

Perception et description acoustique de la qualité vocale dans le chant lyrique

Maëva Garnier, Nathalie Henrich, Michèle Castellengo, Danièle Dubois, Jacques Poitevineau

LAM (UPMC, CNRS, Ministère de la culture), 11 rue de Lourmel, 75015 Paris

Tel. : +33 (0)1 53 95 43 34 – Fax : +33 (0)1 45 77 16 59

Mèl : {garnier, henrich}@lam.jussieu.fr, castel@ccr.jussieu.fr, ddubois@ccr.jussieu.fr, poitevin@ccr.jussieu.fr

ABSTRACT

This paper presents a method for finding outstanding physical parameters which should contribute to the perception of western operatic vocal quality. Based on the results of a previous cognitive study, presented in a companion paper [4], this methodology emphasises the listening point of view and endeavours thinking about the relevance of the evaluation criteria and of the realised analyses. It is shown how the enhancement of the first two formants or of the singing formant may contribute to the perception of brightness ("brillance"), of the character "voix timbrée" or of the presence of air ("air sur la voix").

1. INTRODUCTION

De nombreuses études ont montré l'influence de paramètres articulatoires, acoustiques, spectraux, ou glottiques sur la qualité vocale perçue de la parole [1][2]. Le domaine du chant lyrique possède un vocabulaire spécifique et une conceptualisation de la voix qui lui sont propres, rendant inappropriée l'évaluation de voix lyriques selon des critères provenant de la littérature sur la parole.

L'étude présentée dans cet article s'attache à rechercher les corrélats acoustiques de critères verbaux utilisés pour décrire la qualité vocale dans le chant lyrique. Elle s'appuie sur les résultats d'une analyse cognitive de la perception de la voix lyrique [4], qui a permis de dégager des qualités vocales ayant un sens dans le contexte du chant lyrique et renvoyant à une même conceptualisation pour la majorité d'experts interrogés de ce domaine.

Dans un premier temps, nous allons décrire l'expérience psychoacoustique menée avec des professeurs de chant, puis nous présenterons les analyses acoustiques réalisées et nous discuterons les résultats obtenus.

2. PROTOCOLE

2.1. Stimuli

Le corpus de sons est constitué par les enregistrements audio de 3 voix d'hommes (2 basses-barytons, B3 et B6, et un ténor, T2), effectués dans une cabine insonorisée [3]. Le signal audio est enregistré à l'aide d'un microphone de pression 1/2" (Brüel & Kjaer 4165) placé à 50 cm des lèvres des chanteurs, puis amplifié grâce à un préamplificateur (Brüel & Kjaer 2669) et un

amplificateur de mesure (Brüel & Kjaer NEXUS 2690), avant d'être numérisé sur un DAT (PORTADAT PDR1000) à une fréquence d'échantillonnage de 44.1kHz sur 16 bits. Il est important de disposer de très bonnes conditions techniques d'enregistrement et de restitution pour réaliser une étude sur la qualité sonore.

L'expérience a porté sur les premières mesures de l'*Ave Maria* de Gounod, chantées d'abord en émission normale, puis selon cinq différentes qualités vocales (soit 18 extraits au total).

Cette phrase musicale courte permet d'analyser 3 voyelles différentes (/a/, /e/ et /i/) et des consonnes entre chaque note. L'ambitus restreint de la phrase (une quarte) permet de balayer une plage de fréquences fondamentales semblables (entre 200 et 300 Hz). Afin de préserver l'homogénéité des stimuli, seules des voix d'hommes ont été étudiées. En outre, elles offrent la possibilité de distinguer lors de l'analyse acoustique le fondamental du premier formant des voyelles /a/ et /e/.

2.2. Résultats de l'étude cognitive

Une analyse linguistique préalable du discours libre des chanteurs a permis de dégager des termes importants et pertinents pour caractériser la qualité d'une voix lyrique afin de construire les échelles du test psychoacoustique d'évaluation de voix, présentées sur la figure 1 [4].

Les commentaires des experts sur chaque voix du corpus nous ont fourni des pistes d'analyse acoustique du signal. En particulier, l'étude linguistique a montré que les caractères "clair" et "sombre" semblaient être reliés à la perception "d'antériorisation" et de "postériorisation" des voyelles. Le caractère "timbré" ou la "brillance" renvoient plutôt à des notions de balance spectrale ou de renforcement de certaines bandes de fréquences.

sombre	←	→	clair
détimbré	←	→	timbré
sourd	←	→	brillant
pas nasillard	←	→	nasillard
sans air sur la voix	←	→	avec de l'air sur la voix
sans vibrato	←	→	avec beaucoup de vibrato
sans trémolo	←	→	avec beaucoup de trémolo

FIG 1 : Différentiateurs sémantiques sélectionnés pour le test d'évaluation de voix lyriques.

2.3. Déroulement du test d'évaluation

Le test d'évaluation a été conduit avec 11 professeurs de chant. Il consiste en l'écoute et

l'évaluation des 5 émissions modifiées de chacun des 3 chanteurs, sur des échelles à 5 points constituées par les termes issus de l'étude cognitive (figure 1).

2.4. Méthode d'analyse

La plupart des études sur la qualité vocale cherchent à établir la combinaison de paramètres physiques à laquelle renvoie une qualité vocale. Or, de nombreux paramètres physiques peuvent varier entre deux extraits, sans que notre oreille soit toujours capable de percevoir ces variations ni que celles-ci contribuent toutes à la perception d'une qualité vocale en particulier. Notre objectif consiste à rechercher les paramètres les plus saillants dans la perception d'une qualité vocale, tout en admettant qu'il peut en exister d'autres, moins perceptibles.

Le traitement statistique des résultats du test d'évaluation a montré que les échelles concernant la "nasalité", le "trémolo", le "chevrotement" et "l'air sur la voix" donnaient des résultats peu consensuels. Au contraire, la concordance des évaluations sur les caractères "timbré/détimbré", "brillant/sourd", "clair/sombre" et sur la "présence de vibrato" a justifié la recherche de corrélats acoustiques pour ces critères.

Sur la même base, nous avons repéré les voix les plus représentatives d'un critère donné, celles montrant une opposition maximale sur une échelle, ou encore celles présentant de fortes similarités.

Leur écoute en particulier et les commentaires libres des professeurs de chant à propos de ces extraits nous ont guidé dans la recherche des critères acoustiques les distinguant perceptivement. L'importance dans l'étude de la voix du formant du chanteur [6], de la balance spectrale pour la couleur sonore [9] et des formants vocaliques pour la perception des voyelles a motivé l'utilisation d'outils d'analyse et de représentation appropriés.

3. RESULTATS

3.1. Caractère "brillant/sourd"

Le discours des experts indiquant l'existence d'un lien entre la brillance, l'efficacité, la portée du son et le formant du chanteur, l'analyse s'est portée sur la différence d'amplitude entre le maximum des harmoniques situés dans la sone des deux premiers formants (0-2000 Hz), et l'amplitude spectrale dans la zone du formant du chanteur (correspondant à la zone sensible de l'oreille, entre 2000 et 4000 Hz).

Entre la voix jugée la plus sourde et celle jugée la plus brillante, on constate globalement sur toute la phrase une grande différence quant à l'émergence du formant du chanteur par rapport aux premiers formants. (figure 2)

La figure 3 étend cette analyse à l'ensemble des 15 voix lyriques étudiées. La différence d'amplitude entre les maxima des deux zones spectrales [0-2000 Hz] et [2000-4000Hz] semble bien corrélée à la notion de brillance pour T2 et B6, moins pour B3.

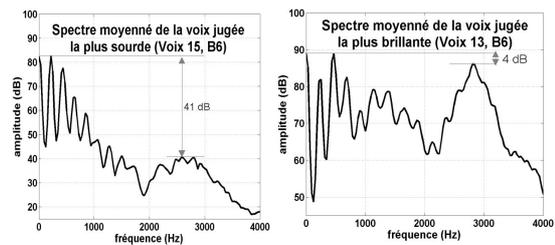
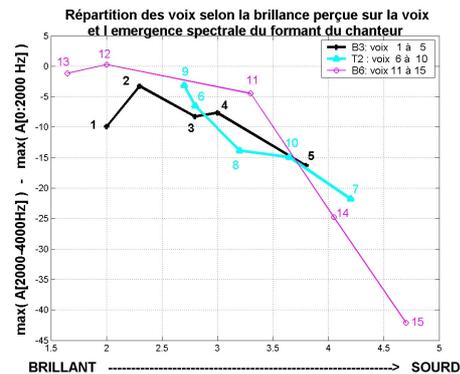


FIG 2 : Comparaison des spectres moyennés sur toute la phrase musicale des voix jugées la plus sourde et la plus brillante.



Chanteur	Amplitude des variations sur l'émergence du formant du chanteur	Jugement (de 1 à 5) du caractère Brillant/Sourd
B3	15 dB	2 à 3.8
T2	20 dB	2.75 à 4.25
B6	42 dB	1.6 à 4.75

FIG 3 : Répartition des voix selon la brillance perçue et l'émergence spectrale du formant du chanteur.

B6 produit les plus grandes variations sur l'émergence du formant du chanteur, à la différence de B3 qui a vraisemblablement exploré d'autres aspects de la qualité vocale lors de ses cinq tentatives. Ces résultats peuvent aussi indiquer l'existence d'un autre paramètre physique responsable de la perception de brillance, sur lequel jouerait plutôt B3.

3.2. Caractère "timbré / détimbré"

La notion de timbrage, assez proche de celle de brillance dans le discours des experts, semble liée au renforcement du formant du chanteur mais également à la richesse/pauvreté du spectre entier. Les analyses menées sur ce caractère ont par conséquent porté sur le formant du chanteur mais aussi sur l'ensemble du spectre.

La figure 4 montre, pour chaque chanteur, que le caractère "détimbré" s'accompagne d'une forte atténuation spectrale dans la région [2000-4000 Hz], de la même façon que pour le caractère "sourd". On remarque également une autre atténuation spectrale au-delà de 4000 Hz pour les voix "détimbrées", particulièrement chez le chanteur B6.

Cependant, ces atténuations peuvent être accentuées par la différence de niveau sonore global entre les émissions "timbrées" et « détimbrées ». Une représentation spectrale en pourcentages de l'énergie globale permet d'apprécier les variations de balance spectrale entre

le grave et l'aigu, indépendamment du niveau sonore. Par ailleurs, une échelle de fréquences de type logarithmique rend mieux compte de la discrimination spectrale humaine [5]

Ainsi, la figure 5 montre que les voix "détimbrées" présentent une atténuation au delà de 4000 Hz (deux dernières bandes de fréquences), mais surtout un renforcement énergétique important dans le fondamental (196 à 261 Hz), correspondant aux deux premières bandes de fréquences. Cela pourrait expliquer la perception "d'aggravation du spectre" mentionnée par les experts à l'écoute des voix jugées "détimbrées". La troisième bande de fréquence, rendant compte des variations du 1er formant vocalique, est renforcée pour les voix "détimbrées" des deux basse-barytons, pas pour le ténor. Enfin, on observe toujours un renforcement du formant du chanteur (6ème bande de fréquence) pour les voix "timbrées". Néanmoins, on constate que ce renforcement chez T2 est davantage lié au niveau sonore qu'à la balance spectrale.

3.3. Caractère "clair"/ "sombre"

Les commentaires des experts ainsi que de nombreux traités de chant [10] présentent les caractères "clair" et "sombre" comme des qualités vocales incontournables du chant lyrique, liées à une variation de prononciation des voyelles. La figure 6 représente l'analyse comparée des spectres moyennés de trois voyelles pour les deux voix jugées les plus distantes sur l'échelle clair/sombre, mais produites par le même chanteur T2 (figure 6), ainsi que les zones de fréquence des deux premiers formants d'un chanteur français masculin [8]. Pour les deux /a/, entre les voix claire et sombre, le maximum d'amplitude dans la zone F1 passe de H4 à H3, et dans la zone F2 de H6 à H5. Une variation similaire est visible pour le /i/ dans la zone de F2. Ces variations, que l'on peut interpréter comme un abaissement des formants, sont semblables à celles qu'on observe lors du passage d'une voyelle ouverte à une voyelle couverte [7]. On observe également un enrichissement de l'aigu du spectre de la voix claire, en particulier dans zone du formant du chanteur.

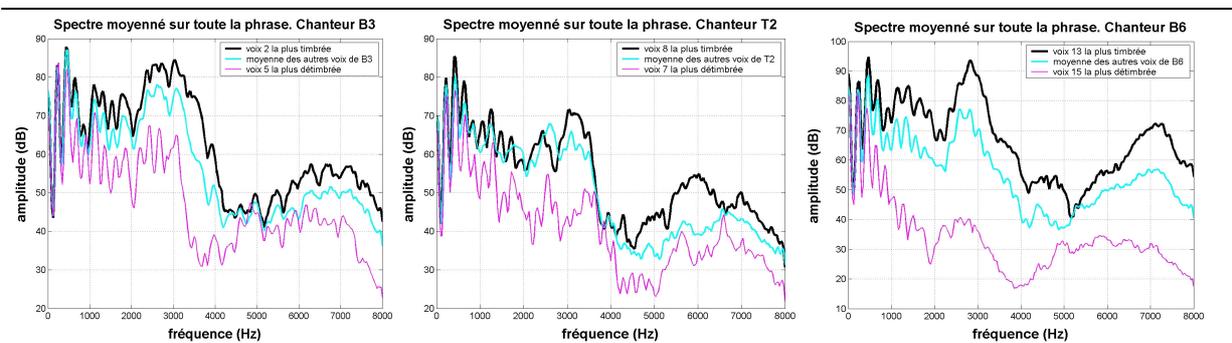


FIG 4 : Spectres moyennés sur l'ensemble de la phrase musicale, pour chaque chanteur, de sa voix jugée la plus timbrée, de la plus détibrée, et de la moyenne de ses autres voix.

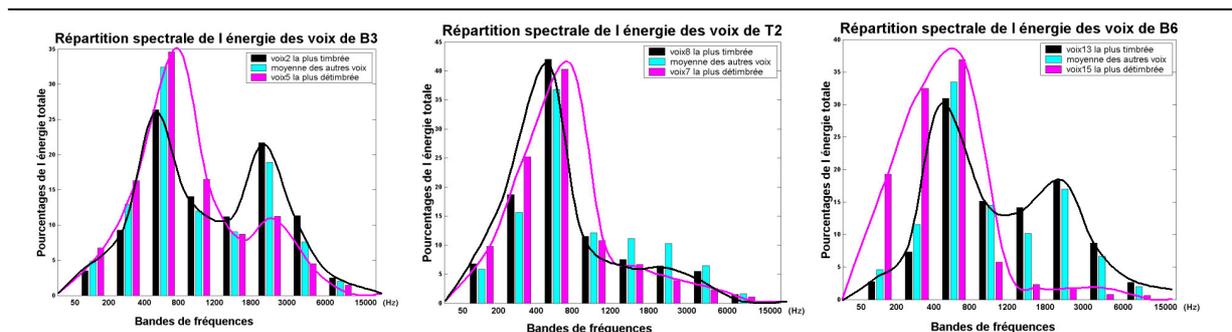


FIG 5 : Répartition par bandes de fréquences, pour chaque chanteur, de l'énergie de sa voix jugée la plus timbrée, de la plus détibrée, et de la moyenne de ses autres voix.

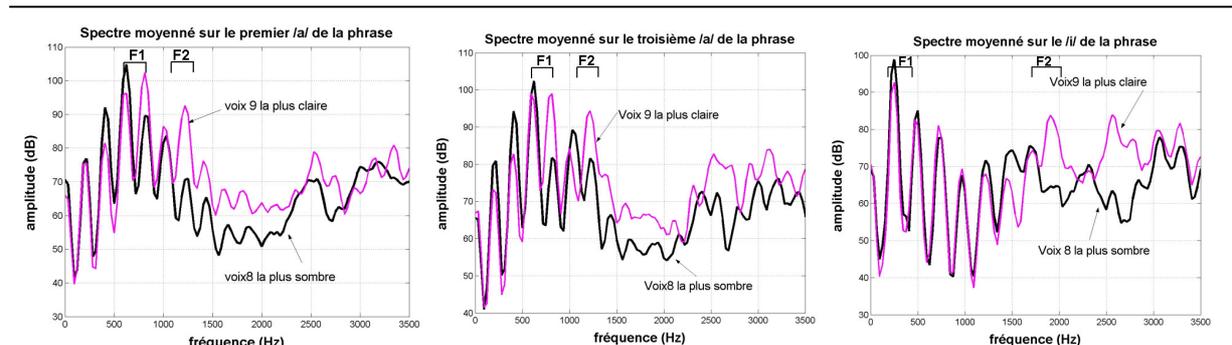


FIG 6 : Comparaison des spectres moyennés sur 3 voyelles de la phrase, des voix jugées la plus claire et la plus sombre.

3.4. Corrélations entre critères

Les résultats du test d'évaluation, présentés figure 7, et l'étude de la verbalisation libre des experts ont fait apparaître de fortes corrélations entre l'air, la brillance, le timbre et le vibrato. Une voix "timbrée" est souvent également qualifiée de "brillante", "avec du vibrato" tandis qu'une voix "détimbrée" est associée au caractère "sourd" et à la présence d'"air sur la voix".

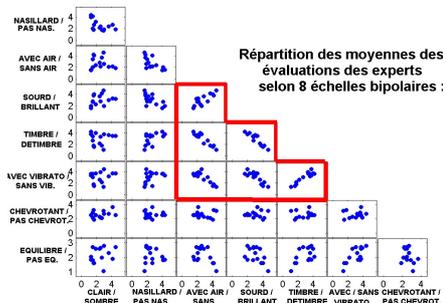


FIG 7: Diagramme des corrélations entre les 8 critères d'évaluation. Les fortes corrélations sont entourées.

Le formant du chanteur semble jouer un rôle important dans la perception de ces différentes qualités corrélées, mais non synonymes. Le renforcement spectral du formant du chanteur par rapport à la zone des deux premiers formants vocaliques correspond à la fois à la qualité de brillance et au timbre d'une voix. A l'inverse, le caractère "sourd", la présence "d'air sur la voix" comme le "détimbre" sont caractérisés par une accentuation spectrale de la zone des premiers formants de la voix au détriment de la partie haute du spectre.

4. DISCUSSION ET CONCLUSION

4.1. Aspect multiparamétrique

La complexité de la recherche de corrélats entre des critères verbaux et acoustiques réside dans le caractère multiparamétrique de ces deux espaces. Nous avons observé qu'une qualité vocale était liée cognitivement et perceptivement à d'autres qualités vocales par des relations de synonymie, d'implication, de partage d'un certain nombre de propriétés sémantiques. Toutes ne sont cependant pas corrélées et s'il est possible de distinguer le caractère le plus saillant d'une voix, une seule qualité vocale ne suffit pas à la décrire entièrement. De même, du point de vue physique la perception d'une qualité vocale ne peut être imputée à un paramètre unique mais à la contribution pondérée de différents paramètres.

4.2. Discussion sur les termes des échelles

Bien que le choix des échelles soit justifié par l'étude préalable de la signification individuelle des termes, certains qualificatifs deviennent pourtant mal partagés par les experts lorsqu'ils sont associés comme pôle sur la même échelle (antonymie non exacte, problème de termes comme « équilibré » qui pousse à un jugement moyen).

Un point reste problématique: un même terme peut qualifier le timbre de l'émission « normale » d'un chanteur, par comparaison à une référence en mémoire ou à un autre chanteur, mais peut également qualifier les modifications de qualité vocale d'un même chanteur. Dans les deux cas, ce qualificatif ne correspond pourtant pas à la même "intensité" de perception, entraînant une ambiguïté dans l'interprétation des résultats de l'évaluation d'une voix suivant un critère, si l'on mélange les différents chanteurs.

4.3. Discussion sur les différences interindividuelles entre les chanteurs

Ces différences soulignent également que plusieurs paramètres sont ajustables par le chanteur pour jouer sur un critère donné. Il est donc plus intéressant de s'intéresser aux variations de qualité vocale de chaque chanteur plutôt que de vouloir agréger les différentes stratégies de tous les chanteurs confondus.

Dans la présente étude, les modifications de qualité vocale étaient effectuées selon l'inspiration du chanteur. Ils ne produisaient donc pas tous les mêmes qualités vocales, et ne les qualifiaient pas par les mêmes termes que les professeurs de chant à la réécoute, ou ceux constituant les échelles du test psychoacoustique. Il nous semble donc nécessaire de poursuivre par une autre expérience où il serait demandé aux chanteurs de produire des qualités vocales « imposées », tout en leur permettant de réécouter leur production.

Des paramètres spectraux ont été mis en avant dans cette étude, résultant de l'écoute et de l'analyse du discours des professeurs de chant. Il est maintenant nécessaire de tester leur validité perceptive relativement aux qualités vocales explorées, par la synthèse d'exemples sonores appropriés.

BIBLIOGRAPHIE

- [1] T. Ehrette, N. Chateau and V. Maffiolo. Predicting the Perceptive Judgment of Voices in a Telecom Context. Eurospeech, 2003.
- [2] D.H Klatt, L.C. Klatt. Analysis, synthesis and perception of voice quality variations among female and male talkers. JASA, 1990
- [3] N. Henrich. Etude de la source glottique en voix parlée et chantée. Thèse de doctorat, 2001
- [4] M. Garnier, D. Dubois, J. Poitevineau, N. Henrich, M. Castellengo. Perception et description verbale de la qualité vocale dans le chant lyrique: une approche cognitive. JEP 2004
- [5] E. Zwicker, H. Fastl. Psychoacoustics, facts and models. Sger verlag, 1999
- [6] J. Sunberg. Level and ceter frequency of the singer's formant. Journal of voice, 2001.
- [7] B. Chuberre. Les registres et passages dans la voix chantée. Thèse de doctorat, 2000.
- [8] X. Rodet. Conférences des journées d'étude du festival du son, Paris, 1980.
- [9] G. von Bismark. Timbre of steady sound : a factorial investigation of its verbal attributes, Acustica, 1974
- [10] M.Garcia. L'art du chant. Paris, 1840.