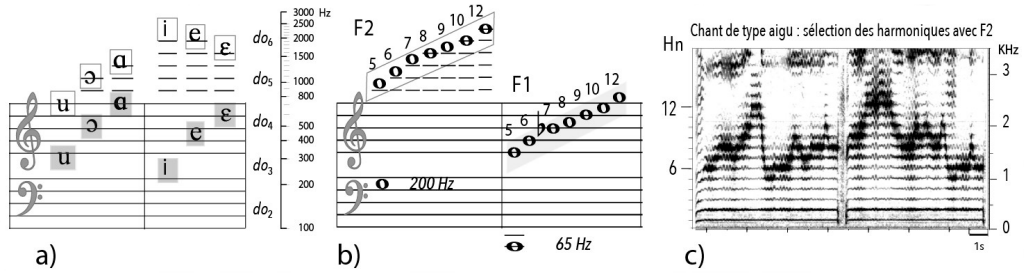
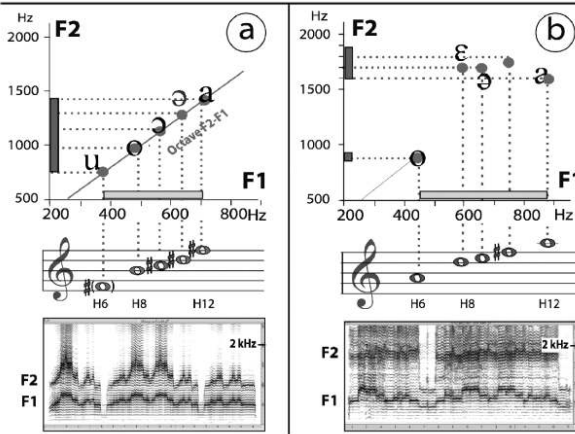


Fig. 4. a) position musicale des voyelles selon leurs formants F1 et F2; b) tessiture des harmoniques utilisés selon le fondamental de la voix; c) chant mongol de style aigu par Taravaj Ganbold (son n° 4).



Mongol aigu, Taravaj Ganbold. Source: Ganbold, *Mongolie*, CD Inédit, 1989, Auvidis W 260009, piste 06.

Deux chants de type grave (même mélodie) - Contrôles différents de F1 et F2



Chant de type grave à deux fondamentaux

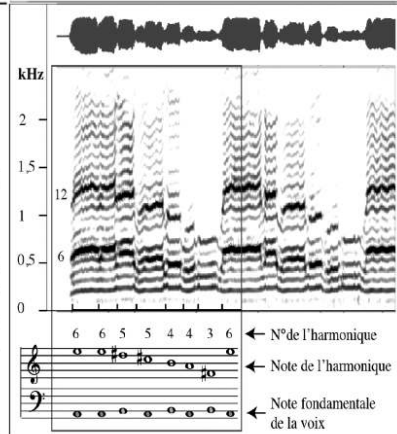


Fig. 5. a) et b) deux chants Touvas en *period-doubling* basés sur une mélodie très populaire, *Artii-Sayir*. À gauche a), Vassili Chazir (son n° 5); au milieu b), Tumat Kara-ool et Andrei Chuldum-ool (son n° 6); à droite c), chant xhosa. (son n° 7).



Khargiraa 1, Vassili Chazir. Source: *Voices from the center of Asia*, 1990, CD Smithsonian Folkways, piste 18.



Khargiraa 2, Duo Tumat Kara-ool et Andrei Chuldum-ool. Source: *Voices from the center of Asia*, 1990, CD Smithsonian Folkways, piste 09



Chant Xhosa, M^{me} Mbizweni, Dargie. Source: *Les voix du monde*, CD CNRS, CD-II-36.

L'exemple étonnant d'une femme xhosa – virtuose du chant grave en *period-doubling* –, présenté figure 5 à droite, couronne ce paragraphe (Dargie 1993, Castellengo 2015: 249). Pour compenser le fait que la fréquence fondamentale de sa voix est trop aiguë pour cibler la bonne tranche harmonique, la chanteuse alterne deux fondamentaux à intervalle d'un ton, 110 et 124 Hz ($/a_1$ et $/si_1$), et développe la mélodie avec les harmoniques de rang 3 à 6, renforcés à l'octave par le choix du groupe des voyelles $/a/$ et $/o/$. Cette technique s'inspire du jeu de l'arc musical.

La sélection harmonique est au cœur de la pratique du chant diphonique qui n'est ni une illusion perceptive ni une mélodie purement spectrale et suscite bien des ambiguïtés. À la première écoute, la plupart des musiciens de culture occidentale captent immédiatement la « note » fondamentale de la voix chantée, ce qui leur suffit. Certains s'intéressent aussi aux variations sonores qui se produisent dans l'aigu de la voix. Dans tous les cas, un long moment d'écoute, et la connaissance qu'il s'y trouve une mélodie sont nécessaires pour soudainement la percevoir et prendre plaisir à découvrir l'immense variété de timbres des productions diphoniques.

Emergence d'une mélodie spectrale en polyphonie : cas de la *quintina* d'un chant liturgique sarde, le « *jesu* »

Cet article serait incomplet sans la mention d'un phénomène perceptif connu sous le nom de *quintina*. Il s'agit d'un chant religieux de Sardaigne – le *jesu* – richement documenté par Bernard Lortat-Jacob (Lortat-Jacob 1998), au cours duquel il est possible d'entendre une voix aiguë qui couronne la polyphonie, deux octaves au-dessus de celle du *bassu*. S'agit-il d'une illusion ? Le phénomène perceptif qui sous-tend la perception d'une « note » musicale dépourvue d'énergie sur la fréquence fondamentale¹⁴ est connu en acoustique (Castellengo 2015: 116, 157) sous le nom de « résidu » ou « son résultant » (Terhardt 1982). Il correspond au codage temporel de la hauteur tonale des sons musicaux¹⁵. Seul un travail de filtrage puis de synthèse avec les sons prélevés dans l'enregistrement original nous a permis de démêler et de confirmer les hypothèses perceptives dans le cas exceptionnel d'une polyphonie vocale.

Écoutons le début du chant (son n° 8a, fig. 6a). Fasciné par la beauté des quatre voix et par la plénitude sonore de la polyphonie, l'auditeur entend rarement la *quintina* d'emblée mais l'expérience suivante sur un accord extrait du chant (son n° 8b, fig. 6b) en révèle l'existence en cours d'écoute: 1) accord

¹⁴ Par exemple un $so/2$, 200 Hz, entendu lors de la présentation des seuls harmoniques 2 et 3, $so/3$ et $ré_3$ (400 et 600 Hz).

¹⁵ Par opposition au codage tonotopique de la hauteur spectrale.

a)  Notation musicale de l'extrait sonore

b) Expérience d'écoute de la quintina : filtres et association des deux sons filtrés

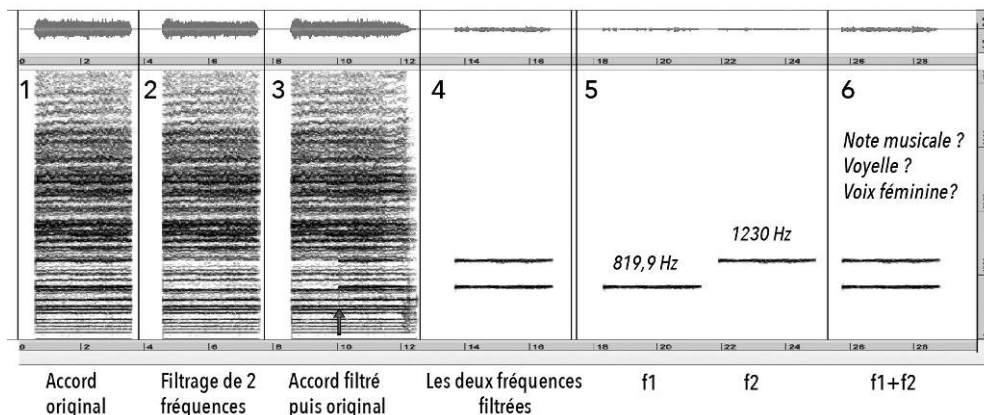


Fig. 6. a) phrase musicale du son n° 8a; b) étapes de l'expérience de filtrage et de resynthèse de la voix créée par deux harmoniques accordés aux résonances vocaliques d'un « a »; son n° 8b (1 à 6).



Quintina, chanteurs de Castelsardo, première phrase du *jesu*. Source: Lortat-Jacob, CD du livre *Chants de Passion*, 1998.

normal; 2) accord privé de deux fréquences; 3) réentrée des deux fréquences au milieu de l'accord; 4) les deux fréquences filtrées isolées qui correspondent à la voix de la *quintina*; 5) écoutés isolément, les deux sons extraits de l'enregistrement se succèdent comme deux sons purs de fréquence instable (819,9 Hz et 1230 Hz). Écoutés simultanément 6), nous pouvons avoir la surprise d'entendre une voix féminine chantant un [a] à 409,9 Hz¹⁶, soit un *la*₄! Au regard de l'ensemble des quatre chanteurs qui comptabilisent un très grand nombre d'harmoniques¹⁷, seuls deux harmoniques suffisent pour évoquer une cinquième voix qui disparaît lorsqu'on supprime l'un d'eux!

Il s'agit une fois encore d'un « jeu » formantique avec les résonances des voyelles. Chaque auditeur pourra remarquer que la deuxième voyelle du *Jesu* chanté est intermédiaire entre un [a] et un [o] très ouvert, mais en aucune manière un « ou » ([u])¹⁸! Parmi les confréries de Sardaigne seules quelques-unes – dont celle de Castelsardo – parviennent à créer les conditions d'un timbre vocal particulier pour faire émerger la voix de la *quintina*. Il leur faut conjuguer strictement des conditions très particulières (Castellengo 2015: 467-471): que les quatre chanteurs aient

¹⁶ Soit un *lab3* bas d'un quart de ton.

¹⁷ De l'ordre de 200 rien que pour la voix de *bassu*!

¹⁸ Voir Lortat-Jacob (1998: 129) pour une analyse des modifications des mêmes voyelles parlées et chantées.

des voix riches en harmoniques, stables et sans vibrato ; qu'ils utilisent certaines voyelles dont ils ajustent les formants F1 et F2 à l'oreille ; que le *bassu*, qui supporte l'harmonie, choisisse sa note aux alentours de 100 Hz afin que les harmoniques 2 et 3 du son résultant se situent dans la zone de fréquence dite « dominante » ; et enfin que les voix fusionnent selon une échelle quasi-harmonique en se calant sur les intervalles naturels, sans battement. Tous les auditeurs n'entendent pas la cinquième voix à la première écoute, mais dès que le phénomène est repéré, il ne les lâche plus. Cet art du timbre vocalique qui suggère « une voix qui n'existe pas » étonne d'autant plus que le spectre harmonique des chanteurs est extrêmement dense et que seuls deux harmoniques suffisent à la faire apparaître.

Registration vocalique des chanteurs d'une polyphonie Aka

La musique des pygmées Aka est essentiellement vocale et polyphonique. La structure musicale des chants, de type périodique, s'inscrit dans un rythme complexe et rigoureux. À l'ordinaire, un chanteur lance un incipit, développe des variations et rapidement le groupe entier des chanteurs présents se partage les différentes parties de la polyphonie, purement vocalique, qui peut compter jusqu'à huit parties. Pour démêler la complexité sonore de ces polyphonies et en comprendre la structure intervallique, Simha Arom a utilisé la technique du *re-recording* (Arom 1976). Lors de l'analyse sonographique des voix isolées nous

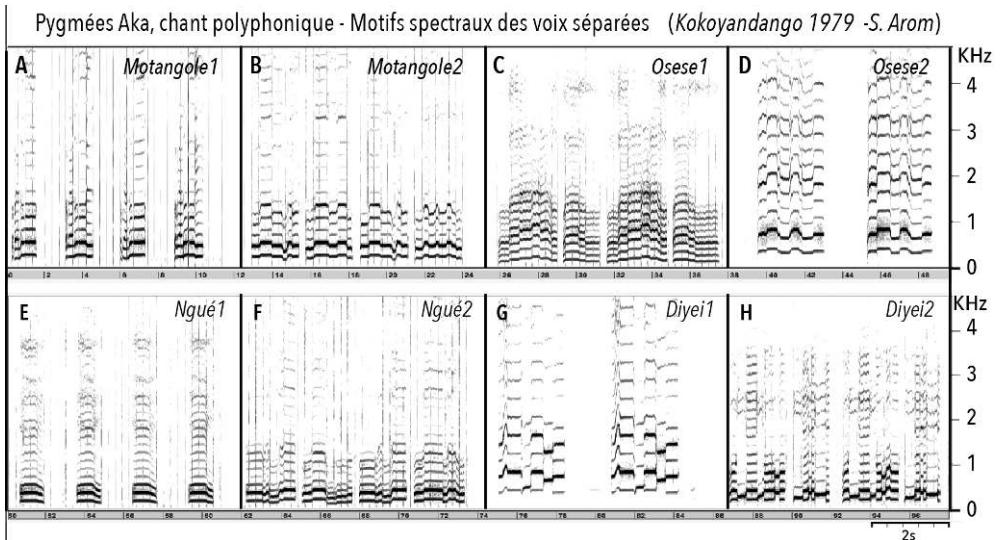


Fig. 7. Analyse spectrale de chacune des voix d'un chant polyphonique des Pygmées Aka (son n° 9).

avons eu la surprise de découvrir que chaque chanteur adoptait un choix de voyelles – principalement /u/, /a/, /o/, /e/, /ɛ/ – et qu'il s'y tenait strictement du début à la fin du chant, ce qui peut sembler paradoxal au regard de l'habileté des chanteurs Aka à varier et à complexifier leur partie.

L'analyse de deux périodes pour chaque partie présentée sur la figure 7 met en évidence le rôle des voyelles dans le suivi précis de certains harmoniques : H1-H2 (A; D); H1-H2-H3 (E); H1-H2-H5 (B); H1-H2-H6 (F); l'alternance H1-H5 et H1-H3-H4 (C). Le yodel des deux *díyeyi* (G et H) impose l'alternance H1 ou H2.

Nous n'avons pas d'indications sur ces choix de voyelles ni sur le sens que les Aka donnent à cette pratique remarquable¹⁹.

Un art du timbre vocalique : la polyphonie Bunun du *pasi but but*

Le *pasi but but* – chant de prière pour une récolte abondante du millet – qui est pratiqué une fois par an par les Bunun aborigènes de Taiwan a fait l'objet d'une étude approfondie (Wu 1996 : 180-248). Ce chant représente une forme musicale unique, associant mouvements mélodiques en glissando et entrée des voix par paliers successifs (cf. Discographie, CD Voix III-06). C'est aussi un chant particulièrement raffiné du point de vue des timbres vocaliques, c'est pourquoi nous l'avons retenu dans ce travail. La version que nous présentons provient de Naifubo (1987), un « village reconnu comme celui entre tous où l'on chante le *pasi but but* de la façon la plus proche de l'archétype » (Wu 1996 : 252). Avant de poursuivre, nous recommandons d'écouter l'exemple sonore (n° 10).

Le choix des tessitures et des voyelles témoigne d'une exceptionnelle sensibilité à l'équilibre spectral de la polyphonie, comme le montre l'analyse spectrographique de la figure 8. La voix supérieure (v1) qui est la plus importante, est chantée avec la voyelle /o/ par trois hommes qui se relaient pour que le son soit continu sur toute la durée du chant (3'35), tout en effectuant une montée insensible de la hauteur (de 192 à 255 Hz soit une quarte, pour cet exemple). Les autres voix s'insèrent selon une succession descendante, en empruntant d'autres voyelles²⁰ : v2 (un homme) à la tierce inférieure de v1, voyelle /e/ ; v3 (deux hommes) un ton plus bas que v2, voyelle /e/ ; v4 (deux hommes) un ton plus bas que v3, voyelle /i/. Puis un nouveau cycle s'effectue en progressant insensiblement vers l'aigu et le chant se termine sur une quinte pure par l'accomplissement du sixième cycle dit de « la plus pleine lune ».

¹⁹ La stratégie que les chanteurs Aka développent spontanément évoque celle des polyphonies instrumentales traditionnelles dont les parties, loin de fusionner, affirment clairement leur indépendance, tant mélodique que rythmique.

²⁰ Les Bunun ont des termes propres pour qualifier le « timbre » des voix (Wu 1996 : 220).

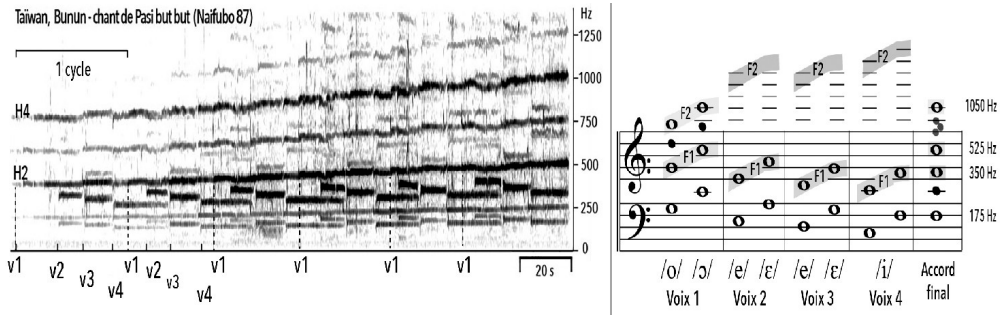


Fig. 8. Sonogramme du chant de *pasi but but* avec indications des entrées des voix et représentation musicale de la position des formants vocaliques pour les quatre voix de la polyphonie Bunun (Castellengo 2015: 461) (son n° 10).

Voici maintenant la description du jeu vocalique entre voyelles ouvertes et voyelles fermées (figure 8).

L'important renforcement des harmoniques H2 et H4 de la voix 1, visibles sur le sonogramme, correspond aux deux zones de résonance – F1 et F2 – de la voyelle /o/, soit environ 400 et 800 Hz (pour une voix d'homme). Pour les voix 2, 3 et 4, c'est aussi l'harmonique H2 qui est renforcé par F1. Cependant /ɛ/, /e/ et /i/ étant des voyelles fermées, le deuxième formant F2 est supérieur à 2 kHz. Les autres harmoniques de ces voix sont faibles et n'interfèrent pas avec la voix supérieure qui reste dominante. Enfin, pour que la plénitude sonore de ce chant rituel s'accomplisse sur l'accord terminal, les Bunun doivent décider soigneusement de la fréquence initiale²¹ et adapter finement leur voyelle au cours de la progression vers l'aigu, pour rester en accord avec les formants vocaliques.

Chant Bunun. Source : archives Wu Rung Shun 1996, Wind Records; *The Music of the Aborigine on Taiwan Island*, Vol 1; piste 1 (3'35).



Conclusion

Dans l'univers de la musique, la voix humaine est une source sonore « à part ». Elle est l'expression d'un instrument vivant, d'un instrument incarné dont le son n'est jamais totalement stable et porte l'empreinte, immédiatement reconnaissable, de la trace spectrale des résonances vocaliques.

Comme pour toute pratique instrumentale, le chanteur développe et travaille sa voix. Luthier de son propre instrument, il le façonne au sein d'une culture donnée, musicale, vocale, mais aussi linguistique, en quête d'une esthétique personnelle ou partagée collectivement.

²¹ Le choix de la note initiale est de la responsabilité d'un chamane qui est aussi le meneur du chant (Wu 1996: 219-223).

Le timbre vocal est l'expression sonore de son jeu d'équilibriste habile, entre souffle, mouvements et résonances. Le chanteur peut sélectionner un mode d'émission par le choix du mécanisme laryngé dans lequel il s'exprime, jouer sur les contrastes – comme dans les techniques de *yodel* et de *tahrir* – ou au contraire les continuités mélodiques, en coordination et interactions avec les résonances vocaliques de sa langue. Il forge l'identité de son timbre harmonique, comme le montre le chant féminin bulgare. La musique naît des jeux vocaliques, des choix de tessitures et du nombre infini des ornements : c'est une musique des variations de timbre, superbement illustrée par les nombreuses facettes du chant diphonique. Le chanteur tout entier est un instrument de timbre.

Les pratiques polyphoniques offrent encore d'autres possibilités de jeu sur le timbre vocal, par la combinaison de plusieurs timbres dont chacun conserve sa spécificité (polyphonie des pygmées Aka), ou qui s'unissent pour faire surgir une voix, comme celle de la *quintina* dans le chant polyphonique traditionnel de Sardaigne, ou encore qui élaborent une registration des tessitures et des timbres vocaliques selon un rituel immémorial comme celui de la polyphonie Bunun.

Ces pratiques vocales mises en lumière et disséquées sous le regard de l'analyse acoustique et physiologique sont tout à la fois uniques et plurielles. Toutes témoignent par leurs spécificités, de l'extraordinaire plasticité de l'instrument vocal humain, de la richesse des expressions et de la pluralité des timbres de la voix.

Références

- AROM Simha
1976 «The use of play-Back Techniques in the Study of Oral Polyphonies», *Ethnomusicology* XX/3: 483-519.
- BAILLY Lucie
2009 *Interaction entre cordes vocales et bandes ventriculaires en phonation: exploration in-vivo, modélisation physique, validation in-vitro*. Thèse de Doctorat de l'Université du Maine. URL : <https://tel.archives-ouvertes.fr/tel-00416587>
- BAILLY Lucie, Nathalie HENRICH BERNARDONI et Xavier PELORSON
2010 «Vocal fold and ventricular fold vibration in period-doubling phonation: Physiological description and aerodynamic modeling», *The Journal of the Acoustical Society of America* 127(5): 3212-3222.
- BAILLY Lucie, Nathalie HENRICH BERNARDONI, Franck MÜLLER, Anna-Katharina ROHLFS et Markus HESS
2014 «The ventricular-fold dynamics in human phonation», *Journal of Speech, Language, and Hearing Research* 57: 1219-1242.
- CASTELLENGO Michèle
1991 «Continuité, rupture, ornementation. Ou les bons usages de la transition entre deux modes d'émission vocale», *Cahiers de musiques traditionnelles* 4 <http://journals.openedition.org/ethnomusicologie/1583>
- 2014 «Perception(s) de la voix chantée: une introduction», in Henrich Bernardoni, dir.: *La voix chantée, entre sciences et pratiques*. Louvain-la-Neuve: Ed. De Boeck, Coll. Voix parole langage (Solal): 35-64.
- 2015 *Ecoute musicale et acoustique: avec 420 sons et leurs sonagrammes décriptés*. Paris: Eyrolles.

- CASTELLENGO Michèle, Jean DURING et Nathalie HENRICH BERNARDONI
 2020 «Electroglottographic assessment of Tahrir, a Persian vocal technique», International Conference on Voice Physiology and Biomechanics ICVPB2020 December Edition, Grenoble, France.
- CASTELLENGO Michèle et Nathalie HENRICH BERNARDONI
 2014 «Interplay between harmonics and formants in singing: when vowels become music», *International Symposium on Musical Acoustics, Jul. 2014*, Le Mans, France.
- CASTELLENGO Michèle, Sylvain LAMESCH et Jean DURING
 2009 «Laryngeal mechanism change and frequency jump: the Persian *tahrir*», Poster presented at the 8th *international conference PEVOC*, Dresden, Germany.
- CURTET Johanni, Nathalie HENRICH BERNARDONI, Michèle CASTELLENGO, Christophe SAVARIAUX et Pascale CALABRESE
 2021 «L'art du Khöömii mongol». Actes des Rencontres Nationales de Recherches en Musique. Paris: Ministère de la Culture (en cours d'édition).
- CASTELLENGO Michèle, Bernard LORTAT-JACOB et Gilles LÉOTHAUD
 2001 «Pitch perception: five voices with four Sardinian singers», *International Symposium on Musical Acoustics*, Perugia, Italy.
- CURTET Johanni
 2013 *La transmission du höömij, un art du timbre vocal: ethnomusicologie et histoire du chant diphonique mongol*. Thèse de doctorat, Université Rennes 2.
- DARGIE Dave
 1993 «Thembu Xhosa umngqokolo overtone singing: the use of the human voice as a type of "musical bow"». Conférence présentée à l'International Council on Traditional Music (ICTM), Berlin.
- FÜRNISS Susanne
 1991 «La technique du jodel chez les Pymées Aka (Cenrafrique). Étude phonétique et acoustique», *Cahiers de musiques traditionnelles* 4.
- HENRICH Nathalie
 2001 *Etude de la source glottique en voix parlée et chantée: modélisation et estimation, mesures acoustiques et électroglottographiques, perception*. PhD thesis, Université Paris 6.
 2012 «Physiologie de la voix chantée: vibrations laryngées et adaptations phonorésonantielles», in Renaud Garrel, Benoît Amy de la Bretèque & Vincent Brun, dir: *La voix parlée et la voix chantée*. Montpellier, Sauramps médical: 17-32.
- HENRICH BERNARDONI Nathalie
 2014 «Vibrations et résonances en voix chantée», in Henrich Bernardoni, dir: *La voix chantée, entre sciences et pratiques*. Louvain-la-Neuve: De Boeck, Coll. Voix parole langage (Solal): 3-20.
- HENRICH Nathalie, Mara KIEK, John SMITH et Joe WOLFE
 2006 «Resonance strategies used in Bulgarian women's singing style: A pilot study», *Logopedics Phoniatrics Vocology* 32/4: 171-177.
- LAMESCH Sylvain
 2010 *Mécanismes laryngés et voyelles en voix chantée: Dynamique vocale, phonétogrammes de paramètres glottiques et spectraux, transitions de mécanismes*. Thèse, Université Paris 6.
- LE GONIDEC Marie-Barbara
 2005 «Du patriarcat au communisme: les femmes et la musique en Bulgarie», *Cahiers de musiques traditionnelles* 18: 173-182.
- LORTAT-JACOB Bernard
 1998 *Chants de Passion. Au cœur d'une Confrérie de Sardaigne*. Paris: Éditions du Cerf.
- ROUBEAU Bernard, Nathalie HENRICH et Michèle CASTELLENGO
 2009 «Laryngeal vibratory mechanisms: the notion of vocal register revisited», *Journal of Voice* 23/4: 425-438.

TERHARDT Ernst, Gerhard STOLL et Manfred SEEWANN

1982 «Pitch of complex signals according to virtual-pitch theory: tests, examples and predictions», *Journal of the Acoustical Society of America* 71 : 671-678.

TITZE Ingo R., Ronald J. BAKEN, Kenneth W. BOZEMAN, Svante GRANQVIST, Nathalie HENRICH *et al.*

2015 «Toward a consensus on symbolic notation of harmonics, resonances, and formants in vocalization», *Journal of the Acoustical Society of America* 137/5 : 3005-3007.

WU Rung Shun

1996 *Tradition et transformation : le pasi but but, un chant polyphonique des Bunun de Taiwan*. Thèse, Université Paris X, Nanterre.

ZEMP Hugo et Tràn Quang Hai

1991 «Recherches expérimentales sur le chant diphonique», *Cahiers de musiques traditionnelles* 4 : 27-68.

Sons et discographie

Son 1 (fig. 2) : Chant Bulgare, Maria Kiek. Source: Henrich, archives UNSW, Sydney Australie.

Son 2 (fig. 3) : Yodel, Barbara Klossner. Source: Lamesch, archives LAM, Paris, France.

Son 3 (fig. 3) : Tahrir, Mohammad Reza Shâjariân. Source: Castellengo, archives LAM, Paris, France.

Son 4 (fig. 4) : Mongol aigu, Tarajav Ganbold. Source: Ganbold, Inédit Mongolie, 1989, Auidis W 260009, piste 06.

Son 5 (fig. 5 gauche) : Khargiraa1, Vassili Chazir. Source: CD Voices from the center of Asia, 1990, Smithsonian Folkways, piste 18.

Son 6 (fig. 5 milieu) : Khargiraa2, Duo Tumat Kara-ool et Andrei Chuldum-ool. Source: Voices from the center of Asia, 1990, Smithsonian Folkways, piste 09.

Son 7 (fig. 5 droite) : Chant Xhosa, M^{me} Mbizweni, Dargie. Source: Les voix du monde, CD CNRS, CD-II-36.

Son 8 (fig. 6a) : Quintina, chanteurs de Castelsardo, première phrase du *jesu*. Source: Lortat-Jacob, CD du livre *Chants de Passion*, 1998.

Son 9 (fig. 6b) : Expériences de filtrages et de recombinaisons des deux fréquences qui génèrent la cinquième voix; Castellengo 2001 et 2015.

Son 10 (fig. 7) : Polyphonie Aka. Source: Archives personnelles Arom, 1979.

Son 11 (fig. 8) : Chant Bunun. Source: archives Wu Rung Shun 1996, Wind Records; The Music of the Aborigine on Taiïwan Island, Vol 1, piste 1 (3'35).

RÉSUMÉ. Tout à la fois luthier, instrument et instrumentiste, le chanteur ajuste et contrôle finement les diverses parties de son instrument pour «jouer du timbre vocal». Le lien entre gestes et timbres est analysé en s'appuyant sur diverses réalisations remarquables mises en œuvre dans quelques chants traditionnels du monde: le chant féminin bulgare, la technique du *yodel*, le *tahrir* iranien, le chant diphonique et quelques polyphonies de Sardaigne, de Centrafrique et de Taiïwan. La diversité et la précision de ces réalisations est l'expression sonore d'un jeu d'équilibre entre le souffle, la vibration de plis laryngés et les résonances vocaliques du conduit vocal. Les pratiques vocales mises en lumière dans cet article et disséquées sous le regard de l'analyse acoustique et physiologique témoignent de l'extraordinaire plasticité de l'instrument vocal humain et de la richesse inouïe des jeux de timbres des chanteurs.