

Approche de la qualité vocale dans le chant lyrique : Perception, verbalisation et corrélats acoustiques

**Mémoire de stage de DEA
Mars - Juillet 2003**

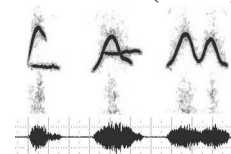


Etudiante : Maëva Garnier

Encadrement : Nathalie Henrich,
Michèle Castellengo, Danièle Dubois,
Jacques poitevineau.

DEA ATIAM (Paris 6, ENST, Aix
Marseille II, UJF Grenoble I)

Laboratoire d'accueil : Laboratoire
d'Acoustique Musicale (LAM)



Sommaire

Table des figures	4
Remerciements	6
Introduction	7
Contexte de l'étude et Bibliographie	10
1) Quelques rappels sur la voix	11
1.1) Principe de production de la voix	11
1.2) Méthodes d'exploration de la source glottique	12
1.3) La voix chantée	13
2) La Qualité vocale : problématiques	13
2.1) Contrôle de la qualité vocale	13
2.2) Ecoute et catégorisation	14
2.3) Problématique : la qualité vocale et le timbre d'une voix	15
3) Une étude fortement pluridisciplinaire	15
3.1) Perception	16
3.2) Verbalisation et linguistique	16
3.3) Relation entre l'acoustique de la voix et la qualité vocale	18
Analyse de la verbalisation des chanteurs	21
1) Objectifs	22
2) Matériel et méthodes	22
2.1) Enregistrements et corpus des chanteurs : Provenance, contexte d'enregistrement et consigne donnée aux chanteurs.	22
2.2) Analyse sémantique et linguistique	23
3) Résultats	24
3.1) Liste des qualificatifs utilisés et consensus sur la qualité vocale	24
3.2) La perception de la voix comme expérience individuelle ou collective	26
3.3) Positionnement du sujet par rapport à la voix	26
3.4) Emergence de plusieurs objets autour du mot « voix ».	28
3.5) Premières relations entre les différents termes utilisés	28
3.6) Enoncés métalinguistique et modalités	32
4) Conclusion : vers les experts	32
Perception de la qualité vocale et analyse du discours des experts	34
1) Description du test d'écoute	35
1.1) Sélection et diffusion des extraits sonores :	35
1.2) Plan d'entretien	36
1.2.1) Accueil du sujet au laboratoire	36
1.2.2) Explications du contexte du test d'écoute au sujet	36
1.2.3) Entretien ouvert dirigé sur l'écoute des voix	36
1.2.4) Evaluation d'une quinzaine de voix par rapport à des échelles bipolaires	37
1.2.5) Critiques et améliorations du test	37
2) Réalisation	37
2.1) Recrutement des sujets	37
2.1.1) Critères	37
2.1.2) Moyens de communication	37

2.1.3) <i>Problèmes liés à leur intérêt pour la physiologie et l'acoustique</i>	38
2.2) Pré-test	38
2.2.1) <i>Problème de durée.</i>	38
2.2.2) <i>Mise au point de nouvelles échelles.</i>	38
2.3) Tests	39
2.3.1) <i>Conditions</i>	39
2.3.2) <i>Premières remarques générales</i>	39
2.3.3) <i>Problèmes du test inchangeables à ce stade de l'étude.</i>	40
3) Résultats descriptifs	41
3.1) Dépouillement des échelles et pistes pour l'analyse acoustique	41
3.2) Analyse des termes des échelles dans le corpus des experts	46
3.2.1) <i>Analyse du souffle</i>	47
3.2.2) <i>Analyse du vibrato</i>	48
3.2.3) <i>Analyse des caractères clair et sombre</i>	48
3.2.4) <i>Analyse de la brillance / surdité</i>	49
3.2.5) <i>Analyse du caractère nasal / nasillard</i>	50
3.2.8) <i>Analyse de l'équilibre</i>	50
3.4) Etude des différents objets entrant en jeu pour parler de la qualité vocale	50
<i>Voici page suivante un schéma récapitulatif de la « hiérarchisation » de ces objet dans le discours des experts</i>	51
3.5) Etude de la modalisation dans le discours	53
Recherche de corrélats acoustiques	58
1) Rappels des pistes d'analyse mises en évidence lors de l'étude de la verbalisation des experts.	59
2) Etude du vibrato.	59
3) Recherche de corrélation entre le caractère clair /sombre et la couverture des voyelles	61
2) Recherche de corrélation entre la brillance et l'intensité du formant du chanteur	62
4) Recherche de corrélats pour le timbrage	64
5) Corrélation entre l'air, la brillance et le timbrage.	66
Conclusion	70
Bibliographie	71
Ouvrages	71
Thèses et rapports de DEA	72
Articles	72

Table des figures

Figure 1 : Schéma illustratif des processus cognitifs de catégorisation des sons	9
Figure 2 : Représentation en coupe de l'appareil vocal dans son ensemble	11
Figure 3 : Principe de l'électroglottographie. D'après Henrich, 2001	12
Figure 4 : Schématisation de la représentation de la qualité vocale par rapport à une référence en mémoire du sujet.....	14
Figure 5 : Triade sémiotique d'Aristote : relation entre mots, objets et concepts	17
Figure 6 : Influence du débit glottique sur la qualité vocale.....	19
Figure 7 : schéma de définition du quotient ouvert	20
Figure 8 : répartitions des marques de la personne renvoyant au sujet ou à un collectif.....	26
Figure 9 : Occurrences des pronoms pour chaque chanteur. Classification des chanteurs en trois groupes plutôt objectif/plutôt subjectif/neutre.....	27
Figure 10 : Répartition des pronoms dans les corpus tessiture et sexe	27
Voici par exemple les résultats cumulés par pupitre :	27
Figure 11 : Répartition des chanteurs dans les catégories objectif / subjectif / neutre	28
Figure 12 : Association des termes dans le discours des chanteurs.....	31
Figure 13 : Nouvelles échelles mises au point pour les tests d'écoute	39
Figure 14 : Chanteur 1 (B3).....	42
Figure 15 : Chanteur 2 (T2)	43
Figure 16 : Chanteur 3 (B6).....	44
Figure 17 : Corrélations entre les différents termes des échelles. Dans le carré rouge, les corrélations fortes sont soulignées.	46
Figure 18 : schéma récapitulatif de la « hiérarchisation » des objets cognitifs dans le discours des experts	52
Figure 19 : Voix 1 du chanteur B3, perçue avec du vibrato (esthétique)	60
Figure 20 : Voix normale du chanteur T2, perçue avec un vibrato trop lent.	60
Figure 21 : Voix 4 du chanteur B6, perçue sans vibrato.....	60
Figure 22 : représentation des 1ers « A » des voix selon la différence des harmoniques 1,3,5.....	61
Figure 23 : représentation des 2èmes « A » des voix selon la différence des harmoniques 1,3,5	62
Figure 24 : représentation des « I » des voix selon la différence des harmoniques 1,3,5	Erreur ! Signet non défini.
Figure 25: Emergence spectrale de la zone du formant du chanteur (entre 2000 et 4000 Hz) par rapport à la zone des premiers formants (entre 0 et 2000 Hz), en fonction du caractère brillant ou sourd de la voix. Cas de la première occurrence de la voyelle [a] dans la phrase chantée.	63
Figure 26 : Emergence spectrale de la zone du formant du chanteur (entre 2000 et 4000 Hz) par rapport à la zone des premiers formants (entre 0 et 2000 Hz), en fonction du caractère brillant ou sourd de la voix. Cas de la seconde occurrence de la voyelle [a] dans la phrase chantée	63
Figure 27 : Emergence spectrale de la zone du formant du chanteur (entre 2000 et 4000 Hz) par rapport à la zone des premiers formants (entre 0 et 2000 Hz), en fonction du caractère brillant ou sourd de la voix. Cas de la première occurrence de la voyelle [i] dans la phrase chantée.....	64
Figure 28 : Spectre moyenné sur l'ensemble de la phrase musicale, dans le cas du chanteur B3. En trait épais plein : voix jugée timbrée (>2.5) , en trait épais pointillé, voix jugée ni timbrée, ni détimbrée (~2.5); en trait fin, voix jugée détimbrée (<2.5).....	65
Figure 29 : Spectre moyenné sur l'ensemble de la phrase musicale, dans le cas du chanteur T2. En trait épais : voix jugée timbrée; en trait fin, voix jugée détimbrée (<2.5).	65
Figure 30 : Spectre moyenné sur l'ensemble de la phrase musicale, dans le cas du chanteur T2. En trait épais : voix jugée timbrée; en trait fin, voix jugée détimbrée (<2.5).	65
Figure 31: Emergence spectrale de la zone du formant du chanteur (entre 2000 et 4000 Hz) par rapport à la zone des premiers formants (entre 0 et 2000 Hz), en fonction du caractère timbré ou détimbré de la voix. Cas de la première occurrence de la voyelle [a] dans la phrase chantée.	66

Figure 33: Emergence spectrale de la zone du formant du chanteur (entre 2000 et 4000 Hz) par rapport à la zone des premiers formants (entre 0 et 2000 Hz), en fonction du caractère timbré ou détimbré de la voix. Cas de la deuxième occurrence de la voyelle [a] dans la phrase chantée.....	67
Figure 34: Emergence spectrale de la zone du formant du chanteur (entre 2000 et 4000 Hz) par rapport à la zone des premiers formants (entre 0 et 2000 Hz), en fonction du caractère timbré ou détimbré de la voix. Cas de la voyelle [i] dans la phrase chantée.	67
Figure 35: Emergence spectrale de la zone du formant du chanteur (entre 2000 et 4000 Hz) par rapport à la zone des premiers formants (entre 0 et 2000 Hz), en fonction de la présence ou non d'air sur la voix. Cas de la voyelle [i] dans la phrase chantée.	68
Figure 36: Emergence spectrale de la zone du formant du chanteur (entre 2000 et 4000 Hz) par rapport à la zone des premiers formants (entre 0 et 2000 Hz), en fonction de la présence ou non d'air sur la voix . Cas de la voyelle [i] dans la phrase chantée.	68
Figure 37: Emergence spectrale de la zone du formant du chanteur (entre 2000 et 4000 Hz) par rapport à la zone des premiers formants (entre 0 et 2000 Hz), en fonction de la présence ou non d'air sur la voix. Cas de la voyelle [i] dans la phrase chantée.	69

Remerciements

Je tiens en premier lieu à remercier Nathalie Henrich, ma maître de stage pour sa gentillesse, son amitié, son dynamisme et son optimisme inébranlable et communicatif. Elle a su me transmettre son intérêt pour la voix, me guider dans mes « premiers pas » dans le domaine de la recherche. Par son exemple d'ancienne du DEA ATIAM, elle m'a également donné chaque jour la confirmation de l'adéquation du choix de ce DEA avec mes envies de continuer à travailler dans ce domaine, ce type de laboratoire, et parmi des personnes comme elle.

Je voudrais également remercier Michèle Castellengo pour sa sympathie et sa disponibilité. Elle a su me faire partager sa grande expérience, ses compétences multiples et ses éclats de rire tout au long de mon séjour au LAM.

Bien sûr, je remercie aussi tout particulièrement Danièle Dubois et Jacques Poitevineau qui m'ont introduite aux sciences humaines et convaincue de l'importance à accorder en priorité à la perception humaine plutôt qu'aux instruments de mesures. Ayant jusque là eu uniquement un parcours de physicienne, cela a été pour moi non pas une découverte mais en tout cas une grande leçon d'ouverture d'esprit et l'occasion de maintes réflexions sur les méthodologies scientifiques. Un grand merci également à Pascale Cheminée qui nous a amicalement aidé sur l'analyse des corpus des experts.

Il arrive parfois qu'on ne se rende pas compte de la chance qu'on peut avoir lorsqu'elle se présente. Pour ma part, j'ai été très heureuse et consciente d'avoir eu l'honneur de travailler avec des personnes d'exception et de référence dans ces différentes disciplines, qui ont pour autant su rester simples et très humaines.

Je voudrais aussi remercier Jean-Dominique Polack, directeur du LAM, ainsi que l'ensemble du personnel pour leur accueil chaleureux.

Une pensée également à l'attention de Alain Zürcher, Nicolas Kern et Bertand Chuberre, chanteurs, ainsi que Jean Nirouët, Michel Hart, Gaël de Kerret, Jacques Chuilon, Yaël Benzaquen, Lucia Di Carlo, Irène Bourdat, Robert Expert, Paolo Zedda et Cyrène Paparotti, chanteurs lyriques et professeurs, qui ont si gentiment accepté de collaborer bénévolement à mes recherches. A leur contact, j'ai appris énormément de choses concernant l'état de la recherche sur la voix, la physiologie de l'appareil vocal, le répertoire et les grands chanteurs lyriques, les techniques vocales...etc. J'espère que nous aurons l'opportunité de pouvoir encore collaborer sur de prochains projets dont nous avons déjà eu l'occasion de discuter.

Enfin, je ne saurais terminer sans remercier tous les stagiaires ayant effectué leur stage au LAM à la même période que moi pour leur enthousiasme, les très bons moments musicaux de chorale ou de jazz collectif, les discussions variées que nous avons pu avoir et qui ont contribué à la très bonne ambiance du laboratoire.

Introduction

La voix est un système de production sonore très particulier : à la fois outil de communication et d'interaction sociale, donc porteuse de sens, elle peut également caractériser une personne, ou être utilisée comme instrument de musique à part entière, dans le cadre de la voix chantée.

L'apprentissage de la parole étant déjà un lointain souvenir pour chacun de nous, nous nous sommes habitués à la perception de notre propre voix et nous prêtons désormais peu d'attention à la façon dont nous parlons. Notre principal souci étant de nous faire comprendre nous assurons l'intelligibilité de notre voix à travers le contrôle de l'intensité sonore, de l'intonation ou de l'articulation.

Pourtant, le message ou les émotions que nous transmettons à un interlocuteur sont loin d'être véhiculés seulement par les mots que nous pouvons prononcer, mais aussi, et parfois bien plus, par la qualité de la voix qui en est le support. C'est le cas par exemple de certaines expressions quotidiennes machinales qui n'apportent de sens que par la façon dont on les exprime. Il en est de même de l'ensemble des qualités d'une voix qui nous permet de discriminer des personnes, voire de les reconnaître.

Mais qu'est ce que la qualité vocale ? Peut-on considérer qu'il s'agit de tout ce qui diffère du contenu sémantique de la production vocale, ou bien ce terme fait-il plus particulièrement appel aux qualités du son, dans le sens où on l'entend pour le timbre d'un instrument de musique ?

Pour l'instant, c'est sur la première conception que se sont basées la plupart des recherches sur la reconnaissance du locuteur, ou les études sur l'émotion véhiculée par la voix. La prosodie, l'intonation, l'articulation, la hauteur, le rythme, l'intensité et le « timbre » de l'émission vocale seraient des indices de la qualité vocale. On voit pourtant que chacun de ces indices est lui-même complexe.

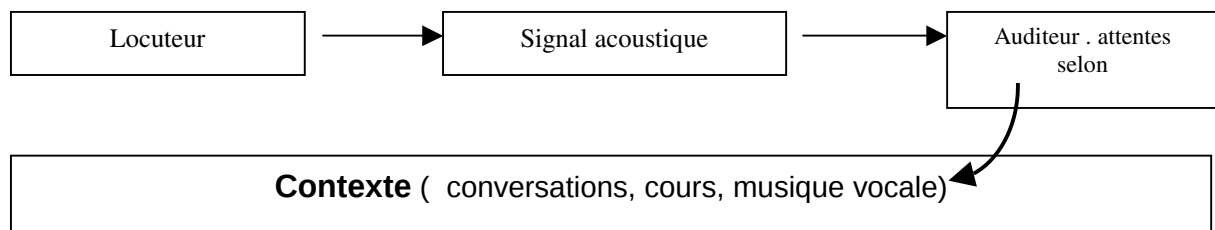
C'est le cas de la notion de timbre auquel la littérature donne plusieurs significations : est-il caractérisé par les attaques et les fins de sons ? Ou bien le timbre est-il ce qui distingue deux sons de même intensité et même hauteur ? Dans ce cas, peut-on considérer pour le cas particulier de la voix que les voyelles font alors partie du timbre, donc de la qualité vocale ? De plus, dans la plupart des études du timbre instrumental, les sons sont étudiés sur une seule note. Peut-on réduire la voix à un son, décorréler la voix de son essence communicative, de son phrasé, de sa musicalité dans le cas de la voix chantée ? En même temps, on perçoit qu'au cours d'une phrase, la qualité vocale semble changer sur chaque syllabe. Se pose alors le problème de localité ou de globalité du timbre. Quand on parle de qualité vocale, parle-t-on du timbre en général ou d'un timbre mot par mot ? Est-il même possible de définir un timbre global ? Si tel est le cas, cela nous ramène au problème de départ : quel sens à donner à la qualité vocale, si l'articulation, la diction, la prosodie qui font partie de la perception globale de la voix font partie du timbre ?

Les différents professionnels de la voix (chanteurs, phoniâtres, acousticiens, orateurs, orthophonistes,...) ont bien saisi à la fois, l'importance et la complexité de la notion de « qualité vocale », et entreprennent des recherches pour la définir et trouver des descripteurs. Dans le cas des voix de serveurs téléphoniques, la qualité vocale renvoie à l'identification du genre, de l'âge ou du comportement supposé d'un locuteur (timide, prétentieux...), à des modifications de timbre du son émis. En revanche, pour des voix pathologiques, ces qualificatifs sont plus liés à la physiologie ou à des fonctionnements anormaux des cordes vocales (aphone, rugueuse, diplophonique...). Dans le domaine du chant, la qualité vocale est souvent décrite de façon très métaphorique (blanc, soufflé, plat,...) ou par un « jargon » lyrique bien particulier, qui de surcroît varie légèrement d'un professeur à l'autre. Elle peut également aller jusqu'à comprendre des notions de style musical.

Il apparaît donc que la perception de la qualité de la voix dépend de la situation de communication – communication téléphonique, musique...etc - qui conditionne les attentes que nous pouvons en avoir. En effet, la perception n'est pas une réception passive de stimulations extérieures, mais plutôt une recherche active d'indices dans la multitude des informations qui nous parviennent du monde. La sélection de ces indices est motivée par nos attentes en fonction du contexte ou de la probabilité d'une situation et vise à donner du sens à nos perceptions.

La question qui se pose alors n'est peut être pas tant de savoir *comment* nous percevons la voix, mais plutôt de se demander *qu'est ce que* l'on cherche à écouter précisément dans la voix, quels sont ces indices que nous sélectionnons auditivement pour caractériser et catégoriser une voix ?

D'ailleurs, la voix faisant fatalement intervenir le contexte d'interlocution, il serait intéressant de savoir si l'auditeur et le producteur perçoivent et caractérisent la voix de façon semblable ou différente. En effet, d'un côté, le locuteur semble avoir des sensations exacerbées, du fait que la voix vient de l'intérieur de lui même. Mais de l'autre, on peut aussi se demander s'il est possible pour l'interlocuteur d'arriver à écouter une voix comme il le ferait d'un son instrumental, c'est-à-dire décorrélée de l'instrumentiste. En effet, il est rare qu'on parle du clarinettiste lorsqu'on cherche, par exemple, à qualifier un timbre de clarinette. Mais en est il de même pour la voix, très liée à l'humain et à l'identité de la personne ? Ainsi, il serait probablement très intéressant d'étudier l'impact de l'humain et du corps sur la perception de la voix et pour cela d'adopter une approche méthodologique multisensorielle.



En bref, si toute l'information est contenue dans le signal sonore, c'est en partant de l'avis des auditeurs que nous pourrions accéder à celle qui est pertinente du point de vue perceptif.

Notre approche est bien donc bien différente de celle de la psychophysique : elle consiste, non plus à évaluer le jugement subjectif de paramètres physiques isolés, mais de partir de l'identification des caractéristiques sémantiques exprimées dans le discours des experts pour les corrélérer à des paramètres physiques. En effet, la démarche analytique de la psychophysique ne permet pas de repérer les structures globales des propriétés physiques dont la conjonction et l'intégration dans un « tout » constitue précisément ce qui fait sens et permet l'identification des différentes catégories de timbres ou de qualité sonore (vocale). En contraste la démarche qui consiste à partir des représentations sémantiques telles qu'elles se manifestent dans les processus langagiers (en particuliers lexicaux) permet le repérage du caractère holistique des catégories de timbre, et de leur structure interne. On fait en effet l'hypothèse que ces catégories sémantiques sont d'avantage structurées selon un air de famille autour de représentations typiques à différents niveaux hiérarchiques que selon une structure logique régulière fût-elle multidimensionnelle.

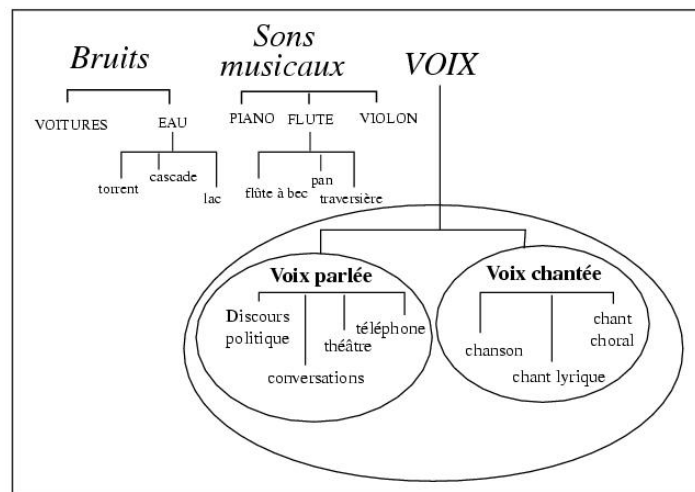


Figure 1 : Schéma illustratif des processus cognitifs de catégorisation des sons

Le choix de la voix chantée lyrique répond à une double exigence. D'une part, la pratique de ce type de chant consiste à cultiver prioritairement le contrôle de la qualité vocale mise au service de l'expression musicale. D'autre part, les chanteurs acquièrent la maîtrise de la production tout en développant une écoute spécialisée des variations de qualités vocales, tant pour leur propre voix que pour celle de leur partenaire. Enfin, l'instrument vocal se dérochant à la vue, il existe un important vocabulaire spécifique pour décrire les « gestes techniques » à accomplir et les qualités des sons à produire.

Le point de départ de notre étude se situe donc en sciences humaines. Bénéficiant d'un corpus d'exemples sonores enregistrés par des chanteurs lyriques confirmés lors de la thèse de N. Henrich, nous commencerons par analyser les termes employés par les chanteurs lors de la production des variations de qualité vocale, et ceux que des auditeurs professionnels de la voix utilisent pour décrire les exemples enregistrés que nous leur donnons à entendre.

Cette première partie du travail implique différentes opérations : la sélection des exemples enregistrés ; l'étude des verbalisations des chanteurs ; la mise en place d'un protocole d'entretien ouvert avec des experts acceptant de se prêter au test d'écoute ; enfin, la transcription et l'analyse des verbalisations qui ont permis d'aboutir à la mise en évidence des différents objets cognitifs entrant en jeu dans la perception de la qualité vocale de voix lyriques et à la représentation que peuvent s'en faire les chanteurs et les auditeurs.

Nous avons ensuite identifié les critères acoustiques qui rendaient compte des variations de qualité vocale perçues par les auditeurs. Nous verrons que ceux-ci renvoient à des processus d'écoute rendant compte des différentes attentes des auditeurs. Ceux-ci ont trait soit à l'écoute de la répartition spectrale (émergence du formant du chanteur, équilibre spectral global, souffle), soit à une écoute spécifique de la production des voyelles, ou encore à une écoute musicale globale (vibrato, attaques, articulations).

Un tel travail n'aurait pu être mené sans l'implication forte d'une équipe de linguistes (Danièle Dubois et ses collaborateurs) s'associant aux acousticiens du LAM. La transcription exacte et l'analyse des verbalisations des 6 experts a pris un temps considérable, mais nous mesurons aujourd'hui l'importance de ce travail qui conserve le contexte d'utilisation des termes employés et à fait apparaître certaines notions que nous n'attendions pas.

Les recherches sur la qualité vocale intéressent de plus en plus les professionnels de la voix, notamment les professeurs de chant ou les orthophonistes, qui cherchent à comprendre le fonctionnement de la voix et à aider leurs élèves ou patients à apprendre à contrôler leur qualité vocale. Elles font également écho à la demande croissante des industriels dans les domaines de la synthèse vocale (afin de tendre vers plus de naturel et d'individualité de la voix synthétisée), des serveurs vocaux ou de la reconnaissance d'individu par leur empreinte vocale.

Nous espérons, par la méthode de recherche mise en œuvre et les premiers résultats obtenus, apporter une contribution nouvelle aux études de qualité sonore.

Contexte de l'étude et Bibliographie

1) Quelques rappels sur la voix

1.1) Principe de production de la voix

Avant de commencer à parler de la qualité vocale, il me semble nécessaire de faire quelques rappels sur la physiologie de la voix et les principes de base de l'émission vocale. Voici un schéma en coupe de l'appareil vocal afin de se repérer visuellement au cours de l'explication.

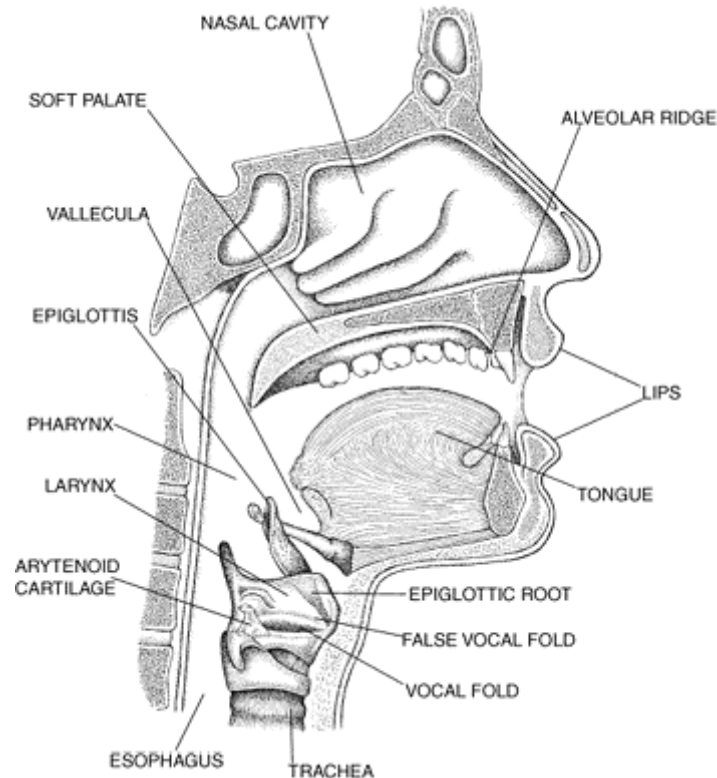


Figure 2 : Représentation en coupe de l'appareil vocal dans son ensemble

Tout d'abord, la production vocale peut se décomposer en 3 phases :

- La production du souffle
- La mise en vibration des cordes vocales par modulation de ce souffle et la génération d'une onde acoustique
- Le filtrage, la propagation et le rayonnement de l'onde acoustique

Les poumons et la trachée jouent le rôle de moteur dans l'émission vocale en agissant comme un soufflet qui produit et contrôle la pression d'air nécessaire à la mise en vibration des cordes vocales.

Le souffle est alors modulé au niveau du larynx par l'ouverture et la fermeture périodique des cordes vocales et le contact entre les cordes vocales génère une onde acoustique. La hauteur du son émis, de 90 à 120 Hz chez un homme en voix parlée, et de 150 à 300 Hz chez une femme, est liée à la fréquence du cycle d'ouverture et de fermeture des cordes vocales. Lorsque la glotte est fermée, c'est-à-dire lorsque les deux cordes vocales sont accolées, la pression sous-glottique, qui est la pression de l'air en provenance des poumons, augmente jusqu'à être supérieure à la force d'accolement des cordes vocales. La glotte s'ouvre alors sous cette pression, laissant passer de l'air. La pression sous-glottique chute alors, provoquant de nouveau l'accolement des cordes vocales. Ce cycle d'ouverture - fermeture glottique se répète de façon périodique.

Au sortir du larynx, le son glottique est théoriquement indépendant de la prononciation et des voyelles. Il n'est donc pas le son que nous qualifions usuellement voix, en tant qu'auditeur.

Ce son final est obtenu à partir du son glottique par filtrage à travers toute une série de résonateurs : conduit vocal, pharynx, cavité buccale modifiée par la langue, les lèvres et le voile du palais, cavité nasale, et l'œsophage, souvent oublié. Ces résonateurs privilégient certaines parties du spectre, notamment les zones formantiques caractéristiques des différentes voyelles contrôlées par la forme de la cavité buccale.

Chez les orateurs et les chanteurs, une autre partie du spectre est renforcée : il s'agit de ce qu'on appelle aussi le « formant du chanteur », situé vers les 2500 Hz pour les hommes, et qui est responsable de la portée et de la brillance d'une voix.

1.2) Méthodes d'exploration de la source glottique

Ces variations temporelles d'ouverture de la glotte peuvent être mesurées et visualisées grâce à plusieurs méthodes, de façon directe par la laryngoscopie stroboscopique, la cinématographie à vitesse ultra-rapide ou la vidéokymographie et de façon indirecte et totalement non-invasive par la mesure de débit et de pression en sortie des lèvres ou par l'électroglottographie.

C'est cette dernière méthode dont nous avons exploité les mesures au cours de cette étude. En simplifiant, cela fonctionne à l'aide de deux électrodes placées sur le cou du sujet, au niveau du cartilage thyroïdien auxquelles on applique un courant alternatif de haute fréquence (environ 1 MHz). Le principe (cf Figure II) est semblable à celui d'un ohmmètre et consiste à mesurer la différence de potentiel entre les deux électrodes. Celle-ci dépend de la résistance au passage du courant à travers le cou (peau, cartilage, muscles, glotte). De par la mauvaise conductivité de l'air par rapport aux tissus corporels, cette résistance est plus grande quand la glotte est ouverte que fermée. Lorsque les cordes vocales s'accrochent bien à la fermeture, on obtient alors un signal de même période que le son émis. A noter que ce signal, bien que périodique, est loin d'être sinusoïdal. Il est donc également très riche spectralement et porteur d'information concernant la qualité vocale.

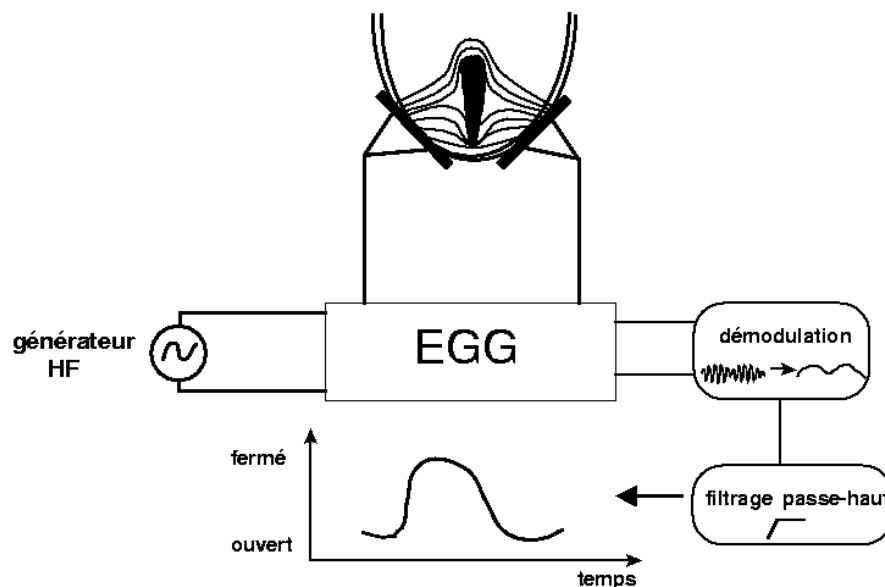


Figure 3 : Principe de l'électroglottographie. D'après Henrich, 2001.

1.3) La voix chantée

Pour bien comprendre certaines verbalisations données par les chanteurs ou les experts de la voix lyrique, il faut préciser certaines notions et certaines contraintes esthétiques de la voix chantée classique.

Tout d'abord, la ligne de conduite du chant classique se base sur le respect de la santé vocale et sur l'efficacité de l'émission : Une bonne émission est une production où la voix n'est pas forcée, où le corps ne subit pas de tension et où toute l'énergie dépensée est utilisée à produire le son et à le porter. Cela passe principalement par une bonne gestion et connexion du souffle, et par une bonne respiration ventrale.

Pour l'esthétiques de timbre de la voix lyrique, l'école italienne propose des maximes d'équilibre de la qualité vocale : Le « Chiaro oscuro », équilibre entre rondeur et brillance par le dosage approprié des graves et des aigus du spectre, mais aussi l' « Aperto/coperto », équilibre entre l'ouverture et la couverture. La couverture du son revient à modifier la phonation des voyelles, en particulier postérieures. Cette techniques est utilisée dans l'aigu d'une tessiture où des sons ouverts seraient dangereux pour le larynx. Mais dans le médium de la tessiture, il est préconisé esthétiquement de trouver un équilibre entre l'ouverture et la couverture, et de garder une constance de timbre sur l'émission des voyelles d'un bout à l'autre d'une phrase musicale..

Pour finir, deux autres caractéristiques principales de la voix lyriques subissent des contraintes esthétiques : l'attaque d'un son, qui doit se faire nettement et être pensée, ainsi que le vibrato qui est apprécié pour sa constance, sa régularité et son naturel.

2) La Qualité vocale : problématiques

2.1) Contrôle de la qualité vocale

Lorsque nous avons appris à parler, nous avons dû écouter et essayer de reproduire la façon dont nous parlaient les adultes, non seulement au niveau articulatoire, mais également au niveau du son produit lui-même. En effet, nous avons dû apprendre à gérer le niveau sonore de nos émissions en fonction de différents contextes (voix murmurée, voix normale, voix forcée dans un environnement bruyant). Nous avons également appris à corréler le son de notre voix avec les émotions que nous voulons exprimer. Par exemple, nous ne parlons pas de manière douce lorsque nous nous mettons en colère, ou à l'inverse, nous ne parlons pas avec une voix criarde dans un contexte de séduction. Toutes ces qualités vocales sont pour la plupart culturelles et non pas innées, c'est pourquoi à un moment de notre apprentissage de la parole, nous avons eu à apprendre à les contrôler.

Désormais, tout cela est pourtant déjà un lointain souvenir pour chacun de nous, et nous ne faisons plus attention à la façon dont nous parlons qui nous paraît naturelle. De même, nous prêtons peu attention à notre propre timbre de voix qui fait partie intégrante de notre identité, tout comme la couleur de nos cheveux ou la taille que nous faisons.

En fait, notre principal souci quotidien dans l'émission de la parole réside plutôt dans la volonté d'être compris, que ce soit en soignant la clarté des idées que nous exprimons ou encore en adaptant l'intelligibilité de notre voix à l'environnement dans lequel nous parlons. Ces différents contrôles relèvent dans ce cas plus de l'intensité sonore, de l'intonation ou de l'articulation.

Pourtant, le message ou les émotions que nous transmettons à un interlocuteur sont loin d'être véhiculés seulement par les mots que nous pouvons prononcer, mais aussi, et parfois bien plus, par la qualité de la voix qui en est le support. Que deviendrait la célèbre réplique du Cid « Va, je ne te hais point ! » sans une qualité de voix appropriée ! Loin de la littérature, c'est également le cas de nos expressions quotidiennes machinales, qui n'apportent de sens que par la qualité de la voix qui les accompagne.

De même, ce n'est souvent pas par le sens de ce que dit une personne qu'on peut la reconnaître ou la discriminer d'un autre locuteur, mais plutôt par son intonation ou le timbre de sa voix.

Le contrôle de la qualité vocale est donc responsable d'une grande part d'information dans l'expression d'un message vocal. En voix parlée, ce contrôle n'est pas conscient et a toujours été perçu comme secondaire par rapport à l'intelligibilité du message. C'est pourquoi il est intéressant de commencer l'étude de la perception de la qualité de la voix par le domaine de la voix chantée, où à l'inverse de la voix parlée, le contrôle subtil de la qualité vocale par le travail du chant prime sur la recherche d'intelligibilité. Cette maîtrise de la qualité vocale est rendue possible grâce à nos sensations internes et au retour auditif de notre voix, mais également par le développement musculaire (larynx, cage thoracique, diaphragme, abdominaux) et la pratique de techniques vocales spécifiques (gestion du souffle, maîtrise des résonateurs). La voix lyrique est en cela une 1^{ère} étude très intéressante car la qualité vocale y est contrôlée à la fois volontairement et consciemment, mais aussi très précisément.

2.2) Ecoute et catégorisation

Les cadres théoriques de la catégorisation «écologique» (Rosch, 1978 ; Dubois ; 1991 ; 1997 ; Guyot, Castellengo & Fabre 1997) permettent d'amorcer une interprétation des phénomènes de catégorisation qui intègrent des éléments perceptifs des propriétés physiques du signal acoustique et des réponses mémorisées par les sujets à partir de l'expérience (ou expertise) qu'ils ont développés en couplage avec ces paramètres perceptifs. Ainsi on pourra évaluer à partir de l'identification des catégories sémantiques des voix la contribution de propriétés du signal telles que la physique peut la décrire mais aussi en tant qu'indice relié en mémoire à l'identification de la source (homme/ femme par exemple) ou aux connaissances des mécanismes de production des effets vocaux que nos experts maîtrisent également.

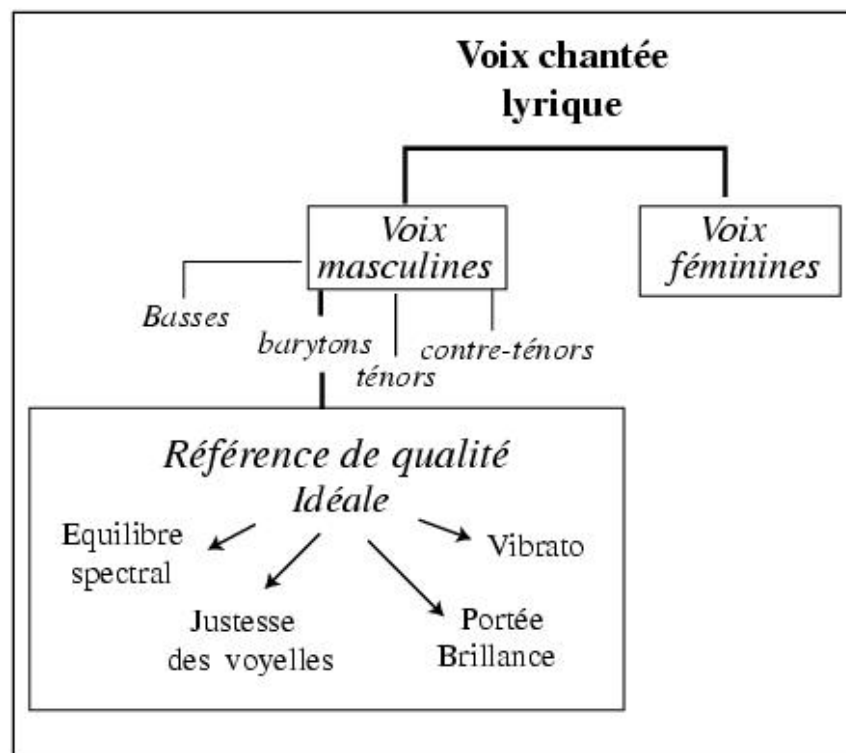


Figure 4 : Schématisation de la représentation de la qualité vocale par rapport à une référence en mémoire du sujet.

Comme illustré sur la Figure 4, l'appréciation de la qualité correspond à l'appréciation des écarts par rapport à une référence en mémoire. Cette référence s'est construite sur la base des expériences personnelles antérieures.

2.3) Problématique : la qualité vocale et le timbre d'une voix

La qualité vocale n'est pas une notion aussi simple qu'il peut y paraître de prime abord. De nombreuses questions se posent, en particulier au niveau de sa définition. Que désigne t'on exactement sous le terme de qualité vocale ?

En effet, si l'on affirme que le sens d'un message ne réside pas seulement dans les mots prononcés, peut-on alors considérer que la qualité vocale fait référence à tout ce qui diffère du contenu sémantique de la production vocale, ou bien ce terme fait-il plus particulièrement appel à la qualité du son, de la même façon qu'on qualifie le timbre d'un instrument de musique ?

Pour l'instant, c'est sur la première conception que se sont basées la plupart des études de reconnaissance d'un locuteur par son empreinte vocale, ou les études sur l'émotion véhiculée par la voix. Les indices de la qualité vocales considérés sont alors aussi bien la prosodie que l'intonation, l'articulation, la hauteur, le rythme, l'intensité ou le timbre de l'émission vocale.

De même, réduire le concept de qualité vocale au timbre ne simplifie pas pour autant le problème. Sa perception et sa verbalisation sont déjà des problématiques en ce qui concerne les instruments de musique et les sons de synthèse (Faure, 2000) car de nouvelles discussions de définition interviennent, selon la signification que la littérature donne au timbre. Certains le caractérisent par les attaques et les fins de sons. D'autres avancent que le timbre serait ce qui distingue deux sons de même intensité et même hauteur.

Dans le cas de la voix, cette dernière conception du timbre nous amènerait logiquement à considérer les voyelles comme en faisant partie. Pourtant, dans la pensée commune, on admet le fait qu'il est possible de faire la même qualité vocale globale sur des voyelles différentes. Et en même temps, on semble percevoir de légères modifications de la qualité vocale sur chaque syllabe d'une phrase. Comment est-t'il alors possible de reconnaître un locuteur s'il change sans arrêt de qualité vocale ?

De plus, dans la plupart des études psychophysiques du timbre instrumental, les sons sont étudiés en ne considérant qu'une seule note. Dans le cas de la voix, cela a-t'il quand même un sens de la considérer comme un instrument de musique comme un autre, de réduire la voix à un son, de s'abstraire du phrasé, de la diction, de la musicalité. Se pose alors un gros problème quant à la localité / globalité du timbre (Payri 2000). Quand on parle de qualité vocale, parle-t'on du timbre en général ou d'un timbre mot par mot ? Et est-il même possible de définir un timbre global ou finalement cette impression générale est-elle fortement influencée une partie du son (en particulier sa fin) ? Si le timbre peut se percevoir de façon globale, alors cela nous ramène au problème de départ de définition du sens à donner à la qualité vocale, car alors l'articulation, la diction, la prosodie qui participent activement à cette perception globale de la voix feraient alors partie du timbre ?

Enfin, un dernier questionnement relatif à la reconnaissance d'individu par leur empreinte vocale et à la définition du timbre est de savoir si le timbre désigne les caractéristiques d'une voix en émission normale, c'est-à-dire cette fameuse empreinte vocale, aussi bien que les variations d'émission au sein d'un même locuteur. Et lorsque qu'on qualifie le timbre d'une voix, est-ce toujours par comparaison à une même référence mentale, auquel cas les voix que nous qualifions semblablement renverraient effectivement à une perception similaire ? Ou bien la qualification de la qualité d'une voix se fait-elle toujours en différentiel par rapport à la voix normale du locuteur (imaginée ou mémorisée) ?

3) Une étude fortement pluridisciplinaire

Jusqu'à présent, les études sur la perception de la qualité vocale ont été menées dans le domaine de la parole. Dans le domaine du chant, les études précédentes se sont plus attachées à décrire uniquement les effets acoustiques et physiologiques de l'émission vocale lyrique. A ce jour, aucune investigation n'a tenté de faire se rejoindre ces deux domaines que sont la perception de la qualité de la voix et son analyse acoustique. C'est donc l'interdisciplinarité et la collaboration entre différentes compétences (linguistiques, perceptives et acoustiques) qui font l'originalité et l'intérêt de cette étude.

3.1) Perception

D'après les résultats des précédentes études sur la perception de la voix parlée, la perception et la verbalisation de la qualité d'une voix semblent donc dépendre de nos attentes en fonction du contexte et des références mentales que nous pouvons en avoir.

En effet, l'humain n'est pas récepteur passif de stimulations extérieures, mais acteur et donneur de sens par sa perception, qui est constituante des objets que nous appréhendons. Bien souvent, les outils de mesure physique du réel ne sont pas capables la plupart du temps de détecter les frontières du monde. S'ils savent le faire, comme en reconnaissance de forme ou de mouvement, c'est parce que leur élaboration a intégré des connaissances propres à l'humain en matière de perception. La catégorisation du monde se fait par la perception qui apporte une interprétation et un sens à la multitude des informations qui nous assaillent à tout moment, et dont la plupart nous sont inutiles. La perception est alors plutôt un processus « top-down » de recherche active d'indices dans ce grand nombre d'informations qui nous parviennent du monde, motivée par nos attentes en fonction du contexte ou de la probabilité d'une situation.

Ainsi, lorsqu'on cherche à étudier la perception de la qualité vocale, la question n'est peut-être pas tant de savoir *comment* nous percevons la voix, mais plutôt de se demander *qu'est-ce que* l'on cherche à écouter précisément dans la voix, quels sont ces indices que nous sélectionnons auditivement pour caractériser et catégoriser une voix ?

Par ailleurs, la voix étant par essence un moyen de communication, d'interaction, d'interlocution, la perception de la voix ne peut pas se concevoir sans dissocier la perception de l'auditeur de celle du locuteur. Il serait intéressant de savoir si ces deux catégories perçoivent et caractérisent la voix de façon semblable ou différente. Il serait également intéressant de voir si cette relation d'interaction entre locuteur et auditeur entre en jeu dans la perception et la verbalisation de la voix. Du point de vue du locuteur, celui-ci semble avoir des sensations exacerbées de par le fait que sa voix vient de l'intérieur de lui-même, mais il n'entend pas directement le son qu'il émet car il ne peut le percevoir que par retour auditif ou par conduction osseuse interne des vibrations. Du point de vue de l'auditeur, la voix ne semble pas être un instrument de musique comme un autre et on peut se demander s'il est possible pour l'auditeur d'arriver à écouter le son d'une voix décorrélée de l'humanité de l'instrumentiste. Certes, dans n'importe quel message musical, c'est bien l'humain qui donne vie au son et fait naître l'émotion. Si nous comparons un son de pure synthèse à un son réel, nous percevons sans problème cette différence. Si nous devons verbaliser cette différence, on peut s'attendre à ce que l'instrumentiste ou son humanité rentre en compte. Cependant, quand on parle d'un timbre d'instrument dans l'absolu, c'est le son et non l'instrument qui est qualifié. Il est rare qu'on parle du clarinettiste lorsqu'on cherche par exemple à qualifier un timbre de clarinette. Mais en est-il de même pour la voix, très liée à l'humain et à l'identité de la personne ? Si on voulait rester sur cette comparaison à un instrument de musique, ne serait-ce pas plutôt la personne toute entière qui en serait l'équivalent plutôt que sa restriction à la voix ? Et par voix, qu'entend-on : la physiologie des cordes vocales et des résonateurs ou bien le son produit ? On voit que même si la voix est source musicale tout comme un instrument acoustique, elle est très différente en tant qu'objet cognitif. Ainsi, il serait probablement très intéressant d'étudier l'impact de l'humain et du corps sur la perception de la voix et pour cela d'adopter une approche méthodologique multisensorielle.

Enfin, l'étude de la perception d'une voix n'a aucun sens si on l'appréhende comme un phénomène instantané. La perception n'existe que par les références mentales construites depuis notre enfance grâce aux expériences sensibles que nous vivons. De plus, nous ne pouvons pas voir l'« instrument vocal » et cela peut avoir une conséquence sur le langage que nous utilisons pour parler de notre perception.

3.2) Verbalisation et linguistique

Pour expliquer le lien entre la perception et le langage, je voudrais commencer par rappeler la fameuse triade sémiotique d'Aristote, représentée figure III

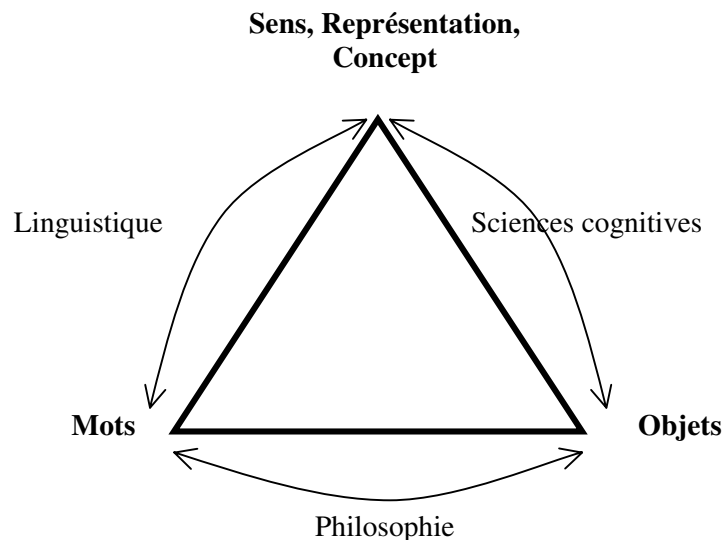


Figure 5 : Triade sémiotique d'Aristote : relation entre mots, objets et concepts

Cette triade permet de représenter les rapports entre le langage, la pensée et le monde. Ainsi, le langage est utilisé pour décrire nos perceptions et renvoie à la fois au réel (dans l'hypothèse qu'il existe) et à nos concepts, c'est-à-dire nos pensées et la représentation que nous nous faisons du monde.

Dans notre cas, l'étude de la perception de la voix nous renseigne sur le rapport qui existe entre la voix réelle (appartenant au monde) et le langage que nous utilisons pour la décrire, mais aussi sur le lien que nous établissons entre les mots auxquels nous avons recours et la représentation mentale que nous avons construit de la voix.

Le langage est bien entendu influencé par notre perception : les mots imitant ou construits sur les sons que nous entendons en sont une bonne preuve (miaulement, craquement, ...). A moins d'y être sensibilisés (tels les linguistes), nous oublions bien souvent que la réciproque est également vraie : nommer quelque chose revient toujours d'une certaine manière à le catégoriser. En cela, le langage a une très grande influence sur notre perception du monde, car il crée des concepts à travers lesquels seulement nous pouvons le percevoir. De nombreux débats ont émergé sur l'existence possible de pensée sans langage, comme dans le cas des animaux par exemple. Il semblerait que des catégories perceptives pures, de l'ordre de l'intuition, précèderaient à la base les catégories linguistiques, mais qu'une fois le langage établi, ces catégories perceptives pures cèderaient la place à des perceptions et des constructions fortement influencées par le langage qui « contamine » le monde dans lequel nous évoluons. Par exemple, nous voyons bien l'influence de la langue maternelle sur la façon dont nous appréhendons le monde du point de vue perceptif.

Jusque là, nous nous posons la question de savoir s'il pouvait exister entre les personnes des perceptions communes de la qualité vocale, ainsi que des descripteurs consensuels pour en parler. A l'inverse, on peut également se demander s'il est possible de reconnaître une voix à partir de sa description ? Cela suggérerait certaines notions relativement univoques renvoyant à des corrélats physiques précis.

Dans le cas du chant lyrique occidental, ceci est un problème puisqu'un jargon bien particulier existe mais semble varier légèrement d'un professeur à un autre. De plus, l'apprentissage dans ce domaine se fait assez tardivement la plupart du temps et le vocabulaire semble assez restreint (ce qui ne veut absolument pas dire qu'on ne décrit pas la voix par des termes extérieurs à ce jargon) Il est donc extrêmement intéressant de pouvoir faire apparaître les termes utilisés pour qualifier la qualité vocale et d'arriver à comprendre les liens qui les mettent en relation, ainsi que de voir jusqu'à quel point les termes que nous utilisons sont ressentis personnellement ou très conventionnels.

De nombreux termes pour qualifier la voix relèvent de la métaphore ou de la synesthésie. Renvoient-ils à des images mentales de la signification première de ces mots, comme la luminosité pour le terme « brillant » ou la couleur pour le qualificatif « blanche », ou ces termes ont-ils désormais acquis une polysémie sonore entière ? On voit ici tout l'intérêt d'adopter une approche multisensorielle pour traiter de la perception de la voix chantée.

Plusieurs recherches ont déjà été menées sur la verbalisation associée à la perception de la voix parlée. Les premières recherches sur la qualification de la voix ont été effectuées par Voiers (1974), qui a introduit le concept de « différentiateur sémantique » en utilisant 49 échelles bipolaires pour évaluer des voix très diverses. Il a ensuite défini un second protocole de 75 échelles pour juger les caractéristiques d'une voix. Le but était de pouvoir décrire entièrement une voix afin par exemple de faire de la reconnaissance ou de la discrimination. Dans les échelles (en anglais) relatives à la qualité de la voix elle-même on trouvait des termes liés au timbre (*loud-soft, heavy-light, clear-hazy, tense-relaxed, rising-falling, low-high, rough-smooth*), les autres critères d'appréciation étant plus hédonistes (agréables/désagréable), psychologiques, indicateur de reconnaissance (*familiar-strange*) ou caractérisant le locuteur. Une bonne partie de la recherche selon des axes décrite dans la littérature s'est basée sur des adaptations de ce protocole. En général, la littérature utilise un nombre moindre d'axes : par exemple Schmidt-Nielsen et Stern (1985) a sélectionné des facteurs parmi les 75 de Voiers, après une analyse factorielle éliminant des redondances. De même, Payri (2000) a réalisé une évaluation en français des voix pathologiques selon des échelles bipolaires. Dans le cadre de la description de voix pathologiques (Payri 2000), ces qualificatifs sont plus liés à la physiologie de l'appareil vocal (voix laryngée) ou à des fonctionnements anormaux des cordes vocales (aphone, rugueuse, diplophonique). Du côté des voix de serveurs téléphoniques (Thibaut Ehrette 2003), la qualité vocale renvoie à l'identification du genre, de l'âge ou du comportement supposé d'un locuteur (timide, prétentieux...), mais aussi à des modifications de timbre du son émis (voix chuchotée, rauque).

Une autre manière d'appréhender les constructions mentales consiste tout bonnement à les faire parler librement du sujet, en leur posant des questions ouvertes ou en leur demandant de décrire leur perceptions. Dans ce genre d'entretien ouvert dirigé, on ne considère pas comme en psychanalyse que les choses sont cachées. La verbalisation libre a pour but de partir à la recherche du sens commun. On ne cherche pas tant à révéler le discours mais plutôt à l'établir.

3.3) Relation entre l'acoustique de la voix et la qualité vocale

S'il paraît inévitable que la perception est une expérience subjective, existe-t-il malgré tout certains consensus sur nos perceptions de la voix et sur la verbalisation que l'on utilise pour la décrire ? Et cela correspond-il à des critères acoustiques objectifs ?

Avant tout : d'où provient le changement de qualité vocale, d'une modification de la forme des résonateurs, d'une gestion du souffle différente, d'un changement de vibration des cordes vocales ?

Sous un angle acoustique, le but est alors de trouver des paramètres physiques pertinents associés à une qualité vocale donnée et de les mettre en relation avec la verbalisation utilisée par le chanteur ou l'expert de la voix.

Pour l'instant, l'analyse acoustique a principalement été effectuée sur des paramètres temporels du signal acoustique audible (Amplitude du son, du bruit, Fondamentale) et spectraux (spectre, spectral tilt). Ainsi la détermination de l'âge ou du sexe d'un locuteur à l'écoute de sa voix a été faite par l'étude de la hauteur (Henton, 1989), de la « breathiness » (Henton) ou de la « crackiness » (Tielen 1992). Les hommes auraient en effet tendance à avoir une voix plus nasalisée et creaky que les femmes.

Pour Miller, ce serait plutôt la source glottique qui serait responsable des différentes qualités vocales productibles. Grâce à l'electroglottographie, cette hypothèse a été confirmée par Monsen&Engrebetson, puis Gauffin&Sunberg. De nouveaux paramètres pertinents ont été extraits du mouvement vibratoire des cordes vocales pour caractériser la qualité vocale. C'est le cas du débit glottique dont l'influence sur la qualité est représentée figure IV :

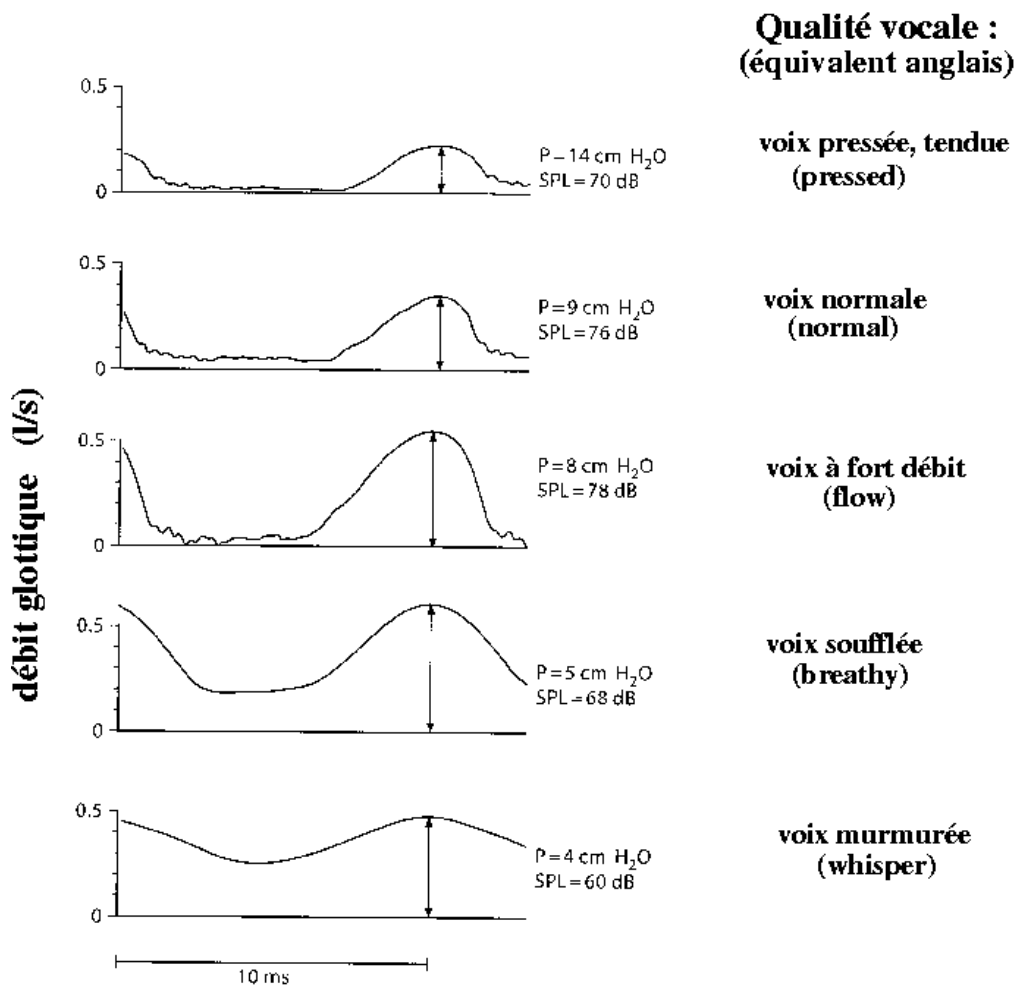


Figure 6 : Influence du débit glottique sur la qualité vocale d'après Henrich, 2001

En particulier, le quotient ouvert serait lié en voix parlée à l'impression de tension de la voix. Le quotient ouvert est défini par le rapport de la phase ouverte du cycle (là où un débit passe à travers la glotte) sur la période du cycle glottique. Les instants d'ouverture et de fermeture de la glotte sont plus particulièrement repérables grâce à la dérivée du signal électroglottographique. Voici un schéma récapitulatif de la définition du quotient ouvert :

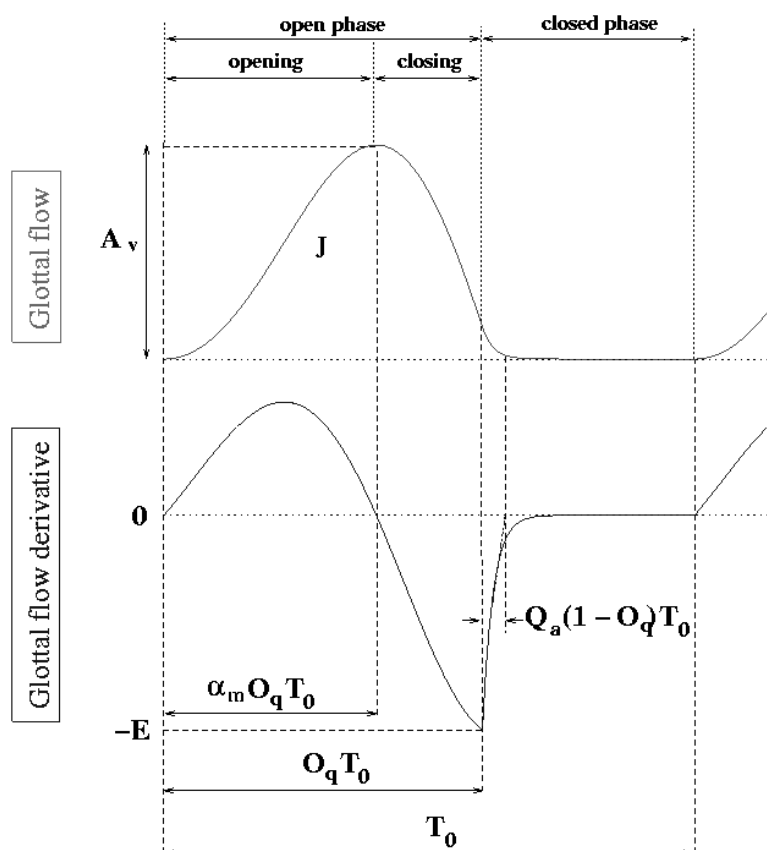


Figure 7 : schéma de définition du quotient ouvert d'après Henrich, 2001

En voix chantée, le quotient ouvert est corrélé à l'I vocale en mécanisme I(voix de poitrine) et à la fréquence fondamentale en MII (voix falsetto chez les hommes, voix de tête chez les femmes). On peut se demander si le quotient ouvert serait des paramètres pertinents pour expliquer certaines autres modifications de qualité vocale. Si tel n'est pas le cas, quels pourraient être les autres paramètres acoustiques ou spectraux à prendre en compte ? Par exemple il existe des paramètres spectraux qui permettent de décrire le timbre instrumental comme la brillance

En conclusion, le but général visé par cette étude est de comprendre comment nous percevons la qualité de la voix, quelles sont nos références mentales dans ce domaine, quels sont les descripteurs verbaux et acoustiques pertinents pour la qualifier et s'il y a corrélation entre ces deux classes d'indices.

Analyse de la verbalisation des chanteurs

1) Objectifs

Afin de commencer cette étude, nous avons utilisé une base de données audio et électroglottographique préexistante de 18 chanteurs classiques, qui contient également la verbalisation donnée par ces chanteurs. Cependant, cette base de donnée n'avait pas du tout été enregistrée dans le but d'étudier la verbalisation libre des chanteurs et par conséquent, les interventions de l'expérimentateur n'étaient pas enregistrées (que ce soient des questions ou des affirmations). Néanmoins, la verbalisation du chanteur est en elle-même intéressante, car il s'agit d'un cas particulier de la perception de la qualité vocale, puisqu'elle concerne le producteur du son lui-même. Elle peut englober des perceptions multisensorielles. Cependant, le fait que ce soit un cas particulier dans la perception de la voix est à la fois intéressant mais ne doit pas rester du tout le seul axe d'observation de cette perception, car l'étude serait alors fortement biaisée. L'analyse de ce corpus des chanteurs ne peut donc être qu'une piste de réflexion préliminaire, pour organiser une étude plus poussée de la verbalisation libre d'experts de la voix lyrique.

De plus, dans l'optique de mettre au point également un test d'évaluation selon des échelles bipolaires, il était intéressant de choisir les termes de ces échelles par rapport à la verbalisation collectée par cette première étude sur la voix chantée, et non pas d'après la littérature sur la voix parlée, beaucoup moins adaptée à notre sujet.

Enfin, cette première étude préliminaire ne sert pas seulement à extraire le vocabulaire pour décrire la voix ainsi que sa fréquence d'utilisation ou les relations entre les termes utilisés. Cela fournit également de nombreuses pistes sur la façon dont on conçoit le terme de qualité vocale et s'il existe un consensus sur ce concept, sur les différents objets rentrant en jeu dans la perception de la voix, sur le comportement des personnes interrogées dans le domaine du chant lyrique, afin d'adopter une attitude en conséquence pour à la fois les rassurer et obtenir des résultats exploitables et intéressants.

Par ailleurs, il faut préciser que ce stage représentait mes premiers pas à la fois en recherche, en linguistique et dans le domaine lyrique. Cette étude préalable m'a ainsi permis de me mettre en même temps à niveau dans ces différents domaines, de façon à être au point pour l'étude des experts.

Toute cette première étude préalable a donc été effectuée sans perdre de vue son but : mettre au point un test d'écoute sur ces mêmes exemples sonores, cette fois avec des experts de la voix lyrique.

2) Matériel et méthodes

2.1) Enregistrements et corpus des chanteurs : Provenance, contexte d'enregistrement et consigne donnée aux chanteurs.

Dans un premier temps, nous avons repris des exemples vocaux enregistrés au cours la thèse de N. Henrich portant sur l'« Etude de la source glottique en voix parlée et chantée ». Cette base de données de signaux audio et électroglottographiques avait été recueillie avec le concours de 18 chanteurs classiques de toutes tessitures (3 sopranos, 3 mezzo-soprani, 3 contre-ténors, 2 ténors, 7 basses), placés dans une chambre insonorisée. Parmi les différentes phases de ce test, il leur a été demandé de chanter deux phrases musicales (Les premières mesures de l'ave Maria de Gounod et l'arpège du petit Rossini) en utilisant différentes qualités vocales laissées à leur libre choix, puis de verbaliser cet essai de changement de qualité vocale.

Voici l'exact intitulé du protocole de ce test :

Ave Maria:

Reprendre cette phrase musicale en essayant de modifier la qualité vocale (à votre appréciation). Précisez oralement les différentes qualités que vous produirez.

Rossini :

Selon votre inspiration, vous pourrez ensuite reprendre l'arpège en essayant de modifier la qualité vocale, comme précédemment.

Le matériau de notre étude préliminaire n'a pas été les signaux acoustiques de ces enregistrements mais leur transcription, avec l'intervention éventuelle de l'expérimentateur et l'instant de verbalisation du chanteur. On ne pouvait malheureusement que déduire les réponses données par le chanteur sans savoir ce qu'avait dit l'expérimentateur. De plus, le chanteur était fortement influencé dans sa verbalisation car on lui suggérait des idées quand il en manquait.

L'intérêt de relever les instants de verbalisation du chanteur est de nous renseigner sur la façon dont a évolué le concept de voix entre l'instant qui précède son émission (connaissance antérieure du monde, intention d'émission, références mentales, voix intérieure au sujet) et celui qui suit (la voix produite est elle alors perçue comme un objet extérieure au sujet, ajustement et rectificatif sur le résultat produit par rapport à l'intention de départ).

Pour le détail des transcriptions, voir dans l'annexe ci-jointe la partie relative au corpus des chanteurs

2.2) Analyse sémantique et linguistique

Cela a été l'occasion d'une première étude linguistique de l'environnement sémantique du chant lyrique, afin de dégager certains termes, certains liens entre eux, les différents objets entrant en jeux pour parler de la qualité vocale et la façon dont le sujet se situe par rapport à l'objet qu'il a à définir.

Dans le corpus des chanteurs, constitué par la verbalisation recueillie autour des essais vocaux des chanteurs, nous avons relevé puis étudié plusieurs pistes sémantiques par pupitre (sopranos, mezzo-sopranos, contre-ténors, ténors et basses), puis par sexe, et enfin tous chanteurs confondus.

Dans le but de mettre en évidence le positionnement des chanteurs par rapport à la voix émise et perçue ainsi que le rattachement ou non de leur expérience sensible à celle des autres individus, nous avons relevé dans leur verbalisation

- les marques de la personne (pronoms personnels, impératifs, marques possessives...) qui peuvent être des indices de constructions subjectives, objectives, individuelles ou collectivement partagées.

je/j'	moi	me/m'	me	Ma	voyons	on
-------	-----	-------	----	----	--------	----

- les références à la perception de la voix en tant qu'objet (aussi bien les *ce*, *ça*, que les objets clairement cités comme la voix, la qualité, la voyelle...)

qualité	voix	voyelle	une voix / des voix	la voix / la voyelle	sur une voix ... / une voyelle	avec une voix	en voix
---------	------	---------	---------------------	----------------------	--------------------------------	---------------	---------

truc/ chose	un	Il	Le	en	y	il y a	rien	n'importe quoi	une espèce /sorte de	de jazz	de façon / de qualité	le plus possible	des résonances de	nez/ tête/ poitrine
----------------	----	----	----	----	---	-----------	------	-------------------	-------------------------------	------------	--------------------------------	------------------------	-------------------------	---------------------------

ça	là dessus	c'	ce que	quoi / qu'	quel/quelle	l'ave maria/ l'arpège/ le reste/ le début/ la version	un, le do /sol/ "e"	la/ une première/ nième fois	les mêmes	d'autres .../ un autre
----	--------------	----	-----------	---------------	-------------	---	---------------------------	------------------------------------	--------------	---------------------------

En examinant les marques de la personne ou de l'objet, on découvre que les adjectifs sont souvent substantivés : les chanteurs parlent « du tubé » ou « du clair », sans que l'on sache quel est l'objet sous-entendu qualifié de clair ou tubé, ou même si un objet est sous-entendu. A l'inverse, certains noms deviennent des adjectifs dans le discours lorsqu'ils représentent une qualité vocale. C'est le cas de « chanson » ou « variété », utilisés exactement comme d'autres adjectifs tels que « brillant » ou « naturel ».

En effectuant ce genre de relevé, on se rend également compte de l'importance du contexte d'utilisation des mots et des prépositions qui les accompagnent, mettant en évidence la polysémie d'un mot et permettant de comprendre les différents objets cognitifs auxquels ils renvoient. De plus, il est difficile de savoir où couper intelligemment les phrases à analyser, et si on doit considérer uniquement les énoncés relatifs à la qualification directe des essais vocaux ou aussi les discours relatifs à la voix en général.

De plus, on se heurte très rapidement à des problèmes de polysémie et d'interprétation obligatoire de la part de l'analyste, même pour celui qui voudrait essayer de rester le plus neutre possible. C'est le cas par exemple pour le pronom « On » où il n'est pas toujours bien évident de déterminer s'il est utilisé à la place de « Je », s'il s'agit d'un collectif incluant le locuteur et l'interlocuteur ou encore s'il fait référence à un « nous, humains » ou plus particulièrement à un « nous, chanteurs ». C'est seulement en fonction du contexte d'énonciation et de l'interprétation de l'analyste que l'on peut statuer. On voit alors le problème posé par l'analyse automatique des corpus proposée par certains logiciels récents.

Nous avons également été frappé par l'utilisation massive de modalités (c'est-à-dire des marques du doute comme l'usage du conditionnel, les expressions comme « peut être », « essayer de », les recherches de confirmations...etc). En parallèle, nous avons remarqué la présence non négligeable de vocabulaire lié au sentiment de contrainte (devoir, falloir, être obligé), d'inconfort ou d'incapacité (pouvoir).

Si nous regroupons ici ces différentes marques car ils nous semblent qu'elles entretiennent des relations entre elles : Les marques d'inconfort ne sont-elles pas une sorte d'excuse pour expliquer leur éventuelle incapacité ? De plus, l'incapacité n'est jamais exprimée en tant qu'affirmation mais plutôt comme doute de ses capacités. On peut donc le relier aux modalités. La sémantique de l'obligation montre également la négociation d'un consensus entre le chanteur et l'expérimentateur. Ce peut être également l'indication d'un discours normatif (il faut, on doit, à la place d'un discours personnel à la première personne). Enfin, les recherches de confirmation peuvent être rapprochées des modalités et des obligations, ne marquant pas forcément l'incertitude des idées exprimées, mais peut être plutôt une façon « douce » de les imposer en les introduisant par « je sais pas » ou sous forme de questions. On voit là encore la difficulté d'une analyse objective sans interprétation.

Afin d'étudier l'impact des propres sensations du chanteur sur sa production vocale, ainsi que son rapport au monde, nous avons décidé d'adopter une approche multisensorielle et de relever les termes liés au corps et à la sensation.

Afin d'examiner plus particulièrement le lien entre la production vocale et sa perception, entre l'intention et le résultat, nous nous sommes penchés sur les verbes d'action et les temps utilisés (présent, passé, futur, participes présents), mais aussi sur les termes d'introduction des commentaires avant et après un essai.

Nous avons également fait une liste de tous les qualificatifs utilisés pour caractériser l'essai effectué, en particulier les termes techniques du jargon lyrique.

Pour finir, nous nous sommes intéressés aux comparatifs et à la synonymie dans le discours (sens proches, contraires, termes allant de pair).

3) Résultats

3.1) Liste des qualificatifs utilisés et consensus sur la qualité vocale

Voici la liste des adjectifs et expressions relevés dans le corpus des chanteurs pour qualifier la qualité vocale, et leur fréquence d'apparition. Nous avons mis en gras les termes les plus fréquemment utilisés et regroupés les termes en catégories par souci de lisibilité seulement.

<u>Jugements</u>	<u>de</u>	<u>valeurs</u>	1 Rural	1 Homogène
<u>hédonistes</u>			1 Populaire	1 facile
1 Joli			1 Comme un benêt	2 difficile
1 bon			1 Qualité meridique	3 évident
1 mal			1 Grotesque	1 délirant
1 Vulgaire			1 simple	

Métaphores

4 voix blanche

1 nasillard

19 nasal

21 clair

17 sombre, 1 sombré, 1

assombrir

1 étouffé

1 brillant

2 serré, 1 en serrant

2 soufflé, 2 avec du souffle,

4 avec de l'air

1 en élargissant

1 plat

1 gros, 3 grossi

2 droit

1 en forçant

1 accentué

2 couvrir

Référence à des expériences

vécues personnelles ou

universelles

1 XVIème arrondissement

1 Charcutière

1 En solfège

1 enfantin

2 castafiore

Noms de chanteurs ou de

personnalités connues

2 Enrico Macias

1 Johnny

1 Florent Pagny

2 Robert Dumay

2 Aznavour

1 Whitney Houston

1 Renaud

1 Jane Birkin

Caractéristiques (physiques)

du son

3 Timbre

2 Timbré/ 3 détimbré

1 tessiture, **3 grave**, 1

octave, 2 hauteur

1 nuance, **3 fort**, 2 piano

1 tempo, 1 lent

3 vibrer/1 vibré

1 piqué, 1lié

1 qui résonne

1 harmoniques graves

1 en sol

Styles musicaux

2 Chanson, 2 chantonné, 1

en chantonnant, 2 chanté.

2 Baroque 17^{ème}

2 Lyrique 19^{ème}

1 Opéra italien

1 avec des ornements

7 variété

3 A la russe

1 Voix bulgare

1 La Walkyrie

1 Crooner

3 Voix parlée

1 jazz

4 style

1 musiques improvisées

Technique vocale /

placement de la voix...etc

5 tubé, 3 engorgé

2 résonances de tête, 2 de poitrine

3 avec vibrato

5 sans vibrato

1 décrochement

1 technique

12 sur telle voyelle

1 sans technique vocale

2 falsetto, 5 fausset, 5 en

voix de tête, 4 en voix de

poitrine, 1mezzo voce un

peu mixte

3 ouvert

1 en pensant en avant, **3 en**

arrière

1 devant, 2 derrière

2 dans le nez

3 dans le pif

1 sur le larynx, 2larynx haut,

3 larynx bas

1 accroché

2 voile du palais tendu

1 ouverture de bouche très

latérale

1 la bouche en hauteur

1 en avalant la voix

1 voix de canard

20 naturel

16 lyrique

4 normal

1 travaillée

4 non chanteur

1 chanteur

Ensuite, nous avons étudié les premières phrases exprimées par le chanteur après lecture de la consigne (« changer la qualité vocale »), les questions qu'il posait afin de la comprendre ou les paraphrases qu'il faisait afin de la préciser. Tout cela nous renseigne sur la recherche d'un consensus sur le terme de la qualité vocale : Où se situe l'évidence concernant ce concept et quelles sont les divergences ?

Nous avons ainsi vu émerger plusieurs conceptions de la qualité vocale dont nous proposons une catégorisation, et qui corroborent les résultats obtenus en dressant la liste des qualificatifs ci-dessus.

Tout d'abord, la qualité vocale renvoie au **SON** perçu, caractérisé par ses **attributs vibratoires ou spectraux** (timbré, détimbré, vibrer, qui résonne, harmoniques graves...), par sa **hauteur** (grave, aigu, tessiture, octave...) ou son **intensité** (nuance, fort, piano), ou décrit par des **métaphores** (blanc, clair, sombre, étouffé, brillant, serré, soufflé, gros, plat, droit). Dans notre cas, c'est cette notion de timbre qu'il nous intéresse d'explorer dans le but de trouver des corrélats acoustiques aux termes utilisés.

Cependant, la qualité vocale, pour les chanteurs, semble renvoyer également très fortement à des **TECHNIQUES VOCALES**, aussi bien en tant qu'**ACTION** sur les **attaques** (piqué, lié), sur le choix de la **voyelle de support**, sur le contrôle du **vibrato** (avec ou sans), sur le placement du son (dans le nez) ou le mécanisme utilisé (voix de poitrine, de tête, fausset), qu'en tant que **SENSATION** de ces techniques (tubé, engorgé, ouvert, en avant, en arrière, devant, derrière, larynx haut ou bas, accroché, voile du palais tendu, ouverture de bouche latérale....)

De plus, les chanteurs semblent parfois manquer de mots pour décrire la qualité vocale produite et recourent à des notions de **STYLE** pour faire partager leur perception. Ce peut donc être des références à des **style musicaux** (variété, baroque, chanson, opéra italien, variété, à la russe, Voix bulgare, Crooner, Voix parlée, jazz, musiques improvisées...) ou plus simplement à des **voix connues** (Enrico Macias, Johnny, Florent Pagny, Aznavour, Whitney Houston, Renaud, Jane Birkin, ...) ou renvoyant à une **mémoire collective** (Charcutière, castafiore, XVIème arrondissement). De façon étonnante et malgré tout marginale, le **tempo** fait également partie du concept de qualité vocale pour certaines personnes.

Enfin, il est évident que la qualité vocale a une grande signification **HEDONISTE**, de part la polysémie déjà du terme « qualité » renforcée par les critères d'esthétiques très exigeants du chant lyrique. La notion de « qualité vocale » est donc directement comprise par certains comme un **jugement de valeur** sur une voix perçue (bon, mal, naturel, normal, travaillée...). De plus, comme dans toute étude sur la perception d'un phénomène, on ne peut passer à côté de la notion d'**agrément** (Joli, Vulgaire, Rural, Populaire, grotesque, homogène, délirant...).

3.2) La perception de la voix comme expérience individuelle ou collective

Le tableau (figure V) représente la répartition en 3 catégories des marques de la personne relevées dans le corpus des chanteurs: un premier groupe rassemble les marques renvoyant au chanteur seul, le deuxième, ceux correspondant à l'expérimentateur, enfin le troisième regroupe les marques du collectif englobant locuteur et expérimentateur ou le chanteur et les autres humains.

	je/j'	moi	me/m'	me	ma	on individuel = je	total	Tu/ t'/ te	impératif 2de personne	Vous	total	on collectif = nous	on général : humanité, chanteurs	voyons	total
%	69,00%	4,00%	1,00%	1,60%	0,50%	4,00%	80,10%	13%	0,50%	0,50%	14%	3%	2%	0,50%	5,20%

Figure 8 : répartitions des marques de la personne renvoyant au sujet ou à un collectif

On voit que la perception de la voix dans un contexte de production est une expérience extrêmement individuelle. En première approche, cela serait l'indication que les termes utilisés pour décrire la voix peuvent donc être considérés comme étant assez personnels, et l'expérience perceptive de la qualité vocale comme très subjective.

On note cependant que si l'expérimentateur n'est pas tant cité que cela, son apparition dans le corpus est toujours associé à des notions d'obligations ou de recherche de confirmation. L'expérience sensible du chanteur semble donc être personnelle mais dans la volonté de faire plaisir à son interlocuteur, dans l'intention de dire ce des choses qu'on attendrait de lui. Nous reviendrons sur ce point dans le paragraphe sur la normativité du discours.

3.3) Positionnement du sujet par rapport à la voix

Le positionnement du sujet par rapport à l'expérimentateur, par rapport aux autres chanteurs et par rapport aux autres humains est déjà une information importante sur la sincérité de sa perception. L'examen des marques de la personne permet également de « mesurer » l'importance de la présence du sujet par rapport à ce dont il parle, par la répartition entre les pronoms renvoyant au locuteur et les indices renvoyant à ce dont on parle.

Cette répartition a permis de dégager 3 catégories de chanteurs :

- des « auditeurs » utilisant des termes renvoyant essentiellement à des techniques vocales, et pour lesquelles la voix est perçue comme un objet.
- Des « producteurs » où les verbes d'actions et le discours à la première personne prédominent.
- Enfin une catégorie de personnes ne faisant aucune thématisation, c'est-à-dire qui n'introduisent pas leurs qualifications. Lorsqu'on leur demande de décrire la qualité vocale, elles ne répondent pas « Je trouve que c'est clair » ou « C'est clair », mais directement « clair » par exemple. Ces personnes étaient probablement influencées par l'expérimentateur qui leur demandait d' « exprimer un qualificatif ».

Auditeurs

	S1		S2		T2		B1		B5		B6		B7	
Locuteur impliqué	1	33%	2	33%	6	26%	1	7%	1	33%	4	25%	6	43%
Ce dont on parle	2	67%	4	67%	17	74%	13	93%	2	67%	12	75%	8	57%
total	3	100%	6	100%	23	100%	14	100%	3	100%	16	100%	14	100%

Producteurs

	S4		A2		A3		CT1		CT2		T1		B4	
Locuteur impliqué	14	61%	5	71%	17	65%	13	65%	17	68%	13	76,50%	3	100%
Ce dont on parle	9	39%	2	29%	10	35%	7	35%	8	32%	4	23,50%	0	0%
total	23	100%	7	100%	26	100%	20	100%	25	100%	17	100%	3	100%

Autres

	S3		A1		CT3		B2		B3	
Locuteur impliqué	9	53%	17	51%	10	48%	18	46%	3	50%
Ce dont on parle	8	47%	16	49%	11	52%	21	54%	3	50%
total	17	100%	33	100%	21	100%	39	100%	6	100%

Figure 9 : Occurrences des pronoms pour chaque chanteur. Classification des chanteurs en trois groupes plutôt objectif/plutôt subjectif/neutre

Lorsqu'on effectue un comptage de certaines marques linguistiques sur un corpus, il se pose un problème, car les sujets ne parlent pas autant en moyenne. Il faut donc tenir compte du poids de chaque locuteur. Les résultats du comptage ne peuvent être comparés qu'en pourcentage et pas en cumulant le nombre d'occurrences des marques linguistiques recherchées.

	4 Sopranes		3 Mezzo soprani		7 Femmes		3 Contre-ténors		2 Ténors		7 Basses		12 Hommes	
Locuteur impliqué	26	53%	39	58%	65	56%	40	61%	19	48%	36	38%	95	47%
Ce dont on parle	23	47%	28	42%	51	44%	26	39%	21	52%	59	62%	106	53%
total	49	100%	67	100%	116	100%	66	100%	40	100%	95	100%	201	100%

Figure 10 : Répartition des pronoms dans les corpus tessiture et sexe

Voici par exemple les résultats cumulés par pupitre :

Les résultats sont alors influencés par la personne du pupitre qui parle le plus, ce qui nous amènerait à une conclusion erronée qui consisterait à dire que globalement, la perception est plus subjective pour les pupitres aigus que pour les pupitres grave.

En fait, si on regarde le tableau suivant, tenant compte du poids du locuteur, et issu de la répartition des chanteurs dans les 3 catégories expliquées plus haut, on se rend compte qu'il n'y a pas de corrélation directe entre le pupitre et le positionnement du chanteur vis à vis de la voix produite.

Par contre, lorsqu'on ne considère que la répartition des marques linguistiques relevées par chanteur et non pas leur cumul, il se pose un autre problème sur ce corpus très petit : certains

chanteurs ont très peu parlé et donc les pourcentages ne sont pas très représentatifs de leur discours. On peut même se demander s'il ne vaudrait pas mieux éliminer ces sujets du corpus.

	Sopranes		Mezzo soprani		Contre-ténors		Ténors		Basses		Tous	
Objectif	2	50%	0	0%	0	0%	1	50%	4	57%	7	37%
Subjectif	1	25%	2	67%	2	67%	1	50%	1	14%	7	37%
Neutre	1	25%	1	33%	1	33%	0	0%	2	29%	5	26%

Figure 11 : Répartition des chanteurs dans les catégories objectif / subjectif / neutre

Pour conclure cette discussion, il faut rappeler que cette étude linguistique n'est que préliminaire et s'effectue sur un corpus non dédié à ce but lorsqu'il a été enregistré. Cela n'a donc pas beaucoup de sens de chercher à en faire une analyse très rigoureuse. C'est plus l'impression générale qu'il faut en tirer.

Si l'on considère toutes les phrases où le sujet parle de ce qu'il produit et non plus seulement les marques de l'objet (« ça », « ce que »...), tous les chanteurs considèrent globalement la voix comme un objet extérieur à eux. Si certains chanteurs s'impliquent personnellement plus que d'autres, cela n'a a priori aucun rapport avec leur tessiture.

3.4) Emergence de plusieurs objets autour du mot « voix ».

En examinant les objets qualifiés par le chanteur pour parler de la voix, nous avons constaté l'importance des différentes prépositions introduisant le mot « Voix » : lorsque la voix est sujet ou COD d'une phrase, le terme de « Voix » est considéré comme l'objet produit. (« La voix est claire » ou bien « j'entends / je fais une voix claire ». Mais la voix ne semble pas être résumée à cette image et est également considérée comme le support de la production (« sur la voix » ou « dans la voix »), comme moyen de production (« avec une voix claire ») ou encore comme style de production (« en voix claire »)

Cependant, la voix n'est absolument le seul objet cognitif à être qualifié lorsqu'on veut décrire la qualité vocale. Dans la « catégorie » des objets produits lors d'une émission, on trouve également la phrase musicale chantée, la qualité. En fait, le chanteur s'en sort par une pirouette pour éviter de citer l'objet de sa production, il dit « je vais faire tubé » ou bien « je vais en faire un tubé », sans que l'on sache exactement ce qu'il qualifie.

Pour ce qui est du support de la production vocale, la voyelle et la note semblent jouer un grand rôle.

Enfin, d'autres objets gravitent autour de la notion de production vocale et leur contrôle est en soit une manière de produire une qualité vocale différente. C'est le cas du vibrato et du souffle (avec ou sans) qui se greffent sur l'objet produit (par exemple la voix) ou sur le support (la voyelle, la note ou la phrase musicale).

3.5) Premières relations entre les différents termes utilisés

L'examen des termes de sens proches, de sens contraire ou des termes allant de pair dans le discours permet de dresser un premier bilan des relations entre objets cognitifs et qualificatifs utilisés pour décrire la qualité vocale.

Commençons par les images mentales relatives aux mécanismes utilisés par le chanteur : la voix de tête, aussi appelée fausset ou falsetto est associée d'une part à la notion de naturel, à la façon dont on peut chanter lorsqu'on n'a jamais pris de cours de chant. Mais pour d'autres, il semble que ce mécanisme renvoie au contraire au caractère lyrique d'une voix. Malgré cette grande différence de

perception entre les chanteurs, la voix de tête est par contre unanimement liée à une impression de hauteur, qu'elle soit spatiale (en haut par opposition à en bas) ou fréquentielle (aigu). Fréquemment, la voix de tête s'accompagne d'un affaiblissement de l'intensité vers le ou d'ajout de souffle dans la voix. A l'inverse la voix de poitrine est perçue comme grave et sans avoir connaissance de la physiologie de l'appareil vocal, les chanteurs perçoivent bien un « décrochement » au passage d'une de ces mécanismes à l'autre.

Le naturel qualifie aussi bien directement une voix que la manière de chanter. Globalement, cela correspond à une absence de vibrato et fait appel à la naïveté devant l'apprentissage du chant ou à l'époque où les chanteurs apprenaient le solfège, en tout cas à l'absence de technique vocale (sauf pour une personne qui dit ne jamais avoir chanté autrement). Les non chanteurs sont perçus comme « ouvrant comme des benêts » et ne liant pas les notes, chantant le plus simplement possible. Pour certains, le naturel correspond à une émission semblable à la voix parlée, pour d'autres à une production vocale en voix de tête, c'est-à-dire tel qu'on peut le faire en étant enfant. La notion de naturel n'est pas loin de celle de normalité, associée à des chanteurs de variété, ou au concept de chanson (dans sa polysémie d'émission enfantine mais aussi comme style musical comparable à la variété).

A l'inverse, la voix lyrique est une voix technique et travaillée, avec du vibrato. Les chanteurs eux-mêmes ne coupent pas à l'image caricaturale de la castafiore et imagine la voix lyrique plutôt en voix de tête, d'intensité sonore forte et sur la voyelle « a ». Pour un des sujets, la voix lyrique se réfère seulement à la voix chantée, par opposition à la voix parlée.

En ce qui concerne le vibrato, il semble associé au concept de brillance. Il ne semble pas évident que le verbe « Vibrer » et le vibrato renvoient à la même perception auditive. En tout cas, indiscutablement, le vibrato est associé à la voix lyrique et son absence aux styles baroque, chanson, et à la voix naturelle. Ensuite, les notions de son droit et d'absence de vibrato semblent aller de pair. Sont-ils synonymes ? De même, il paraît approprié d'adapter le vibrato en fonction de la clarté de la voix (sombre, beaucoup de vibrato, clair, pas trop). Est-ce une adaptation entre les deux concepts ou bien serait-ce justement une explication acoustique de la perception de la clarté dans une voix ?

Car pourtant, indépendamment du vibrato, la clarté est manifestement liée à l'ouverture (notamment à celle de la bouche sur la voyelle « é ») et à l'accrochage. Ce sont la plupart du temps la voix ou les voyelles qui sont ainsi qualifiées. La notion de clair / sombre renvoie également à des styles de chant, plutôt « à la russe » pour les voix d'hommes sombres et chantonné pour les voix claires. Au niveau timbral, la notion de sombre est plutôt associée à la présence d'harmoniques graves et la notion de clair semble corrélée à celle de détimbré.

A noter qu'une position haute du larynx semble provoquer une perception d'un son détimbré. De même, on peut remarquer que les qualificatifs « tubé » et « engorgé » vont de pair avec celui de « à la russe ».

En première approche, nasal et timbré ont l'air synonymes. La nasalité renvoie la plupart du temps à des positions du corps et des résonateurs de tête : très devant, avec une ouverture de bouche latérale. Son contraire semble être engorgé, en arrière. Pour sa part, le terme « en arrière » est associé aux qualificatifs « ouvert » et « gros », à l'impression d'avaler la voix. De son côté, l'adjectif engorgé revient presque tout le temps avec le terme tubé. Peut-on alors considérer que sombre, tubé, engorgé, en arrière renvoient à la même qualité vocale ? Cela ne nous paraît pas si simple. En effet « tubé » est lui-même associé (et non pas synonyme) à une position basse du larynx (mais est-ce pour autant la même chose que « appuyé sur le larynx » ?). Lorsque le chanteur agit sur la position du larynx, il fait très souvent des commentaires sur la tension du voile du palais, sans qu'il y ait de corrélation entre les deux positions, mais cela nous paraît intéressant à souligner.

Enfin, la perception d'air sur la voix amène à qualifier le son ou la voix de « soufflé ». Il semble que le chanteur ait une sensation similaire lorsqu'il est momentanément malade.

Une seule déception : nous obtenons dans ce corpus très peu d'information en matière de position ou de timbre en ce qui concerne la notion de voix blanche. Le seul exemple qui nous soit donné de ce type de voix est le chanteur Renaud. A nous d'interpréter : s'agit il d'une voix détimbrée, soufflée, rauque, sans voix ou tout cela à la fois?

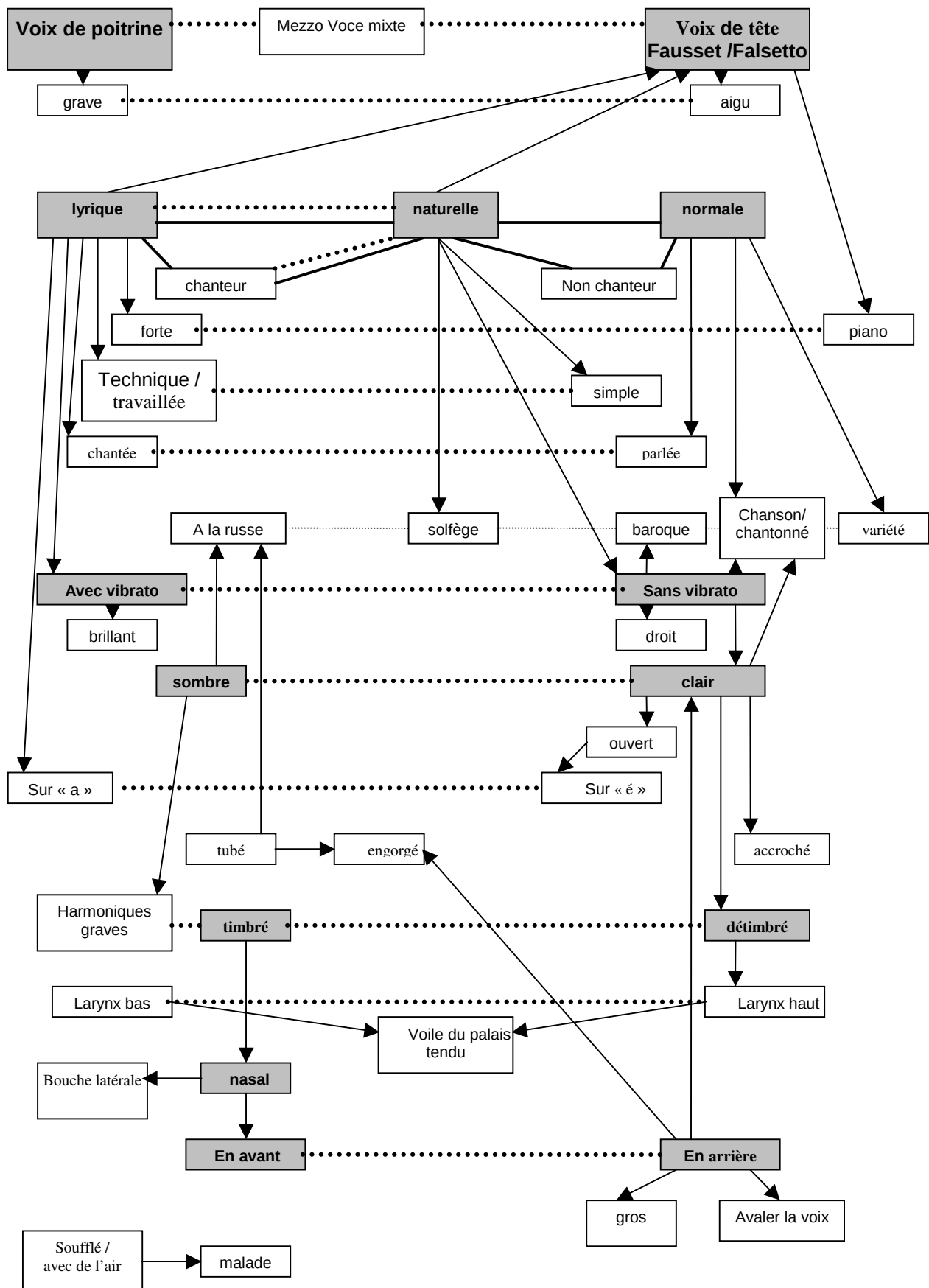


Figure 12 : Association des termes dans le discours des chanteurs

Ci dessus, nous avons fait un bref récapitulatif de ces relations entre qualificatifs sous forme d'un diagramme. Les termes en gris se dégagent du corpus par leur fréquence d'apparition et le rôle de pivots pour la compréhension du vocabulaire lyrique. Les flèches traduisent une implication d'une qualité vocale sur la perception d'une autre. Enfin, les lignes horizontales en pointillés représentent les antonymes et les lignes latérales en gras les synonymes.

3.6) Énoncés métalinguistique et modalités

Les énoncés métalinguistiques caractérisent l'attitude du locuteur vis à vis de son propre discours. Dans notre cas, la forte présence de modalités, de recherches de confirmations traduisent le manque de sécurité du chanteur qui se trouve, c'est vrai, dans une situation d'interaction sociale non neutre puisqu'il se retrouve dans une chambre sourde, devant un micro, en « terrain étranger » puisqu'ils ont accepté de se déplacer au laboratoire. Il y a déjà déséquilibre dans la relation de part le fait qu'il se trouve dans une position d'invité par rapport à l'expérimentateur. Se rajoute le fait que certaines personnes sont parfois « impressionnées » par l'aspect scientifique du laboratoire et des personnes qu'ils rencontrent. Ils sont à la fois sûrs d'eux dans leur domaine, mais mal à l'aise vis à vis du milieu scientifique dont ils craignent de ne pas comprendre les questions ou de ne pas répondre aux attentes de l'expérience.

C'est pourquoi cela peut expliquer les nombreuses modalités et en même temps l'importance des termes techniques « normatifs ». Une façon de se rassurer et de garder une contenance en situation de stress peut consister à faire étalage des connaissances dans le domaine où on est reconnu, de façon à rééquilibrer le rapport d'influence. Parler des choses de façon compliquée ou originale peut être considéré comme une marque d'excellence. Cette réaction n'est pas forcément un signe de prétention mais se fait de façon inconsciente pour chacun de nous dans une situation d'interaction déséquilibrée.

Le problème réside plutôt dans le manque de sincérité du discours que l'on peut alors recueillir, car on a alors un risque de ne pas capter le ressenti personnel de la personne mais plutôt un propos normatif, scolaire et attendu.

4) Conclusion : vers les experts

Cette étude préliminaire a été riche en pistes à creuser au cours des entretiens avec les experts de la voix lyrique. Mais elle a également mis en évidence des comportements à éviter ou à adopter vis à vis d'eux pour obtenir des résultats intéressants.

En effet, si on veut justement éviter d'obtenir un discours convenu de la part des experts, il faut avant tout essayer de les mettre le plus à l'aise possible. Dans ce but, il faudra soigner leur confort, être le plus accueillant, chaleureux et attentifs possible à leur égard. Il n'est pas inutile non plus de les rassurer, voire de les « flatter » un peu sur leurs compétences, quitte à jouer au naïf pour les impressionner le moins possible. Par ailleurs il est important de leur réexpliquer de façon détaillée les conditions du test et ce qui les attend, afin d'éviter qu'ils n'angoissent au sujet des questions qui vont leur être posées, en leur répétant bien que nous n'attendons rien de leur part, excepté qu'ils nous fassent partager leur expérience d'écoute.

En outre, on ne peut jamais évacuer complètement les discours normatifs au départ. Il peut être alors judicieux de commencer la rencontre par une pause café afin de faire connaissance avant le test lui-même, et donc d'essayer de « tarir » ce flot de paroles destinées à impressionner avant de commencer l'entretien de recherche.

Au cours de l'entretien, notre propre discours aura de l'influence. Il faut donc éviter à tout prix les impératifs auxquels on préférera des modalités pour introduire les questions (« Pouvez vous me parler de... » au lieu de « Parlez moi de... »). De même, il est à proscrire d'exprimer nos propres opinions, tant pour ne pas biaiser l'enquête que pour ne pas avoir imposé quoi que ce soit à l'interrogé. Il est très important d'amener l'enquête à coopérer.

Par ailleurs, il faut essayer de se fondre dans le terrain étudié le plus possible, ne serait-ce que par le vocabulaire. En cela, l'étude préliminaire a permis de dégager les termes importants de ce domaine, et nous a donné des pistes pour approfondir la compréhension des relations entre termes ainsi que pour trouver des échelles bipolaires selon lesquelles nous avons demandé aux experts d'évaluer des voix.

Ainsi, tout un groupe de termes semblent très liés : timbré, clair, sombre, ouvert, couvert, brillant, rond...etc. Il serait très intéressant de parvenir à les clarifier, à dégager si certains sont synonymes ou contraires, ou bien si l'un des termes implique un autre.

De plus, d'autres termes comme le vibrato, le souffle, la nasalité semblent être caractéristiques de la qualité vocale. Ce peuvent être des échelles à envisager.

De plus, nous avons vu qu'il existait plusieurs conceptions de la qualité vocale, différents objets cognitifs associés à la voix, ainsi que plusieurs autres concepts entrant en jeu dans la perception de la qualité vocale. Il serait intéressant de voir s'il en est de même pour les experts, qui n'ont plus le même rapport double de production et de perception de la voix

Perception de la qualité vocale et analyse du discours des experts

1) Description du test d'écoute

Après avoir réfléchi aux conclusions et aux pistes qu'apportaient l'étude préliminaire du discours des chanteurs, un premier protocole de test d'écoute avec des experts de la voix lyrique a donc été imaginé.

1.1) Sélection et diffusion des extraits sonores :

Sur les deux airs différents de la base de donnée enregistrée (L'ave Maria de Gounod et l'arpège du petit Rossini), nous avons tout d'abord choisi un air commun à tous les exemples sonores que nous allions diffuser aux experts, afin de pouvoir comparer la qualité vocale indépendamment de variations de mélodies. Nous avons choisi l'Ave maria pour deux raisons : c'est sur cet air là que les chanteurs avaient donné le plus d'exemples et de verbalisations différentes, ce qui nous permettait d'avoir un plus grand choix et de pouvoir sélectionner au moins 5 modifications vocales très différentes par chanteur, de bonne qualité de surcroît. De plus, l'ave Maria contient plusieurs voyelles contrairement à l'arpège du petit rossini, ce qui permet de faire des analyses acoustiques à différents instants de la phrase musicale, sur des voyelles différentes. Cela permet également de voir si certaines qualités vocales sont plus caractéristiques sur une voyelles qu'une autre.

Nous avons également décidé de restreindre l'échantillon des exemples vocaux aux voix d'hommes parmi lesquelles nous avons sélectionné 3 chanteurs(2 basses et 1 ténor), avec leur émission naturelle et 5 essais de changement de qualité vocale pour chacun. En effet, cela se conçoit assez bien qu'il est plus difficile de différencier subtilement le timbre de deux voix si elles proviennent d'individus de sexes opposés. Tout ce que nous risquerions, c'est que les sujets décrivent justement la différence de timbre entre les deux voix en terme de genre, alors que ce sont des détails plus acoustiques et timbraux qui nous intéressent. De plus, méthodologiquement, ce serait une erreur de ne pas prendre la précaution de séparer la perception des voix d'homme et de femmes, tant que nous n'avons pas prouvé qu'elle est semblable. En effet, lorsque nous percevons des couleurs par exemple, la notion de clair n'a pas le même sens pour un rouge ou pour un bleu. Ces deux couleurs existent bien dans la nuance clair, certes, mais la clarté est-elle perçue avec la même intensité ? Dans le doute qu'il en soit de même avec la clarté d'une voix de féminine comparée à la clarté d'une voix masculine, nous avons préféré limiter l'étude aux voix d'hommes pour commencer.

Afin de se borner vraiment à la comparaison de la qualité vocale sans être influencé par d'autres paramètres, nous avons normalisé les extraits entre chanteurs, c'est-à-dire que nous avons fixé une amplitude égale pour toutes les émissions normales des 3 chanteurs puis nous avons réadapté proportionnellement les autres extraits du même chanteur par rapport à sa voix normale. Nous n'avons évidemment pas normalisé tous les exemples vocaux d'un même chanteur à la même intensité car cela n'aurait pas de sens musicalement et nous nous doutons d'après le sens commun et l'étude préalable que des modifications de qualité vocale ont un lien parfois avec l'intensité du son produit.

Nous avons également rehaussé les fréquences fondamentales de certains extraits de façon à ce que tous soient à la même hauteur.

Pour le premier chanteur (B3), nous avons retenu les essais qu'ils qualifiait de Derrière gros / Serré / Variété / Nasal / Soufflé, avec de l'air.

En ce qui concerne le deuxième (T2) les extraits sélectionnés ont été Accroché / Chanson / A la russe / Nasal / Clair.

Pour finir, nous avons retenu pour le troisième chanteur (B6) les exemples vocaux correspondant aux verbalisations: Bouche en hauteur / Ouverture de bouche latérale / Voile du palais tendu,

Il a été choisi de commencer le test par un entretien ouvert non directif sur l'écoute de cette quinzaine de voix sélectionnées, tout d'abord en qualifiant les émissions « normales » des 3 chanteurs dans l'absolu (c'est-à-dire par rapport à une référence mentale personnelle), puis en comparant pour chaque chanteur leur émission normale à leurs 5 essais de changement de qualité vocale, ce qui représentait donc 15 couples de voix. Pour finir, nous avons mis au point une dizaine d'échelles à

deux termes selon lesquelles nous demandions au sujet d'évaluer les voix modifiées vocalement (c'est-à-dire la deuxième des 15 couples précédents).

En annexe, on pourra trouver le CD des extraits sonores ainsi que la liste de diffusion avec la correspondance entre les numéros de piste du CD et la numérotation des voix pour leur évaluation à l'aide des échelles, ainsi que la numérotation des voix lors de l'analyse acoustique.

1.2) Plan d'entretien

1.2.1) Accueil du sujet au laboratoire

Comme nous en avons parlé précédemment en conclusion de l'étude préliminaire sur le discours des chanteurs, cela peut être une bonne solution de commencer l'entretien par une pause café informelle. Cela a tout d'abord comme objectif désintéressé d'être accueillant et de mettre le sujet à l'aise. Mais cela peut également permettre d'évacuer avant le commencement du test le discours pédant de l'expert qui cherche à impressionner et rééquilibrer la relation d'influence jusque là en sa défaveur.

1.2.2) Explications du contexte du test d'écoute au sujet

Ensuite, il est indispensable de réexpliquer au sujet les conditions du test, c'est-à-dire à la fois le résumé du sujet de recherche, sans évidemment leur expliquer ce que nous recherchons exactement par les questions que nous lui posons, car cela n'aurait alors plus aucun intérêt étant donné que le sujet répondrait ce que nous avons envie d'entendre. Il est également important d'expliquer à l'expert quel type de test il va subir (entretien ouvert, échelles), sans oublier de lui préciser que nous n'attendons absolument rien de sa part, si ce n'est qu'il nous fasse part de son expertise en matière d'écoute et de qualification de la voix chantée.

1.2.3) Entretien ouvert dirigé sur l'écoute des voix

Pour la partie du test concernant la verbalisation libre à l'écoute des voix normales des chanteurs puis des couples de voix normales / timbre modifié, nous avons opté pour un entretien ouvert dirigé. Concrètement, cela signifie que nous devons essayer le plus possible de ne pas influencer le sujet dans le vocabulaire qu'il utilise ou dans sa conception des choses, mais que nous pouvons dresser un cadre afin de l'orienter à nous parler de certaines choses en particulier.

Notamment, nous avons remarqué dans l'étude préliminaire que les chanteurs n'avaient pas tous la même compréhension du concept de qualité vocale. Cependant, ils étaient parfois influencés par les interventions du chercheur qui les enregistrait et donc dans cette nouvelle étude, nous avons voulu examiner l'image que les experts pouvaient avoir de ce terme. Ainsi, nous avons adopté une approche du problème par l'intermédiaire de « questions-entonnoirs ». Plus explicitement, cela consistait à ne demander d'abord que la qualification du chanteur, puis si le sujet ne parlait toujours pas de la qualité vocale, de lui demander de décrire plus particulièrement la voix du chanteur, et en fin en dernier recours de lui demander explicitement de parler de la qualité vocale qu'il percevait. Si alors le sujet posait une question quant à la définition du terme « qualité vocale », il lui aurait été retourné une question semblable pour le pousser à expliciter sa propre conception.

Nous avons vu dans l'étude préliminaire que la qualité vocale était liée à 4 types de représentations mentales : le son, la technique vocale, l'hédonisme ou le style.

Dans la première version du protocole, « l'entonnoir » se poursuivait alors ensuite sur des questions correspondant aux 3 autres types de conceptions de la qualité vocale que celle qu'avait l'expert. Par exemple, si spontanément, il nous répondait en terme de technique vocale à une question sur la qualité vocale, nous lui posions ensuite des questions sur sa perception du son produit, du style perçu et sur l'agrément et la valeur de l'émission vocale écoutée.

Ces mêmes questions ont donc été posées pour qualifier les émissions normales des 3 chanteurs, puis pour les comparer à leur 5 essais de changement de qualité.

Cependant, il faut être attentif à ne pas trop poser de question. Car le risque est alors que le sujet deviennent passif vis à vis du discours et perde en proximité.

1.2.4) Evaluation d'une quinzaine de voix par rapport à des échelles bipolaires

D'après les termes pivots ressortant spontanément du discours des chanteurs, nous avons mis au point 5 échelles horizontales bipolaires (c'est-à-dire des échelles constituées de 2 termes aux extrêmes, considérés comme antagonistes).

clair	sombre
sans vibrato	avec beaucoup de vibrato
timbré	détimbré
Peu de souffle	Beaucoup de souffle
Sourd	Brillant

Nous avons choisi des échelles avec un nombre impair de points, ce qui permet de cocher au milieu lorsque l'équilibre entre les deux composante semble être atteint, ou bien lorsqu'aucun des deux termes aux extrêmes ne caractérise la voix entendue pour le sujet. Une échelle à 5 points nous a donc semblée raisonnable (Très A, un peu A, Equilibre, un peu B, très B).

Nous avons donc demandé au sujet d'évaluer les 5 changements de qualité vocale pour les 3 chanteurs (donc 15 voix) selon ces échelles.

1.2.5) Critiques et améliorations du test

Pour finir le test d'écoute, il est usuel de demander au sujet s'il a des critiques à formuler vis à vis du test ou des améliorations à suggérer. Cela permet effectivement de faire progresser notre protocole, mais aussi de le faire partir sur une bonne impression, celle d'être écouté et pris en compte par les expérimentateurs. De plus, si le test s'est mal passé, cela lui permet d'extérioriser sa gêne.

2) Réalisation

2.1) Recrutement des sujets

2.1.1) Critères

Ce qui nous importait était que les experts aient eu un cursus d'apprentissage du chant lyrique et qu'eux même enseigne le chant classique depuis au moins 4-5 ans. Le fait que ce métier soit leur occupation principale pouvait être un plus car on peut penser qu'ils aient alors plus d'expérience et d'implication dans le questionnement de la qualité vocale.

2.1.2) Moyens de communication

Afin de faire passer le test à des experts, il a fallu réaliser un petit travail de communication pour contacter des professeurs de chant classique et de les convaincre de bien vouloir venir passer le test bénévolement au laboratoire. Pour cela, nous avons contacté directement certains professeurs de CNR ou du CNSM de Paris et utilisé les listes d'adhérents de l'AFPC (association française des professeurs de chant) ainsi que des listes de diffusion internet.

Nous avons alors contacté une soixantaine de personnes par mail, par courrier et par téléphone. Lors de la prise de contact, nous leur exposons le thème de notre recherche sans leur expliquer les buts visés.

2.1.3) Problèmes liés à leur intérêt pour la physiologie et l'acoustique

Cependant, le bénévolat du test a entraîné le problème que pour accepter de venir au laboratoire passer ce test, les sujets devaient être très intéressés par le sujet et par les problématiques du laboratoire. Nous avons constaté au cours des entretiens que les experts parlaient beaucoup - en plus ou moins bonne connaissance des choses - de physiologie, d'acoustique et de recherches antérieures menées sur la voix chantée. Il nous paraît douteux que les professeurs de chant soient tous aussi cultivés dans ces domaines et qu'ils parlent de cela à leurs élèves.

2.2) Pré-test

2.2.1) Problème de durée.

Pour tester la validité de notre protocole, nous avons fait passer un pré-test à un professeur de chant. Un gros problème de durée s'est imposé, indépendamment du fait que cela représentait mon premier entretien, peut-être plus maladroit que les suivants. Nous avons donc décidé de maintenir l'entretien, le nombre d'extraits et leur ordre de diffusion dans le protocole. Cependant, dans les 4 types de conceptions de la qualité vocale, nous avons restreint l'entretien et les questions entonnoirs à l'écoute du son produit, qui nous intéressait en priorité afin de trouver des corrélats acoustiques à certaines qualités vocales. Du moins, nous laissons les sujets nous parler en priorité de leur propre idée de la qualité vocale, sans les influencer, puis nous l'orientons ensuite sur l'écoute du son si cela n'avait pas été fait.

2.2.2) Mise au point de nouvelles échelles.

Si l'entretien s'était relativement bien passé, la phase d'évaluation au moyen des échelles ne convenait absolument pas car les termes choisis n'étaient manifestement pas du tout clairs pour le sujet, car comme on l'avait vu lors de l'étude préliminaire, tous les termes étaient assez synonymes ou très liés. Il semblait alors que certains couples de qualificatifs d'une échelle ne soient même pas vraiment perçus comme contraires ou qu'il y ait redondance entre deux échelles. Avec l'aide de cet expert, nous avons donc remis au point de nouvelles échelles plus appropriées. Celui-ci nous a donné également d'autres caractéristiques pertinentes de la qualité vocale comme l'attaque par exemple (glottique / soufflée). Nous n'avons pas utilisé toutes ses suggestions car nous avons souhaité ne pas perdre de vue le but des échelles, c'est-à-dire la recherche de corrélats acoustiques aux variations évaluées de certaines composantes du son. De plus, il était déjà assez lourd pour le sujet de devoir évaluer 15 voix selon 8 échelles. Nous avons donc restreint les échelles aux caractéristiques les plus pertinentes auditivement et les mieux comprises. Voici donc ci-dessous les nouvelles échelles mises au point :

<u>Voix 1</u>						
						Avec beaucoup de vibrato

Figure 13 : Nouvelles échelles mises au point pour les tests d'écoute

2.3) Tests

2.3.1) Conditions

Nous avons donc fait passé ce nouveau protocole de test d'écoute à 10 experts, de début Mai à fin Juin, à cause des grèves de transport qui nous ont obligés à remettre sans arrêt les rendez-vous.

Les sujets étaient donc invités à venir bénévolement au laboratoire. Les tests se sont déroulés dans une pièce du laboratoire dédiée à l'écoute, où les enceintes acoustiques sont de très bonne qualité, mais pour le confort et la convivialité du test, nous avons jugé préférable de ne pas effectuer les tests dans une chambre sourde.

Le temps prévu pour l'entretien était d'environ 1 heure. Dans la réalité, les premiers entretiens duraient plutôt entre une heure et demie et deux heures, malgré le raccourcissement du protocole effectué à l'issue du pré-test, notamment par manque d'expérience de ma part pour arriver à interrompre gentiment les sujets lorsqu'ils partaient dans une direction inexploitable pour notre étude. Sur la fin, le test arrivait effectivement à tenir dans le « temps imparti ».

Pour enregistrer l'entretien, nous avons utilisé un DAT et un microphone placé en retrait, de façon à ne pas impressionner le sujet.

L'étude linguistique étant faisable seulement sur un support textuel, il a fallu transcrire les enregistrements effectués, ce qui n'a pas été réalisable pour tous les sujets étant donnée le peu de temps dont nous disposions et le caractère fastidieux de cette tâche (pour information, ½ heure de requiert une journée de travail pour être transcrit). En annexe, on trouvera la transcription du pré-test ainsi que celle de 5 autres sujets.

2.3.2) Premières remarques générales

Quelques éléments du discours se sont immédiatement dégagés des entretiens, et différaient de ce qu'on avait pu remarquer dans le corpus des chanteurs.

Tout d'abord, nous nous attendions à recueillir en premier lieu une qualification du chanteur ou sa catégorisation par sexe et par pupitre. (« C'est une voix d'homme, plutôt un baryton » par exemple).

Cela n'a absolument pas été le cas. Au lieu de cela, les premières remarques étaient souvent dirigées vers des jugements sur l'efficacité de la position vocale et corporelle du chanteur, sur la santé de sa voix.

De plus, vis à vis de la qualité vocale, leur attitude n'était pas tant de la qualifier que d'essayer de la modifier. Leur réflexe de prof de chant semblait prendre le dessus - à moins qu'ils n'aient pensé que c'était le discours que nous attendions d'eux - et ils imaginaient immédiatement des exercices ou des tests à faire exécuter par le chanteur pour corriger une qualité qui ne leur convenait pas ou pour confirmer une impression qu'ils avaient du chanteur. En cela, il semble que la qualité vocale pour un expert ne soit pas vraiment celle qu'il perçoit sur un extrait du chanteur, mais l'ensemble de ce qu'il perçoit et de ce qu'il imagine sur la voix d'un chanteur. Plus clairement, lorsqu'on leur demande de décrire nous parler de la qualité vocale qu'ils perçoivent, nous nous serions attendus à ce qu'ils nous parent par exemple de l'air qu'ils entendaient dans la voix ou de la mucosité entre telle et telle syllabe. Au lieu de cela, ils nous parlaient de l'air ou de la mucosité, non pas pour caractériser la qualité vocale, mais parce qu'ils étaient les indices d'une mauvaise position du chanteur, d'une mauvaise gestion du souffle ou parce que cela montre que le chanteur ne pourrait pas faire un piano dans l'aigu, par exemple. En résumé, la qualité vocale ne semblait généralement pas correspondre à ce qu'ils entendaient mais à toute la technique vocale maîtrisée ou non par le chanteur. En cela, on peut se demander si cela était une bonne idée de recruter des professeurs de chant comme experts de la voix lyrique. Cette attitude vis à vis de la voix est corroborée par l'usage abondant de termes techniques ou physiologiques.

De plus, on remarque que les experts semblent manquer de mots pour décrire une voix et la meilleure façon pour eux de faire comprendre ce qu'ils perçoivent consiste à imiter : ils finissent tous à un moment ou un autre à rechanter l'air pour préciser leur pensée, ou bien utilisent beaucoup d'onomatopées. C'est donc un problème pour nous qui cherchions à étudier leur verbalisation, car dans ce cas, la perception ne semble pas passer par le langage. L'étude des modalités nous permettra de trancher entre deux explications de cela : est-ce un manque de certitude ou de vocabulaire de la part de l'expert, ou bien la perception d'une voix s'effectue t'elle finalement par la production, par le ressenti de l'imitation ?

Une constatation nous a également surpris : bien que nous leur précisions que lors de l'écoute des couples de voix, la première était toujours la même, c'est-à-dire l'émission normale du chanteur, les experts ne la reconnaissait pas et la qualifiait de façon différente à chaque écoute. Il était prévisible que la perception est une expérience subjective qui dépend de la personne et du contexte, de même que l'ordre de diffusion des extraits était important mais nous ne attendions quand même pas à cela. En fait, il semblait que le sujet renvoyait sa propre acclimatation au test sur la voix du chanteur. En effet, l'ordre de diffusion des extraits a été décidé afin que les variations d'amplitude entre deux modifications de qualité consécutives ne soient pas trop grandes, de façon à ce que la comparaison ne se fasse pas immédiatement sur ce critère d'intensité sonore. Les extraits n'étaient donc pas dans l'ordre chronologique de production du chanteur. Pourtant, les experts, imaginaient que les chanteurs devenaient plus fatigués ou s'accommodaient aux conditions d'enregistrement, ce qui n'était pas possible vu la compilation des extraits dans un ordre quasi-aléatoire.

2.3.3) Problèmes du test inchangeables à ce stade de l'étude.

Lors de la phase de critique à la fin du test, et au cours de l'entretien, plusieurs problèmes se sont posés vis à vis du protocole, mais contrairement au pré-test, ce n'étaient plus là des détails modifiables aussi facilement. En particulier, les enregistrements étaient en cause au niveau de la qualité des chanteurs. Les experts leurs reprochaient en effet de n'être pas d'excellent niveau, et donc de ne pas pouvoir bien évaluer leur qualité vocale. De plus, l'enregistrement en acoustique sèche semble avoir des répercussion sur l'effort vocal des chanteurs ont les experts perçoivent un forçage. D'un autre coté, ces remarques sont troublantes et ont été prises en compte car elles ont été exprimées par plusieurs experts. En même temps, on peut trouver d'autres explications à cette remarque, comme par exemple qu'ils se cherchent des excuses au manque de mots qu'ils peuvent avoir pour décrire une voix. Car au

contraire, si l'on compare les 3 chanteurs, c'est finalement le « meilleur » qui faisait les modifications de qualité vocales les plus subtiles et par conséquent les moins faciles à qualifier ni à discriminer par des analyses acoustiques. Dans l'absolu, leur remarque sur la qualité des chanteurs va donc à l'encontre du sens commun. N'est pas le signe que leur conception de qualité vocale se situe dans le jugement de valeur et de technique vocale du chanteur ?

Un autre problème se pose quant au choix de la mélodie choisie, l'ave maria, qui ne couvre pas un grand ambitus pour pouvoir juger d'une voix sur ses différents mécanismes (passage de la voix de tête à la voix de poitrine, blocages éventuels). On retrouve ici leur conception de qualité vocale qui ne représente pas le timbre d'un chanteur à un moment précis mais toute sa palette vocale, que ce soit timbral ou technique. En même temps, cela poserait problème au niveau de l'analyse acoustique, car le quotient ouvert par exemple se comporte différemment selon les deux mécanismes vocal. De plus, au niveau spectral, comparer deux voyelles à un ton d'écart n'a pas d'importance. A une octave d'écart, cela en a déjà beaucoup plus.

Ces points étaient à noter et pourraient être reconsidérer dans le cas où l'on voudrait approfondir cette étude de la qualité vocale de façon plus « rigoureuse ». Cependant, dans le cas de l'étude, le temps du stage ne nous permettait pas, même si nous en avions eu envie, de réenregistrer une nouvelle base de données.

3) Résultats descriptifs

3.1) Dépouillement des échelles et pistes pour l'analyse acoustique

En annexe, vous trouverez à la fois les tableaux de résultats des évaluations de 15 voix par les échelles, les figures 14, 16 et 17 représentent ces résultats sous forme de diagrammes et serviront de référence à plusieurs discussions par la suite. Les notations des experts sur ces échelles à 5 points ont été codées de 1 à 5, 1 représentant le terme de gauche de l'échelle et 5 celui de droite.

Histogrammes de représentation des résultats

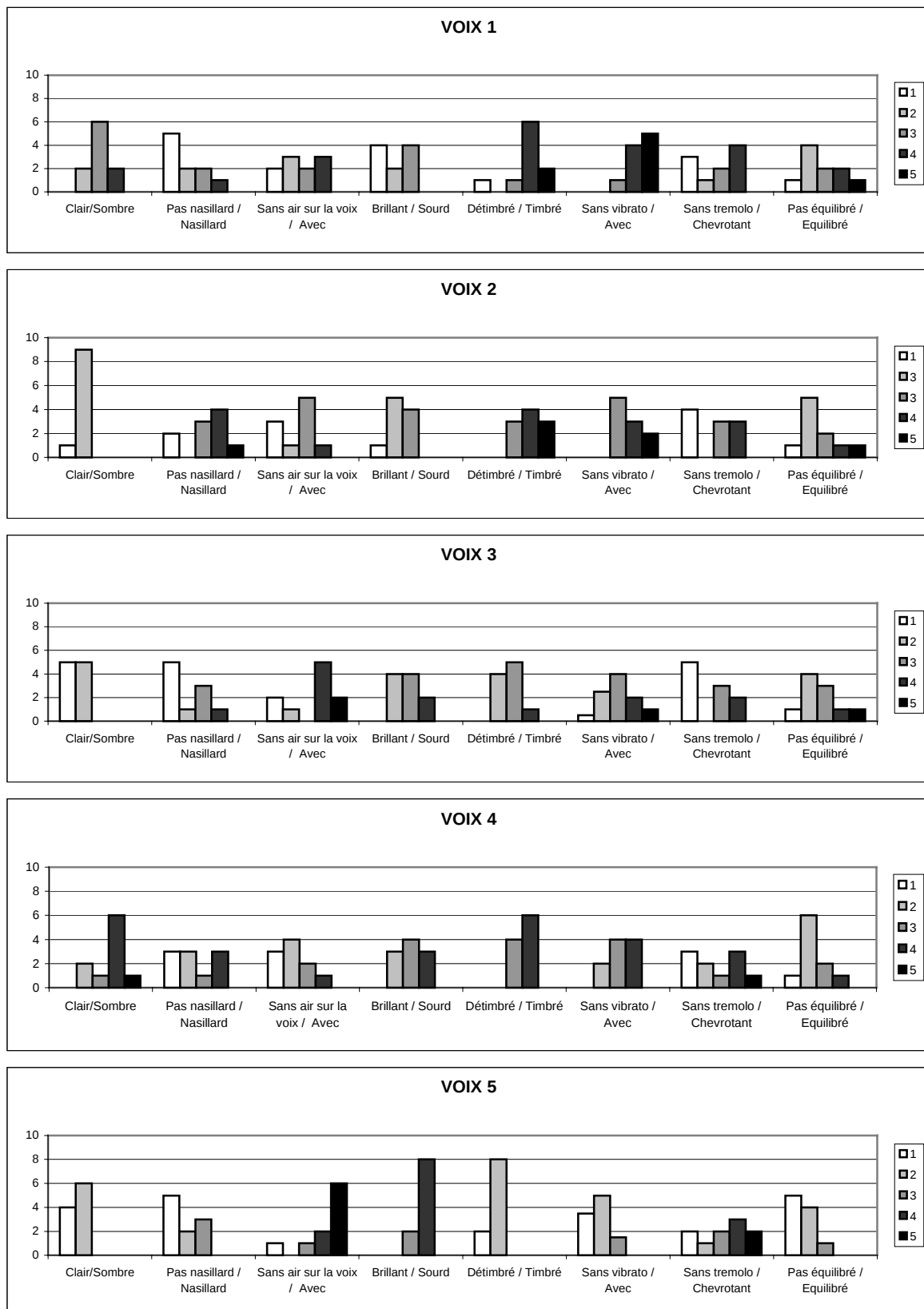


Figure 14 : Chanteur 1 (B3)

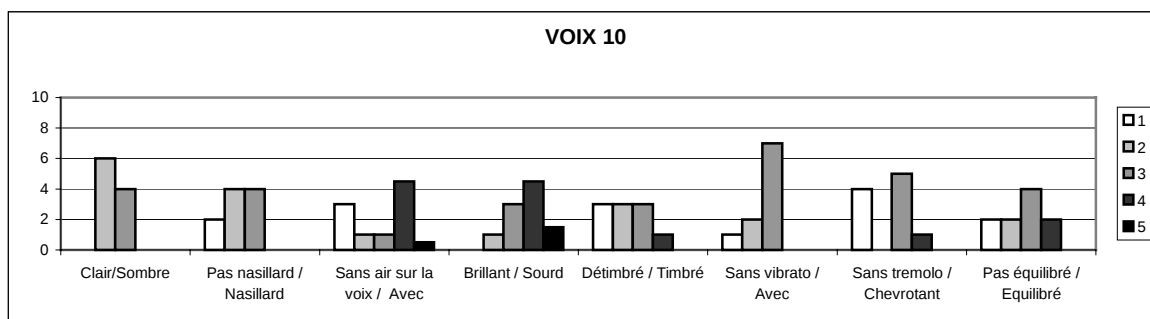
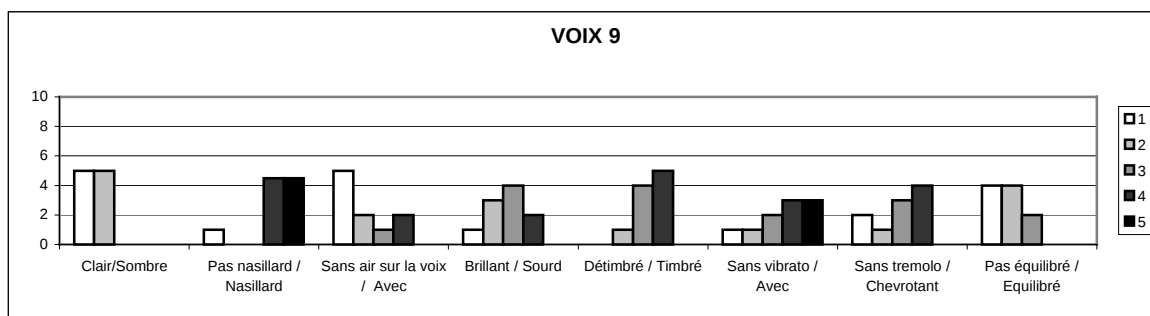
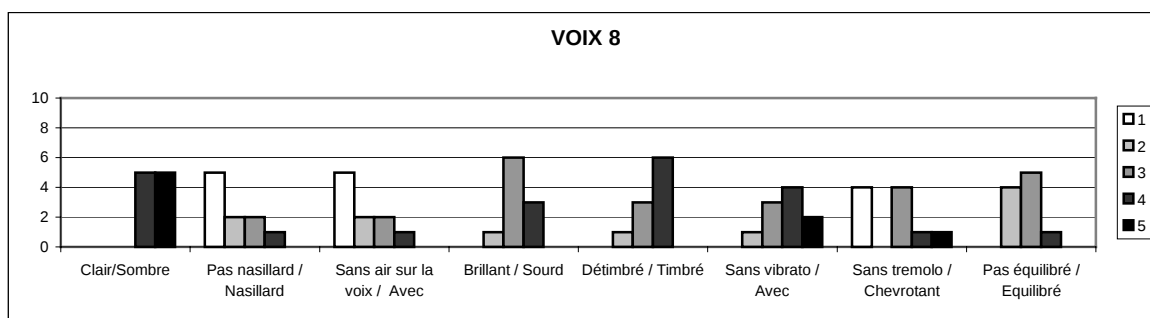
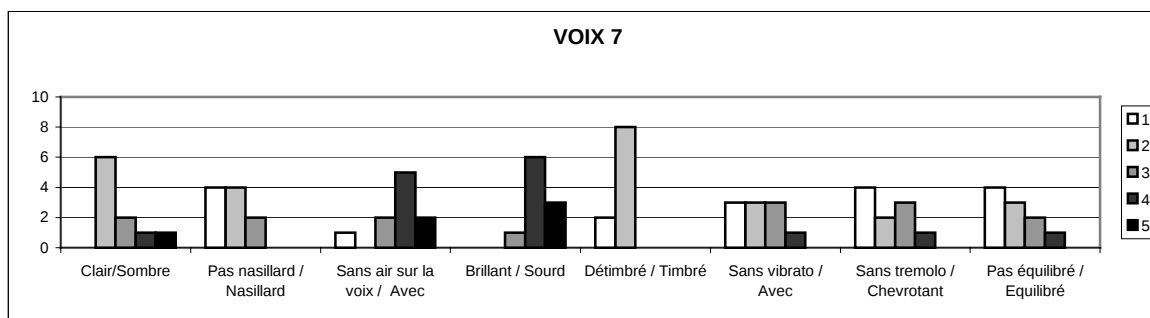
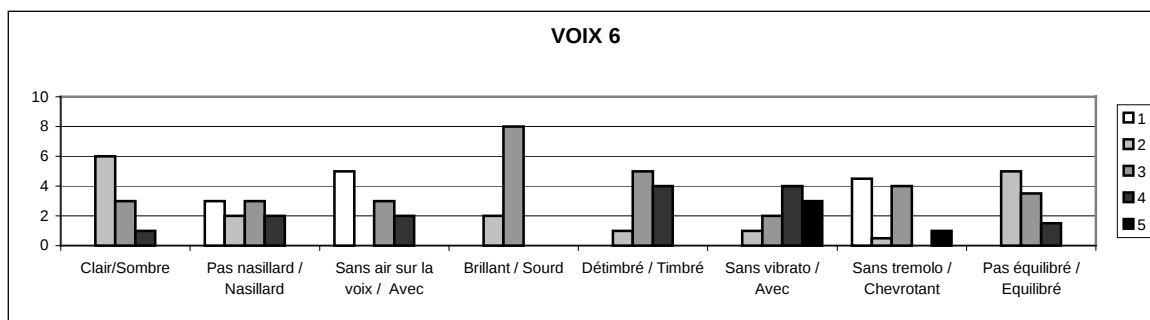


Figure 15 :Chanteur 2 (T2)

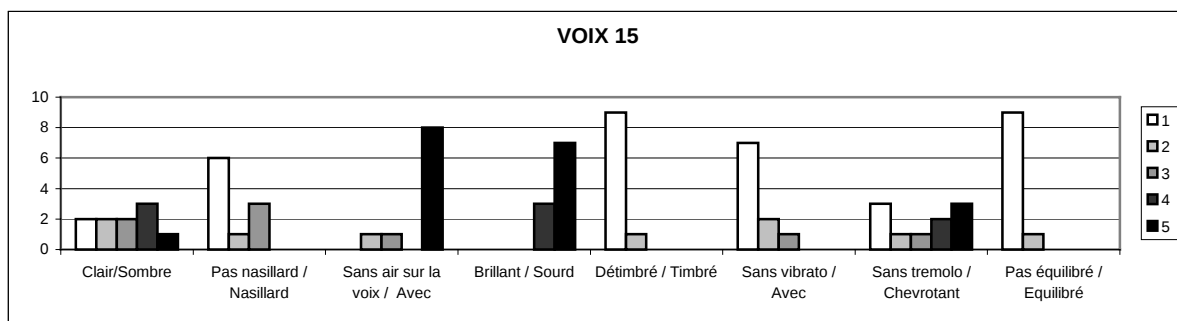
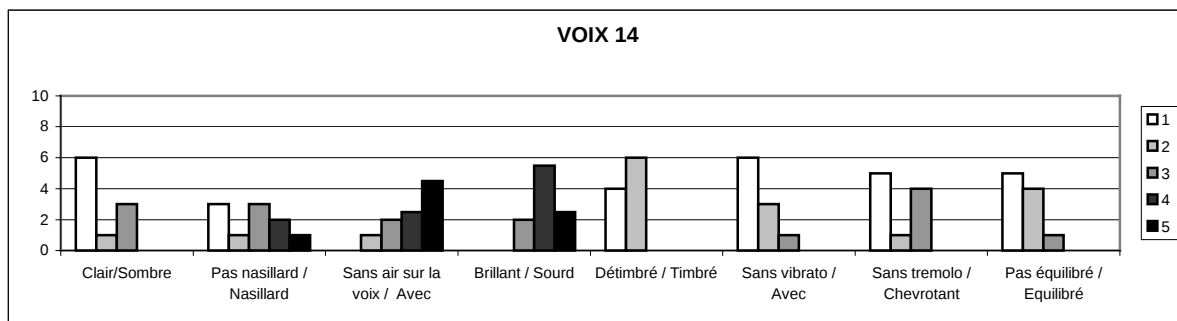
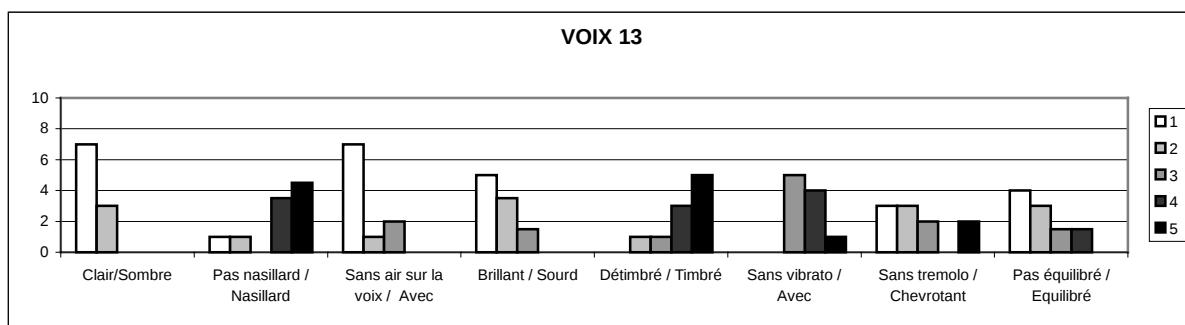
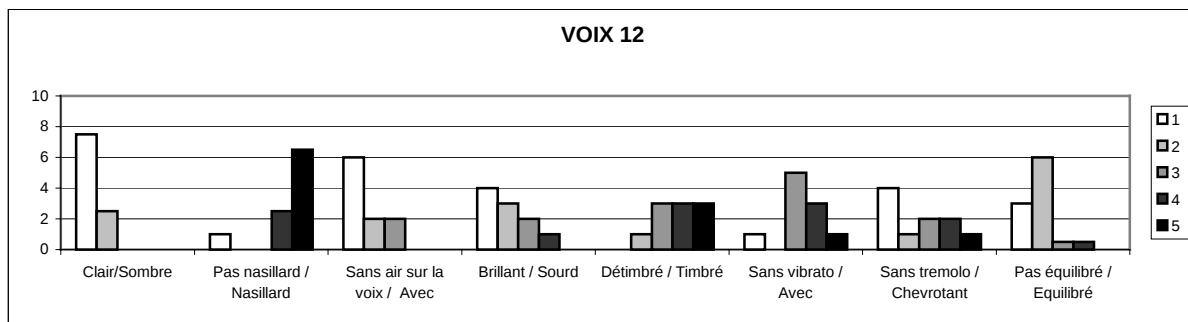
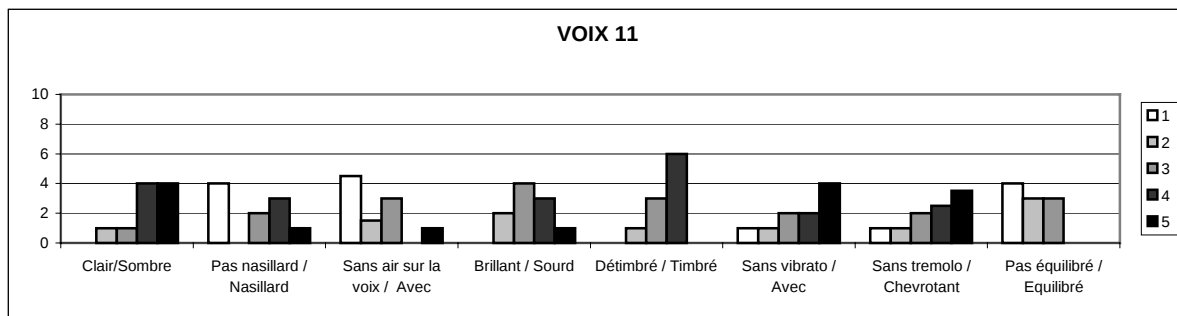


Figure 16 : Chanteur 3 (B6)

Les diagrammes et le calcul des écart-types des jugements sur chaque échelle des 15 voix différentes a permis de mettre en évidence les échelles les plus consensuelles et celles les moins bien comprises, mais aussi les qualités vocales les plus caractéristiques d'une voix.

Les échelles les plus évidentes ou les plus remarquables sont les critères **clair/sombre**, **détimbré/timbré**, la **brillance**/ le **caractère sourd** d'une voix et le jugement du **vibrato**. Par contre, le trémolo semble avoir plusieurs significations différentes selon les experts. C'est ce qui ressort du discours libre des entretiens, et c'est effectivement confirmé par les évaluations sur ces échelles. De même, la nasalité ou le souffle sur la voix souffrent d'un manque de définition et d'accord entre les sujets, car ces deux notions semblent renvoyer par moment à des aspects naturels et inhérents à la voix (le souffle comme matière première de l'onde acoustique ou la nasalité comme résonance non exagérée et même au contraire responsable de rondeur dans le son), et d'autres fois à des défauts (voix nasillarde, souffle qui amène le détimgement du son par son manque de gestion). Cette polysémie peut expliquer les mauvais résultats obtenus sur ces échelles, et fait que nous les éliminons au niveau de l'analyse acoustique : Il n'y a déjà aucun consensus au niveau déverbal, cela ne sert donc à rien d'en chercher des corrélats physiques. Mais cela nous apportent quand même des informations à creuser lors d'une prochaine étude, notamment l'approfondissement de ces polysémies.

Ensuite, la moyenne des jugements des experts sur chaque voix et pour les différentes échelles permet d'examiner les corrélations et les « airs de familles » entre les différentes voix d'un chanteur mais aussi toutes voix confondues.

A l'inverse, cela permet de repérer les voix qui diffèrent le plus sur une échelle donnée, car ce sont elles qui pourront nous donner des pistes sur les paramètres acoustiques liés à une qualité vocale perçue.

Ainsi, la voix 1 ne semble caractéristique d'aucune des échelles proposées à l'expert (ne serait pas aussi du fait qu'elle est la première à être évaluée ?). Par contre, les voix 2, 3, 5, 9, 12, 13 et 14 sont assez unanimement perçues comme claires et les voix 4, 8 et 11 comme sombres. De même, les voix 12 et 13 ont été fréquemment qualifiées de brillantes et les voix 5, 7, 14 et 15 de sourdes. Enfin, il existe un accord sur le caractère timbré des voix 4, 8, 11 et 13 ainsi que le détimgement des voix 5, 7, 14 et 15.

A l'écoute, il nous semble très difficile de juger de la différence de qualité vocale entre deux voix issues de chanteurs différents. La corrélation entre deux voix est souvent plus forte à l'intérieur d'un même chanteur. En particulier, les voix 1 et 4 (du chanteur B3), les voix 12, 13 et 14 (c'est-à-dire les voix 2, 3 et 4 du chanteur B6) sont très associées pour la perception de la clarté. Pour la perception de la brillance, les voix 12 et 13 sont également très liées, et pour le critère timbré/détimgement, il existe un rapport entre les voix 11 et 14, et entre les voix 12 et 13.

En vue de l'analyse acoustique, nous commencerons donc par comparer des voix opposées sur un certain critère, au sein d'un même chanteur. Par exemple, les voix 8 et 9 (« à la russe » et « nasal » selon la verbalisation du chanteur), et les voix 11 et 12 (dites « bouche en hauteur » et « bouche latérale ») nous semblent intéressantes à juxtaposer afin d'approcher la notion de clair/sombre.

Le calcul des corrélations entre échelles montre qu'il existe principalement des liens entre la brillance/surdité et le timbre d'une voix, entre le chevrottement ou le vibrato et l'air sur la voix, ainsi qu'entre le vibrato et le chevrottement. Ensuite, on constate que le chevrottement, la surdité, la présence d'air dans la voix et le caractère sombre d'une voix sont des indices du déséquilibre d'une voix. Au contraire, la nasalité et la brillance/surdité ne semble pas être corrélées, de même que le chevrottement et la clarté/sombritude (semble plus prévisible) ainsi que l'équilibre et le timbre d'une voix.

Cependant les corrélations entre échelles sont ici calculées pour l'intégralité de l'échelle bipolaire. Cependant, il peut y avoir corrélation malgré tout entre certaines branches « monopolaires » de ces échelles, au cas où nous nous serions trompés en considérant deux termes d'une même échelle comme antonymes. Par exemple, la corrélation des échelles timbré/détimgement, avec vibrato/sans vibrato et avec du souffle/sans souffle semble n'apporter aucun résultat. Pourtant, dans le discours libre des experts, il apparaît que la perception de timbre est très liée à la présence de vibrato. A l'inverse, le détimgement, n'est pas lié à l'absence de vibrato mais plutôt à la présence de souffle, de même que le timbre n'est pas associé à l'absence de souffle. Il apparaît donc que timbré/détimgement, bien qu'apparemment

antagonistes ne serait ce que par leur construction lexicale, ne serait pas si contraire que ça du point de vue cognitif.

Cela est un exemple pour rappeler que dans une étude perceptive, il ne est facile de se laisser happer par les résultats statistiques mais c'est pourquoi il est important de toujours baser sa réflexion sur l'écoute du son lui-même.

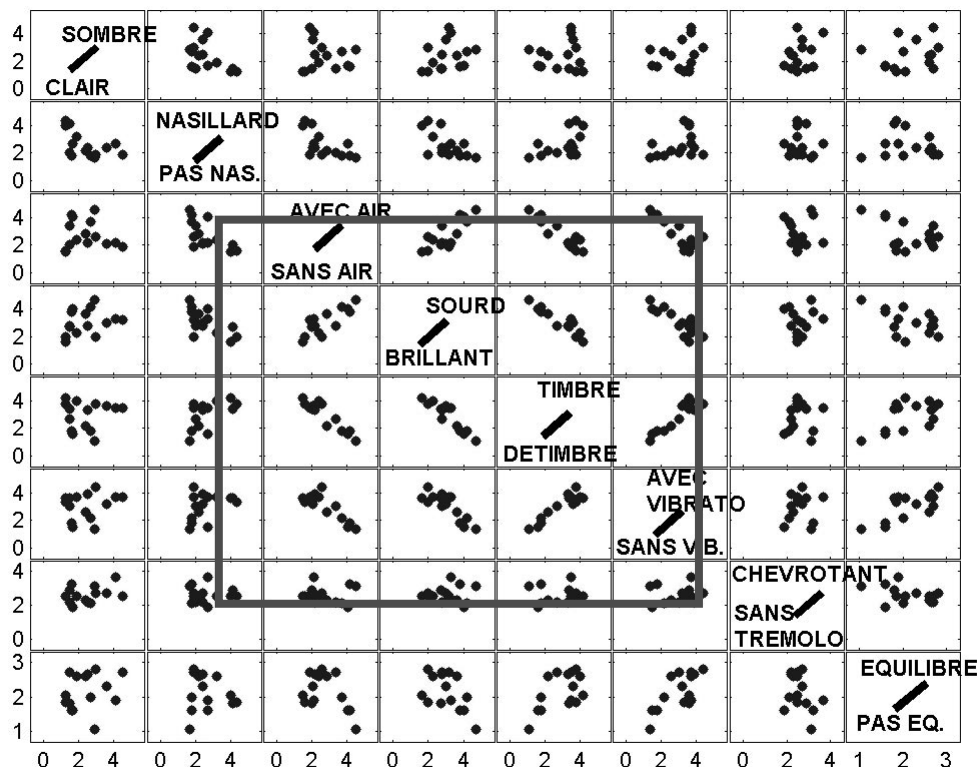


Figure 17 : Corrélations entre les différents termes des échelles. Dans le carré rouge, les corrélations fortes sont soulignées.

La figure 17 illustre graphiquement les liens qui peuvent exister entre différents termes utilisés pour décrire la qualité vocale. Sur cet exemple, il s'agit des échelles bipolaires suivantes (jugées de 1 à 5): clair-sombre, pas nasillard-nasillard, sans air-avec air, brillant-sourd, détimbré-timbré, sans vibrato – avec vibrato, sans tremolo – chevrotant, pas équilibré – équilibré. Les fortes corrélations concernent l'air, la brillance, le timbre et le vibrato, qui sont présentés à l'intérieur du carré rouge. On observe que l'axe sans air – avec air est positivement corrélé avec l'axe brillant-sourd et négativement corrélé avec l'axe détimbré-timbré et sans vibrato – avec vibrato. Cela signifie qu'une voix qui sera jugée « sans air » sera également généralement jugée brillante, timbré et avec du vibrato. Une voix « avec de l'air » sera plutôt jugée également sourde, détimbrée, sans vibrato.

3.2) Analyse des termes des échelles dans le corpus des experts

À travers ces six corpus, nous voyons les experts se livrer à toute une série de caractérisation des voix et de mise en équivalence ou en opposition d'adjectifs censés rendre compte de leur expérience acoustique / musicale. La personne qui les interroge leur propose 8 couples de mots de sens opposé, qu'un linguiste appellera des « antonymes ».

3.2.1) Analyse du souffle

Le « souffle » et l'« air » sont très souvent utilisés dans un contexte de **gestion**, de consommation, que ce soit dans le sens négatif de gaspillage, de perte, de fuites (« air qui passe », « air qui part »), ou dans une conception d'alimentation.

Un autre situation d'employer le terme « souffle » semble liée à l'idée de **contrôle** : « Placer, maîtriser, connecter, libérer, tenir, appuyer *le* souffle » ou « commander, connecter *avec le* souffle ».

On peut remarquer que ces mêmes verbes de contrôle sont également utilisés pour parler du son (« relâcher, pousser, connecter le son, organiser le départ du son »), de la voix (« pousser, soutenir la voix ») ou de l'appareil vocal (appui, relâchement du diaphragme, tenue du larynx). À noter que le verbe « connecter » est utilisé justement pour mettre en relation le souffle et le son.

Enfin, une troisième idée récurrente associée au souffle fait appel à l'**effort**, au travail et à l'efficacité de l'émission, tant pour la fatigue ou le danger qu'elle représente, que pour décrire l'ampleur et la portée d'une voix. On peut éventuellement rajouter à ce groupe l'utilisation du mot « souffle » dans une idée de **Régularité** et de détente corporelle (Souffle constant, irrégulier, secousse de souffle).

On constate également que le souffle est lié très souvent à l'**attaque**, au début du son. Pour parler des sons tenus, ils utilisent plutôt la notion de vibrato ou de vibration. D'ailleurs, souffle et **vibrato** semblent très liés. On peut émettre l'hypothèse que le vibrato serait perçu comme une sorte de souffle (dans le sens naturel et positif) sur les tenues. D'ailleurs, les chanteurs ne semblent pas pouvoir tant contrôler leur vibrato que ça. Il disparaît avec l'intention de piano ou de détimbrage, et est décrit comme quelque chose de naturel qui ne s'assiste pas volontairement, qui serait lié à une légère suppression de souffle, mais se placerait correctement lorsque la gestion du souffle est acquise.

D'ailleurs, un autre lien entre le vibrato et le souffle apparaît lors de la qualification de l'attaque « soufflée » ou « de façon « droite » (donc plutôt sans vibrato) qui s'opposent toutes les deux à une attaque « glottique » dans le discours des experts. On peut donc considérer que ces deux notions sont de nature proche.

Le souffle est très lié à la **pression** (« Pression de souffle, suppression, pression sous-glottique, poussée du souffle), celle-ci étant qualifiée en termes quantitatifs ou relatifs à la santé vocale (suffisante, minimaliste, insuffisante, néfaste), et manifestement responsable en partie de la notion d'**ouverture**. De plus, la suppression du souffle obligerait à serrer plus fort les cordes vocales, créant un son qualifié de **cravaté, serré, pointu**. Cet appui (du souffle, ou du larynx ou autre, nous nous posons la question) empêcherait l'air de « tourbillonner », créant les harmoniques aiguës perçues dans ce type de qualité vocale.

Dans la série des conceptions pseudo-scientifiques, l'« air » renvoie à des descriptions proches de la mécanique des fluides (« débit d'air, filet d'air, courant d'air, turbulences, tourbillon »). Nous remarquons également que l'air est perçu *dans* la voix ou *dans* le son alors que le souffle est plutôt *sur* la voix, *sur* les cordes vocales, *sur* l'attaque.

Quand au terme « **soufflé** », il réfère à une perception négative du son, contrairement à l'utilisation du terme « souffle » comme quelque chose de naturel. D'ailleurs, « soufflé » est utilisé pour qualifier principalement des attaques, tandis que le souffle sur la voix s'applique plutôt à des notes tenues. Bien qu'ici de la même racine lexicale, ces deux termes semblent donc renvoyer à des notions différentes. D'ailleurs, le souffle n'est pas toujours indésirable, il peut donner de la « joliesse » à une voix, notamment dans la « chanson de charme », mais est également utilisée comme effet dramatique dans la « Voce Finta » ou comme technique vocale en studio d'enregistrement, notamment pour des lieders.

Dans ce paragraphe, nous aimerions aborder des pistes d'influence du souffle sur la qualité vocale perçue. En effet, en présence de souffle sur la voix, le son est bien souvent qualifié de **sourd**, et quand il y en a peu, de brillant : « Plus y'a d'air moins y'a de son » dit une des experte. Du moins, il est semble effectivement se dégager que plus il y a de souffle, moins la le spectre est riche en harmoniques. Ainsi, une **voix blanche**, une voix de débutant et une voix avec du souffle semblent être assez proches dans l'absence de vibrato, de résonance et de **richesse spectrale**.

Le souffle est perçu comme central dans l'apprentissage du chant : c'est par sa gestion que la diction peut bien se placer. Dans le cas contraire, les experts perçoivent des « mots mal placés, des notes sans poids, une voix **dépouillée** »

Quand à la physiologie concernée par la production et la gestion du souffle, la verbalisation des experts évoque le resserrement de la **Gorge**, le **Nez** à sa base ou dans l'idée de « naso profundo », le **Larynx** qui peut être détendu ou au contraire déstabilisé par une trop grosse poussée du souffle, et sur lequel on appuie. De plus, un souffle important est manifestement inconciliable avec une bonne qualité d'accolement des **cordes vocale**. Enfin, au niveau de la « soufflerie », le **Diaphragme** doit travailler régulièrement et subir également une pression, tout comme les **abdominaux**, et la **cage thoracique** doit tenir cette pression.

Pour finir, voici quelques objets gravitant autour de la notion de souffle qu'il serait intéressant avec un peu plus de temps d'explorer : le bâillement, le soupir, l'expire, l'inspire, l'amplification, la pression.

3.2.2) Analyse du vibrato

Le vibrato est un objet très complexe gravitant autour de la voix. En fait, il semble lié au souffle, comme nous l'avons vu dans la partie précédente. Les experts insistent unanimement sur l'importance du fait que le vibrato est **naturel** et non pas assisté comme le font certains chanteurs en contractant les abdominaux ou par mouvement volontaire du larynx. Au contraire, le vibrato apparaît de l'équilibre physiologique du chanteur et de la maîtrise de la gestion du souffle ainsi que de la recherche d'un position de stabilité du larynx. Lorsque le vibrato est lié à une pression d'air mal contrôlée ou justement assistée, c'est-à-dire à une poussée du son puis son relâchement, les experts perçoivent ce qu'ils appellent un **trémolo** ou **chevrotement**, qui se mélange au vibrato (et non seulement du trémolo).

Quand le relâchement intervient en plus dans quelque chose de déjà mal structuré au niveau de l'équilibre physiologique, cela devient même du flattement.

Lorsque le vibrato est lent et audible, les experts appellent cela des **oscillations** et vont jusqu'à compter le nombre de battements par seconde (5.5 par seconde ou 7, selon la tessiture du chanteur et la hauteur à laquelle il émet un son)

3.2.3) Analyse des caractères clair et sombre

La clarté semble être très fortement relié à la notion d'**aigu**, d'émission haute et accrochée, à la présence d'harmoniques aiguës dans le spectre (« brillant », « nasal », « ténorisant » et « **timbré** » ne sont pas très éloignés). Dans ce sens, clair semble s'opposer à « bouché » ou « grossi ».

Au contraire, « clair » est également très utilisé lorsque le son est « **détimbré** », « léger », lorsque le son est harmoniquement pauvre, que la voix est blanche.

Cette contradiction apparente des experts sur la conception de la clarté, qui manifestement peut être à la fois liée au timbrage et au détimbrage est peut être explicable par l'association qu'ils font entre les voix de type blanches et les voix d'enfants, généralement aiguës. Auquel cas, leur qualification de clair renverrait toujours d'une certaine façon à l'idée d'aigu.

Dans tous les cas, « clair » est par contre toujours lié à la **légèreté**, par opposition à la lourdeur mais également dans un contexte de baisse d'**intensité** sonore car il est utilisé pour qualifier des mezzo voce, des voix piano, des voix finta.

Par ailleurs, il ressort de la verbalisation des experts un lien assez net entre la clarté, l'**ouverture** et l'**antériorisation des voyelles** et de la prononciation. Ce point sera une de nos pistes au moment de la recherche de corrélats acoustiques.

Pour finir, il faut noter qu'un timbre peut être caractérisé comme clair, à la fois dans l'absolu (empreinte vocale d'une personne) mais aussi pour qualifier une modification vocale par rapport à l'émission normale du chanteur. Cela est intéressant car ce n'est pas à ce point le cas de toutes les qualités vocales.

Le discours recueilli chez les experts valide l'échelle « clair/sombre » choisie pour évaluer les voix dans le sens qu'il semble bien que non seulement ce soient effectivement des descripteurs importants et pertinents de la qualité vocale, mais aussi que l'antinomie des deux termes soit consensuelle.

La qualité « sombre » semble pouvoir être produite par deux « techniques vocales » différentes : La première liée à la **couverture des voyelles** et à la diction, la deuxième effectuée par la **protrusion des lèvres** et un engorgement, à la manière de l'école **russe** ou dans l'intention de caricaturer les chanteurs d'opéra.

De plus, l'impression de voix sombre est influencée par la **gravité** ou la **baisse d'intensité** du son émis.

3.2.4) Analyse de la brillance / surdité

Quant à elle, la perception de surdité semble reliée à l'**efficacité** de l'émission, à la notion de consommation, de gaspillage du souffle. C'est en cela un concept proche du détimbrage, utilisé de façon négative comme synonyme de « mat » ou « terne ». « Feutré » ou « voilé » désignent le même genre de timbre mais de façon positive. Il semble que l'utilisation du sens positif ou négatif dépende seulement du contexte et de l'attente d'esthétique par rapport au style musical en particulier.

Techniquement, la réalisation d'un son sourd est reliée à un placement du son au niveau de la **bouche**, des joues, à une résonance pharyngée. Le son est alors aussi proche voire équivalent d'un timbre « bouché ».

Le couple sourd/brillant paraît assez complexe, car la brillance ne s'oppose à la fois à la surdité mais également à la **rondeur**. Pourtant, rond et sourd ne renvoient pas du tout à la même chose.

A l'inverse, la brillance n'est pas spécialement liée au souffle mais plutôt aux harmoniques aiguës, à la clarté et la ténorisation, ce qui est bien exprimé par des synonymes comme « mordant », « extraverti », « strident ». Il semble se dégager un lien très fort entre la brillance et le **formant du chanteur**, aussi nommé « focus ». C'est d'ailleurs cela qui distingue le timbrage et la brillance, souvent utilisés ensemble. La brillance serait plutôt un timbrage dans l'aigu, alors que l'aspect timbré d'une voix renverrait plutôt à la richesse spectrale dans le médium. Du moins, c'est ce que perçoivent les experts et cela constituera également des pistes à l'analyse acoustique.

Enfin, indépendamment de la tessiture, la brillance et la clarté sont des descripteurs qui influenceraient la catégorisation d'une voix en tant que **Ténor** ou **Baryton**, les ténors ayant un timbre plus brillant et plus clair tandis que les barytons sont perçus plus sombres et plus sourds. A travers les verbes « ténoriser » et « barytoniser », ce sont ces deux qualités qui sont modifiées.

3.2.5) Analyse du caractère nasal / nasillard

La nasalité semble reliée à la présence **d'harmoniques aiguës** dans le spectre, mais non pas seulement à leur renforcement mais plutôt leur sélection ou leur concentration, en renonçant aux harmoniques graves. C'est un concept proche de la ténorisation ou de la clarté.

La nasalité en tant que technique vocale est décrite d'une première manière par l'action sur la **bouche**, soit de façon horizontale, latérale, dans le sourire ou alors en faisant une bouche dédaigneuse, vers l'expression du dégoût.

Une autre technique pour réaliser un son nasillard consiste à agir sur le **nez** en le **pinçant**, le serrant, à placer le son devant, dans le masque, vers les dents, à concentrer le son au dessus du nez, avec une tension à l'arrière de la langue et de l'épiglotte. Un tel son est alors qualifié de « pointu », « tendu », « serré » ou encore « étroit ».

Cette fermeture du nez qui donne un son nasillard (négatif) s'oppose à l'exploitation de la profondeur du nez (« naso profundo ») et à son relâchement qui donne alors un son nasal (positif), plus proche de la **rondeur**. De plus, le pincement du nez s'oppose à l'ouverture de la **gorge** et du rhinopharynx qui peuvent contrer les « intentions » nasillardes. La nasalité est donc l'opposé de la gutturalité (cela oppose notamment le timbre de voix de certaines nationalités comme les latins et les germaniques)

3.2.8) Analyse de l'équilibre

L'équilibre de la voix avait été assez naïvement choisi pour une des échelles, afin principalement d'évacuer le jugement de valeur sur la voix ou son agrément, qui risquait de « polluer » les autres évaluations. En fait, cette notion d'équilibre renvoie à bien des notions du chant lyrique.

En premier lieu, l'équilibre d'une voix désigne l'**homogénéité** du timbre, entre rondeur et brillance (le *chiaro oscuro* de l'école italienne), entre brillance et surdité, enfin au « dosage » **harmoniques et résonantiel**, c'est-à-dire un spectre riche à la fois dans la partie basse du spectre, mais aussi dans la partie haute, au niveau du formant du chanteur, et pas seulement une de ces régions qui serait renforcée au détriment de l'autre.

L'équilibre est également physiologiques. Il renvoie à la **posture générale** du corps, à sa détente, sa tranquillité, et plus particulièrement encore à une position de stabilité à trouver pour le **larynx**. Le déséquilibre est repéré par les experts par une irrégularité du **vibrato** ou une perception de **forçage**, de labeur ou de fatigue vocale chez le chanteur. L'équilibre est donc lié à la santé de la voix et à son efficacité. En effet, un expert décrit particulièrement l'équilibre comme « la balance entre l'énergie nécessaire à faire le son, et le son ».

Cela nous amène à une autre conception de l'équilibre qui est celui du souffle et de la pression sous-glottique. Un équilibre classique semble se situer entre **l'attaque glottique et soufflée**. La gestion du souffle est donc clairement un indice d'équilibre de la voix.

Pour finir, l'équilibre désigne aussi pour certaines experts la boucle audio-phonatoire, c'est-à-dire l'accord entre sensation internes et audition, ou encore la priorité de faire passer la diction avant la voix car c'est la diction qui en serait le support et donc garant de son équilibre.

3.4) Etude des différents objets entrant en jeu pour parler de la qualité vocale

Pour parler de la qualité vocale qu'ils perçoivent, les experts ne qualifient pas forcément la voix ou le timbre, mais bien d'autres objets rentrent en jeu sur le même plan que la voix ou gravitent autour. La voix n'est pas perçue dans sa globalité mais est sous-catégorisée en de nombreux indicateurs.

Tout d'abord, au plus haut niveau, la qualité vocale s'applique à « la version », « l'extrait », « la phrase », « les exemples », « l'Ave maria », « le contexte », « le mot » « le flux », « l'enregistrement ». Les qualificatifs sont alors plutôt hédonistes ou stylistiques.

La qualité vocale est liée également au « chanteur », aux « gens ».

Les qualificatifs utilisés désignent bien sûr « La voix », « le son », « l'attaque », « la voyelle » ou « la note », mais aussi « la phonation », « la diction » et « l'émission » ou directement « la qualité » elle-même ou « Le timbre » (« La fatigue », « la nasalité », « la brillance »...)

A l'intérieur du son lui-même, les experts cherchent à écouter particulièrement « Les harmoniques », « les formants », « le focus », « le formant du chanteur » ou encore « le spectre ».

De plus, deux objets essentiels du chant lyrique se rajoutent à la voix elle-même : « le souffle » et ses dérivés (« l'air », « la pression », « une mucosité ») et le vibrato et ses synonymes (« l'oscillation », « le flattement », « le trémolo », « le battement »).

Voici page suivante un schéma récapitulatif de la « hiérarchisation » de ces objets dans le discours des experts

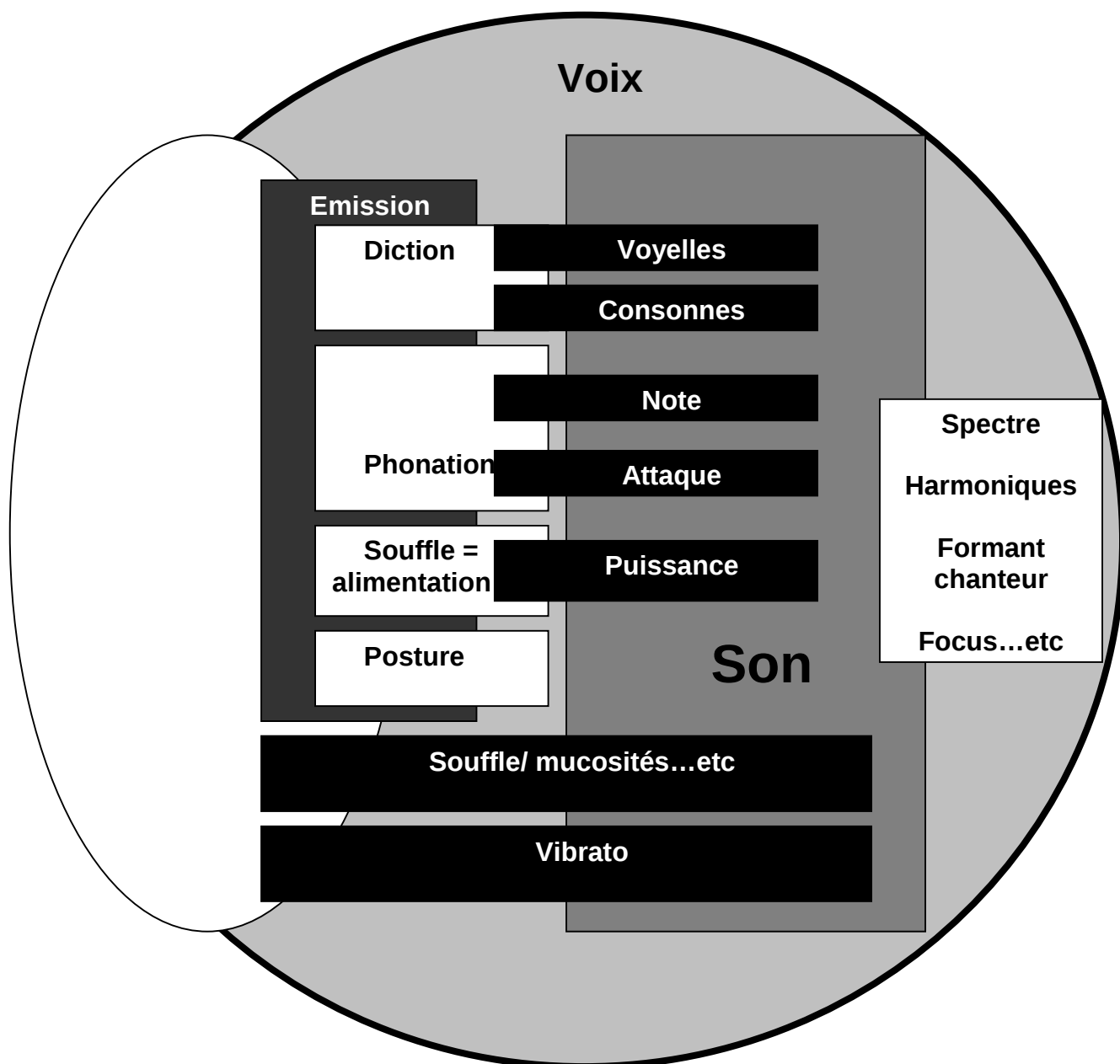


Figure 18 : schéma récapitulatif de la « hiérarchisation » des objets cognitifs dans le discours des experts

3.5) Etude de la modalisation dans le discours

Les six corpus sont des enquêtes réalisées auprès de professeurs de chant, que l'on peut donc légitimement considérer comme des experts du domaine.

Dans l'étude qui suit, nous avons analysé ce que les linguistes appellent aujourd'hui le *discours* de ces experts.

On considère en effet, depuis le linguiste français Benveniste et les travaux qui s'en réclament, que toute *parole* est contrainte par deux types de régularités :

celles qui sont dues à la langue, ou système abstrait de relations et de règles (Tant phonologiques que morphologiques, lexicales, sémantiques ou syntaxiques), tout autant qu'institution sociale et moyen de communication commun à tous les locuteurs d'une même langue ;

celles qui sont dues au discours, ou l'ensemble des réalisations individuelles qui résultent de la mise en œuvre de ce système, et aux « conditions de production » de ce *discours*. Benveniste appelle *énonciation* cet acte individuel de création par lequel tout locuteur fait fonctionner le système de la langue.

Dans la perspective d'*analyse de discours* qui est la nôtre, le texte est pris non pas pour ce qu'il dit mais pour la façon dont il le dit. On considère en effet qu'il existe une pluralité de structurations transphrastiques qui s'articulent entre elles et qui donnent à chaque texte, en fonction de ses conditions de production, une *cohérence textuelle, discursive* particulière, que l'analyse minutieuse des corpus permet de dégager.

Chaque discours manifeste ainsi à travers sa cohérence une appropriation *individuelle* de l'objet collectif qu'est la langue, et ainsi, si on se place d'un point de vue cognitif cette fois (et non plus strictement linguistique), chaque corpus peut également être considéré comme *une* expression des *représentations* que chacun des locuteurs se fait de ce dont il parle. C'est en ce sens qu'à travers une analyse des discours on escompte accéder à une connaissance des représentations cognitives que se font les chanteurs de la qualité des voix.

Sans nous engager très avant dans une typologie des discours on considérera ici que nous avons affaire à des discours spécialisés, la caractéristique principale de ces discours étant d'être produits par des experts.

Parmi les nombreux angles d'attaque que permet l'étude de l'*énonciation*, nous étudierons quelques aspects de la *modalisation*, c'est-à-dire « la composante du procès d'énonciation permettant d'estimer le degré d'adhésion du locuteur à son énoncé » (Dubois1994). afin précisément de viser l'identification des propriétés caractéristiques de la qualité vocale pertinentes pour les locuteurs que nous avons interrogés. On étudiera dans cette perspective les adverbes, les modalisations verbales, les marques de la personne et les énoncés métalinguistiques.

Ainsi, les adverbes et les marques de la modalité (je pense que, je crois que ...) ainsi que les pronoms personnels qui introduisent ces formes permettent-ils de repérer l'état des connaissances du sujet, les certitudes (ou incertitudes qu'il leur accorde...) et à partir des quels on pourra émettre des inférences sur les caractéristiques de ces représentations. Des travaux antérieurs (David, 1997 ; Dubois, 2000) ont en effet permis de repérer que ces marques en discours constituaient des indicateurs sensibles (comme on parle de la sensibilité d'un instrument de mesure physique) des représentations cognitives relatives aux perceptions (visuelles, olfactives mais auditives en particuliers).

L'étude des modalisateurs permet d'étudier le positionnement de l'énonciateur vis-à-vis de ce qu'il énonce, et plus précisément, dans ce corpus d'experts, d'observer la certitude et l'incertitude des connaissances et des jugements dont il fait état, vis-à-vis de sa difficulté à l'exprimer

Ainsi, on sait en effet que l'une des caractéristiques principales des discours d'experts (C. GUASTAVINO(2002) , outre l'emploi, sur le plan lexical, d'un certain nombre de *termes* définis par la discipline, est l'abondance, sur le plan syntaxique, d'énoncés de type *assertif* (ou *déclaratif*) qui « assertent », affirment quelque chose. Ces énoncés sont du type :

*Son vibrato est beaucoup plus régulier que tout à l'heure
Il chante trop haut... il est pas professionnel
Moi, j'ai chanté, j'ai été dans le milieu lyrique pendant 20 ans*

énoncés tout à fait assumés par le locuteur, où il affirme sa position par rapport au monde.

En contraste, ce qui est remarquable dans ces corpus que nous avons (partiellement) étudiés ici, et donc contradictoire, c'est précisément que les assertions sont peu nombreuses comparativement au nombre et à la diversité de l'expression de la modalisation. L'assertion paraît comme enchâssée dans la modalisation, ainsi :

Nan, c'est assez équilibré comme voix, finalement, à la longue. Euh, je vais mettre un tout petit peu sourd quand même... bof, et puis entre clair et sombre, j'arrive pas à me décider donc ça va être au milieu.

En même temps, on entend une sorte d'oscillation dans quelque chose de plat, donc on sait pas... donc... disons qu'il y a un peu de vibrato..

En outre, dans ces énoncés, la présence d'adverbes restrictifs (*assez, un tout petit peu*), de déterminants indéfinis (*une sorte de*), d'auxiliaires d'aspect (*je vais mettre*), l'effacement du sujet d'énonciation (*on entend, donc on sait pas* : qui parle, quel est ce « on » ?) trahissent une distance, une tension entre le sujet et ce qu'il dit, comme s'il lui était difficile, voire impossible d'adhérer à ce qu'il dit.

Cette mise en perspective des connaissances des locuteurs à travers l'étude des modalités ne constitue pas un simple détour concernant ce qui est dit et les contenus de connaissances des locuteurs relativement à la qualité des voix et donc à l'objet central de notre investigation, mais bien un prérequis à des inférences valides sur ces évaluations subjectives. En effet, sur le plan méthodologique, on peut considérer que c'est un passage obligé pour émettre des inférences valides quant aux connaissances en question, que ne permet pas un simple inventaire des formes lexicales simples. (Une telle analyse peut conduire à des contresens).

Sur le plan théorique, on considère que ces formes sont des marques des processus même de construction de la connaissance objective de ce qu'est la qualité de la voix et donc constitue pour nous de précieux repères dans l'identification de notre objet d'étude.

Nous n'avons pu dans le délai du stage explorer exhaustivement l'ensemble des ces formes et en tirer les conséquences quant à l'identification des propriétés des voix, nous donnons ici quelques repères qui illustrent la démarche dans laquelle nous nous sommes engagés, et qui seront travaillés plus précisément ultérieurement.

Aussi avons-nous (commencé de) résumer dans un tableau les différentes formes de la modalisation, qu'elle soit adverbiale ou verbale. Leur abondance montre comme à la loupe l'état d'incertitude des connaissances. Ce tableau est disponible en annexes après les corpus des experts.

On peut noter que, même si ces marques sont particulièrement fréquentes, elles varient notablement en quantité d'un expert à un autre :

Ainsi par exemple, on remarquera une différence importante dans la fréquence du modalisateur adverbial *un peu* entre le corpus du Pré- est et celui de l'expert 5 : 120 occurrences dans le premier contre 20 dans le second, sur un corpus respectivement de 19 et 18 pages. Cette différence est certainement due à la personnalité de chacun des sujets, et, comme telle, ne relève pas de notre champ d'intérêt ici. Par contre on relèvera que ces « traces » sont non seulement présentes mais très nombreuses chez tous ces experts. On remarquera également l'emploi de formes adverbiales telles que :

À peu près, presque, juste, plutôt
Assez, relativement

Peut-être, sans doute
quand même, finalement, enfin

où *enfin* n'est pas employé au sens temporel, mais sert uniquement à « préciser ou corriger ce que l'on vient de dire » (Glose du Petit Robert à l'entrée *enfin*).

Ou encore de formes niées, telles que :

Pas vraiment, pas forcément, pas tellement, pas tout à fait, pas complètement

Autant de précautions oratoires par rapport à l'opinion de l'autre, de l'allocutaire, ici la personne qui pose des questions, mais bien au-delà, à celle du monde musical dans son entier, autant de formules qui permettent tout à la fois de s'engager et de se retirer dans le même mouvement, de progresser dans son dire en fonction de l'assentiment de son allocutaire, toujours recherché en l'absence de consensus solidement établi.

Dans la même tonalité discursive, on observe également l'emploi en grand nombre de modalisations verbales qui, tout comme les adverbes cités, mettent en jeu un axe d'évaluation : à travers eux se profilent de façon assez floue les rapports du vrai, du faux et de l'incertain.

On ne peut qu'être frappé, en particulier dans le corpus du pré-test, par l'absence totale de verbes d'opinion assertant quelque chose comme vrai pour le sujet :

Je crois que, je pense que, je trouve que

contrastant avec une sur-représentation (24 occurrences) de :

je sais pas

et une représentation non négligeable de :

on a / on n'a pas / on peut avoir / l'impression ; on / j'imagine, on peut imaginer

qui dénote une étrange incertitude sur les connaissances de ces experts.

Ce caractère indécis peut s'observer aussi au niveau de la détermination : les déterminants indéfinis viennent souvent se substituer aux articles définis, ainsi

Une sorte de, une espèce de

introduisent un flou et un refus de la clôture censés rendre mieux compte de la nature de l'objet à décrire : ceci est évidemment très éloigné de toute échelle évaluative.

Enfin, les « marques de la personne » choisies ne font que renforcer ce faisceau d'observations.

En effet, dans les discours d'experts ainsi sollicités, la prise de parole est le plus fréquemment à la première personne (ès qualité d'expert), or on remarque dans ces corpus un très grand nombre d'occurrences de « on ». Que représente ce *on* ?

Dans la *structure d'allocation personnelle* décrite par Benveniste, le *moi* s'oppose au *non-moi*, ce qui se traduit, sur le plan grammatical, par les pronoms *je / tu*, qui, à leur tour, s'opposent à *il*. *Je* et *tu* représentant la *personne subjective* et la *personne non subjective*, tandis que le pronom *il* a été appelé par Benveniste (PLG tome 1, pp. 225 à 236 ; tome II pp. 98-99) le pronom de la *non-personne*.

Dans cette structure, le pronom personnel *on* peut avoir plusieurs valeurs. Il s'interprète traditionnellement dans sa valeur indéfinie, comme *l'ensemble indéfini des êtres non-personnels* : c'est le cas dans les énoncés de ce type :

C'est ce qu'on appelle le spinto

Où *on* représente des locuteurs qu'on ne nomme pas, la *doxa*. Mais dans la pratique, ce pronom a pour caractéristique de pouvoir se substituer à tous les autres pronoms. Par exemple comme un

substitut du pronom personnel *nous*, qui équivaut à « moi + eux », ou « moi + vous », selon la situation d'énonciation. Ici, elle ne s'y prête guère, et des cas comme le suivant sont assez rares :

Ça correspond à ce qu'on avait enregistré il y a deux ans ?

Il existe un emploi particulier à l'un des six corpus : en effet, l'un des experts interrogés désigne le chanteur dont il entend l'enregistrement non par *il*, mais par *on*, emploi généralement destiné à marquer une distance par rapport à la personne représentée par ce pronom :

C'est comme si on jouait au haute-contre à la française[...] comme si on cherchait un timbre...

Le cas le plus fréquent, dans ces corpus, est l'emploi du *on* comme substitut du *je*, comme dans l'énoncé suivant :

On a l'impression qu'il veut soulever sa voix, tu vois ?

Quel que soit le cas, l'emploi du pronom personnel indéfini a pour résultat de rejeter le référent dans l'anonymat, ou, comme disent les linguistes parfois, d'*effacer l'agent*. La sur-représentation du pronom *on* dans certains corpus, qui s'ajoute à l'emploi récurrent de tournures impersonnelles telles que :

Y a un tout petit peu de mucosités

Ou :

ça donne simplement une déformation de la voyelle

où le , le *ça*, sur-représenté (102 occurrences !), est employé majoritairement comme sujet de verbe, dans des tournures impersonnelles.

L'ensemble de ces tournures confère aux énoncés une très grande indétermination, et une imprécision de la référence. Comme si les experts ne pouvaient pas s'engager dans leur dire.

Tous ces éléments, le peu d'assertions, les modalisations, l'effacement du sujet de l'énonciation à travers l'emploi de tournures impersonnelles et de pronoms indéfinis concourent au même résultat : tout se passe comme si les locuteurs, ici experts de la voix, tenaient à marquer de façon évidente pour leur allocutaire la distance qu'ils entendent garder vis-à-vis de ce qu'ils disent, comme s'ils n'adhéraient pas vraiment à ce qu'ils énoncent. Cela traduit, nous semble-t-il, une difficulté à se positionner par rapport à ce qui est exprimé, et la difficulté à l'exprimer.

Ceci est en partie lié à la difficulté de parler de cette expertise. Le nombre et la diversité des formules métalinguistiques (tableau X) utilisées dans tous les corpus, du type *on peut dire, comment dire, Je pense qu'on peut dire, on dit aussi, comme on dit, ça veut dire, c'est difficile à expliquer, c'est difficile à définir, disons, je veux dire, c'est-à-dire*, exprimées qui plus est assez souvent au conditionnel : *On pourrait dire, je saurais pas décrire* en...disent assez long sur cette difficulté à établir le consensus parmi les experts. Révélatrice aussi de cette difficulté la phrase d'un expert interrogé sur l'existence ou non de « souffle » dans une réalisation : *On va mettre au milieu pour ne pas se mouiller*

Il est difficile d'exprimer mieux l'incertitude du jugement...

Beaucoup de formes énonciatives indiquent la difficulté à définir et à répondre à la question posée par l'expérimentateur, à savoir la couleur ou la qualité de la voix et se réfugient dans la **description des techniques vocales**, c'est-à-dire un savoir faire, en outre **codifié** dans la pratique (les pratiques éventuellement différentes) des conservatoires et des cultures.

On peut donc précisément illustrer ici l'intérêt d'une méthodologie attentive à l'expression en discours à savoir :

- permet d'identifier l'état des connaissances des locuteurs, qui pourrait apparaître annexe
- mais travers elles les **traces des processus de construction de cette connaissance**. contraints à la fois par les connaissances communes transmises par des discours convenus et standardisés (les discours du conservatoire - que l'on pourrait par ailleurs analyser en tant que tels, - et les connaissances individuelles de chaque chanteur, transmises dans des discours spécifiques de chaque personnalité.

En résumé, et ici nous avons ainsi pu identifier que pour ces experts la qualité de la voix pouvait être décrite précisément à travers la **qualité (experte) d'exécution des modes de production vocales** : l'évaluation de la perception de la qualité de la voix est donc intimement liée, dans la connaissance des processus de production de la voix, et ce sont donc les inférences faites à partir des indices repérés dans le signal qui permettent à ces experts) la voix d'identifier et de juger la qualité des voix.

On ne peut que constater qu'il n'y a pas d'évidence sur le consensus, étant donné l'importance de ces termes qui incluent « dire », renvoyant ainsi à une difficulté dans l'expression de la qualification des voix. C'est ce que nous tenions à mettre en évidence avant d'aborder plus précisément ce qu'une analyse de la sémantique lexicale des formes produites dans le corpus nous permet d'avancer quant à la qualité des voix.

Recherche de corrélats acoustiques

1) Rappels des pistes d'analyse mises en évidence lors de l'étude de la verbalisation des experts.

Pour commencer, il me semble important de rappeler l'originalité et la pertinence de notre méthodologie. Elle ne consiste surtout pas à effectuer des quantités exhaustives d'analyses physiques dont il faudrait ensuite trier les résultats, nous partons toujours de l'écoute des extraits sonores. Avec l'aide de la verbalisation des chanteurs et des experts, nous dégagons tout d'abord des qualités vocales qui ont un sens dans ce contexte et qui renvoient au même concept pour la majorité des experts, pour ensuite rechercher les corrélats acoustiques de ces notions.

L'étude de la verbalisation des experts a permis de valider ou non le choix des échelles dont nous sommes servis pour faire évaluer 15 voix de la base de données audio et électroglottographiques, par les experts. Les résultats de ces échelles ont pu également tout d'abord mettre en évidence les échelles unanimement comprises par les experts : le caractère clair ou sombre d'une voix, sa qualité brillante ou sourde, le timbre de la voix, et le vibrato. Le stage n'étant pas de durée illimitée, nous sommes donc restreints à l'étude de ces qualités vocales, ce travail pouvant être appliqué par la suite à d'autres qualités exprimées dans le discours des chanteurs, mais non quantifiées par des échelles.

Tout d'abord, il s'est dégagé de la verbalisation libre que la clarté serait due en partie à l'antériorisation des voyelles et à la notion d'aigu. Au contraire, le caractère sombre pourrait être provoqué par une postériorisation de la diction et une « aggravation » du spectre.

Quand à la brillance et au timbre, il semblerait que ce soit la répartition des harmoniques dans le spectre (plutôt dans l'aigu pour la brillance et dans le médium pour le timbre) qui soit corrélée à ces qualités vocales. Le formant du chanteur et l'étude des premiers formants du spectre sont donc des pistes à examiner. Enfin, nous avons remarqué que la perception de timbre est liée à celle du vibrato, tandis que la sensation de voix détimbrée provient plutôt de la présence de souffle.

2) Etude du vibrato.

Par rapport aux autres indices acoustiques de la qualité vocale, le vibrato est sans doute le plus simple à repérer, même si sa production physiologique est toujours aussi mystérieuse. Il s'agit en effet d'une modulation du son en fréquence (environ 6 par seconde), dont l'oscillation est plus ou moins rapide et l'amplitude plus ou moins grande.

Les qualificatifs utilisés par les experts pour le décrire sont mêmes parfaitement compréhensibles au regard d'un spectrogramme : En effet, un « vibrato serré » correspond également à des oscillations de période assez petites, donc resserrées lorsqu'on visualise le spectre au cours du temps. A l'inverse, lorsque la fréquence de modulation est faible, le vibrato est alors qualifié « d'oscillation » par les professeurs de chant. On entend effectivement à l'oreille une variation périodique de fréquence et non plus une seule hauteur plus « vibrée » que la normale. Les experts qualifient alors même le vibrato en comptant le nombre de battements par seconde. Mais les experts prétendent que l'impression générale du vibrato dépend de l'adaptation du serrage du vibrato à la hauteur de la note : plus la note est aiguë, plus le vibrato doit être serré pour être apprécié esthétiquement.

Quand à l'amplitude de ces oscillations en fréquence, elles participent à l'impression de flattement lorsqu'elle est trop importante. Quand elle est nulle, le vibrato est inexistant et le son est qualifié de « plat » et « droit ». C'est en effet, ce qu'on observe au spectrogramme, car la note tenue non vibrée reste alors sur la même hauteur.

Cela est illustré sur les deux figures suivantes.

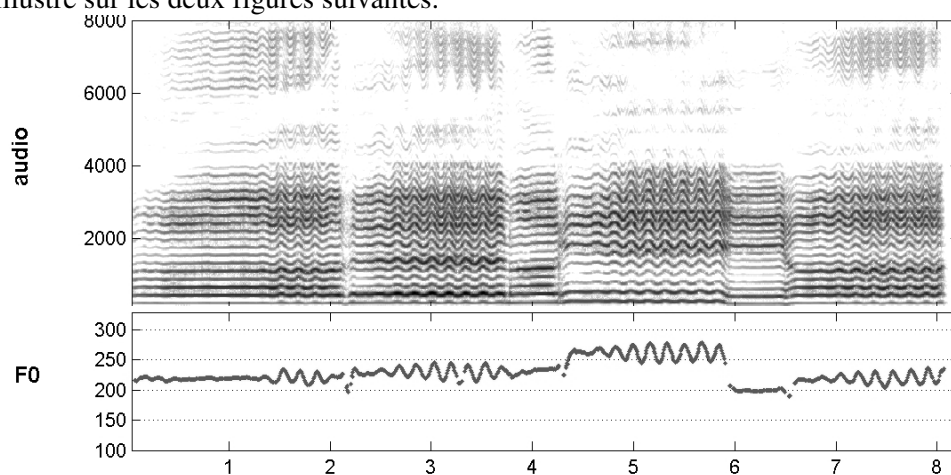


Figure 19 : Voix 1 du chanteur B3, perçue avec du vibrato (esthétique)

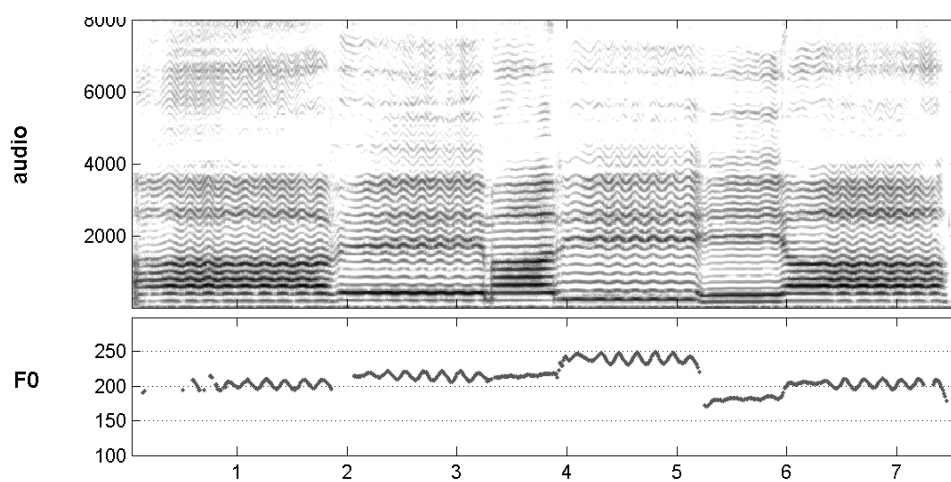


Figure 20 : Voix normale du chanteur T2, perçue avec un vibrato faible.

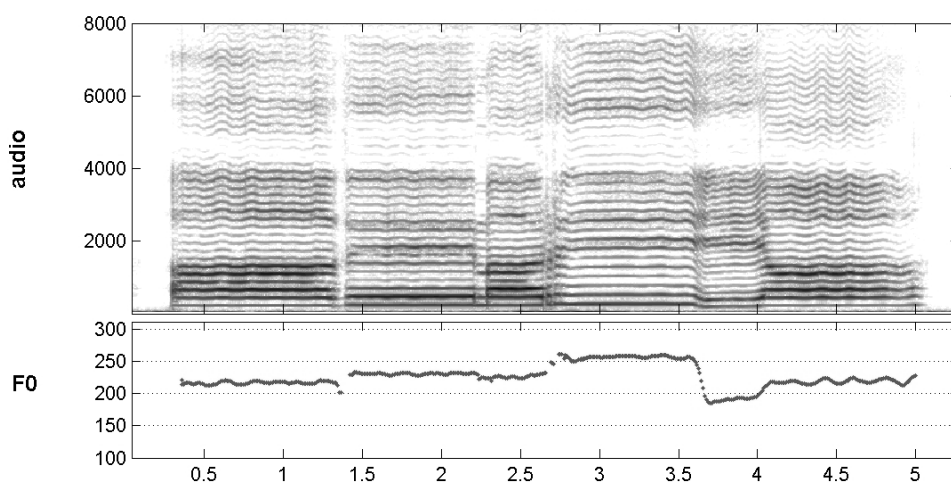


Figure 21 : Voix 4 du chanteur B6, perçue sans vibrato.

3) Recherche de corrélation entre le caractère clair /sombre et la couverture des voyelles

D'après le discours des experts, il semblerait que le caractère sombre ou clair serait dû à l'ouverture et la couverture. C'est pourquoi nous avons étudié le spectre moyenné sur les deux « A » et le « I » de la phrase « Ave maria ». Nous avons observé des variations d'amplitude sur les premiers harmonique dus aux variations de position des deux premiers formants de la voix. La hauteur de la note émise par les chanteurs (vers les 200 / 250 Hz) est bien en dessous de la fréquence du premier formant. C'est tout l'intérêt également d'avoir choisi de restreindre notre étude à des voix d'hommes. Pour le « A », les deux premiers formants se situent vers le 3^{ème} et le 5^{ème} harmonique (que nous appellerons H3 et H5). Nous avons donc choisi d'étudier la différence entre le 1^{er} harmonique (le fondamental) et le 3^{ème}, et entre le 3^{ème} et le 5^{ème}.

Toutes ces analyses acoustiques se sont faites à partir du signal audio et du signal egg pour le quotient ouvert. Nous avons repris et enrichi des programmes matlab développés par Nathalie Henrich au cours de sa thèse.

En fait, nous voulions comparer au départ les voix les plus claires et les plus sombres d'un même chanteur les unes par rapport aux autres. Cela ne donnait aucun résultat intéressant. En revanche, en comparant les voix claires et sombres d'un chanteur à son émission, nous obtenons les figures suivantes de répartition des voix d'un même chanteur selon la différence entre H1 et H3 et la différence entre H3 et H5, par rapport à leur émission normale.

D'après les calculs statistiques et les histogrammes représentés figures 14, 15 et 16 page 45, les voix 2, 3 et 5 du chanteur 1 (B3), la voix 9 du chanteur 2 (T2) et les voix 12, 13 et 14 sont unanimement perçues comme **claires** et les voix 4, 8 et 11 comme **sombres**.

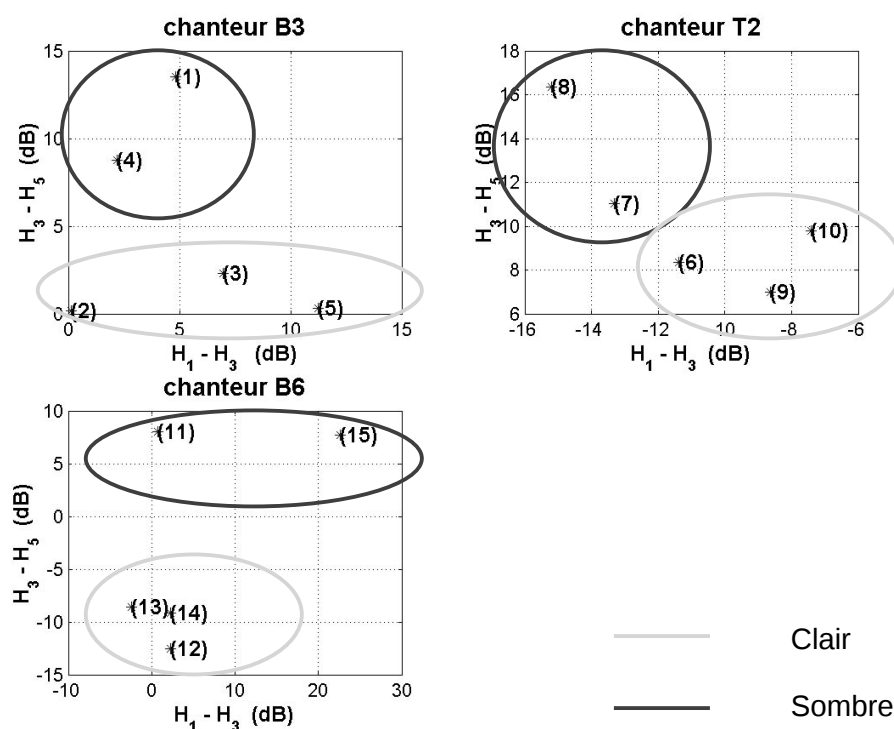


Figure 22 : représentation des 1ers « A » des voix selon la différence des harmoniques 1,3,5

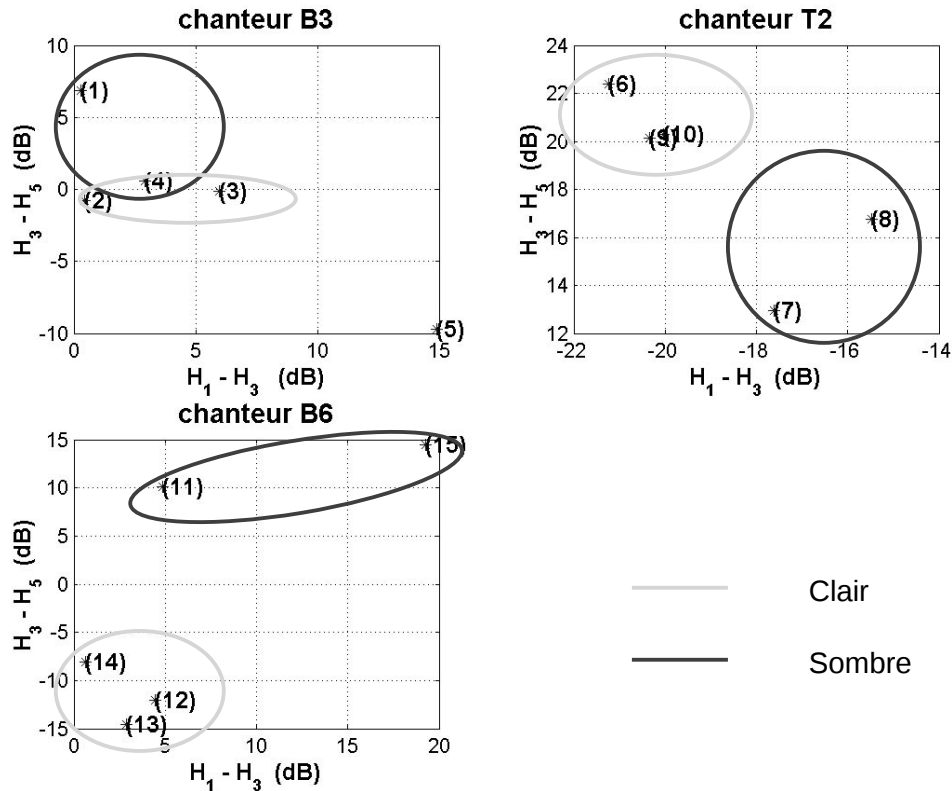


Figure 23 : représentation des 2èmes « A » des voix selon la différence des harmoniques 1,3,5

On constate tout d'abord que le « A » est à peu près placé de la même façon pour le chanteur B6 dans ses deux occurrences, ce qui n'est pas du tout le cas pour les deux autres chanteurs. Cela est intéressant, car au regard des résultats du 1^{er} « A », il semblait que l'axe H_3-H_5 soit discriminant des qualités sombres ou claires : les voix sombres ayant un fort harmonique 3 par rapport à l'harmonique 5 (c'est-à-dire un déplacement du spectre vers le grave par rapport à l'émission normale) est les voix claires une différence moins grande, voire un harmonique 5 supérieur au 3^{ème} (donc un déplacement formantique vers l'aigu).

Cependant, ceci n'est pas vérifié pour le 2^{ème} « A » du chanteur T2. Cependant, il nous semblait à l'écoute que la fin des extraits étaient moins représentatifs de l'intention de qualité vocale du chanteur.

2) Recherche de corrélation entre la brillance et l'intensité du formant du chanteur

En ce qui concerne la brillance, notre piste résidait dans l'étude du formant du chanteur et de la balance harmonique générale. Nous avons donc décidé de nous intéresser à la différence d'amplitude entre le maximum des deux premiers formants (donc des premiers harmoniques, situés dans la zone [0 2000 Hz]) et l'amplitude du formant du chanteur (cette fois dans la zone d'écoute privilégiée de l'oreille, entre 2000 et 4000 Hz).

Voici représenté sur les figures ci-après la répartition des voix selon leur évaluation par les chanteurs sur l'échelle Brillant/Sourd et sur la différence d'amplitude entre les deux zones spectrales.

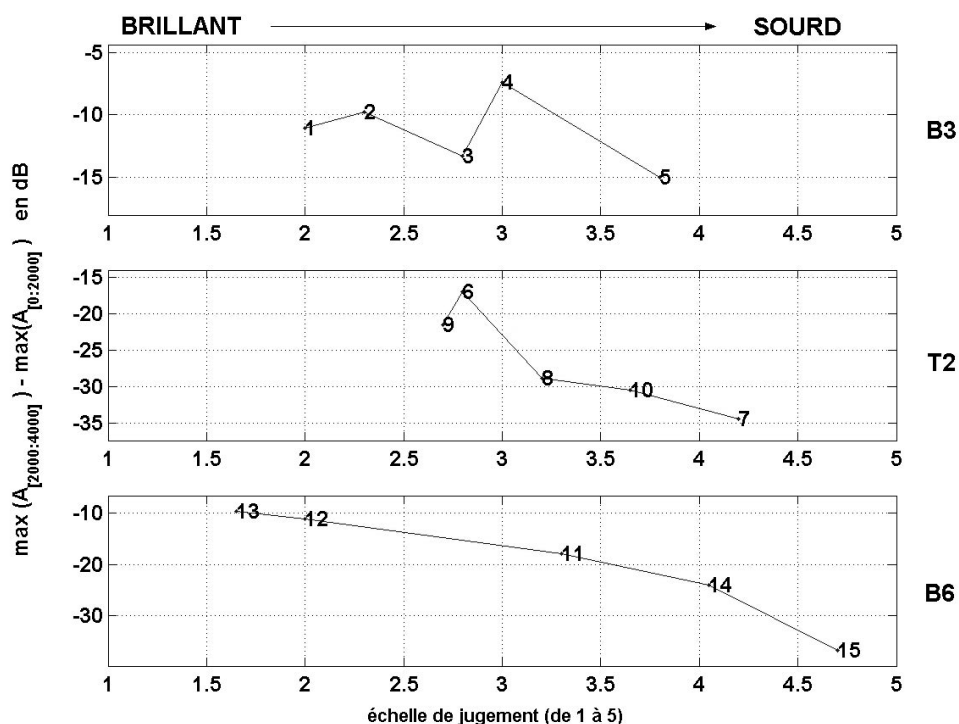


Figure 25: Emergence spectrale de la zone du formant du chanteur (entre 2000 et 4000 Hz) par rapport à la zone des premiers formants (entre 0 et 2000 Hz), en fonction du caractère brillant ou sourd de la voix. Cas de la première occurrence de la voyelle [a] dans la phrase chantée.

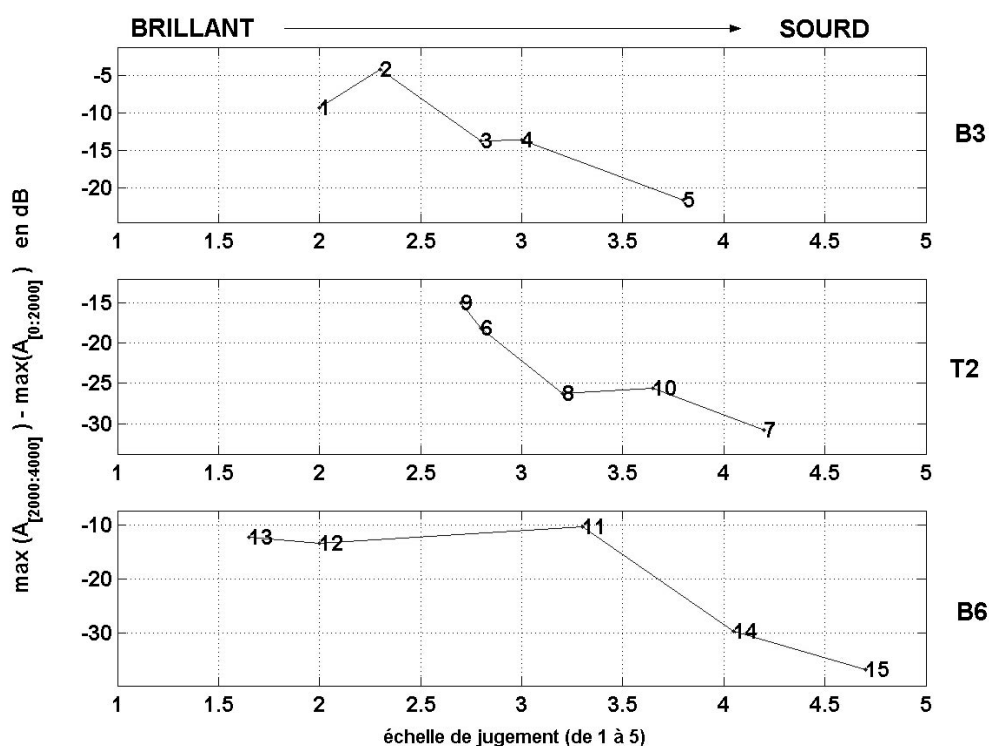


Figure 26 : Emergence spectrale de la zone du formant du chanteur (entre 2000 et 4000 Hz) par rapport à la zone des premiers formants (entre 0 et 2000 Hz), en fonction du caractère brillant ou sourd de la voix. Cas de la seconde occurrence de la voyelle [a] dans la phrase chantée

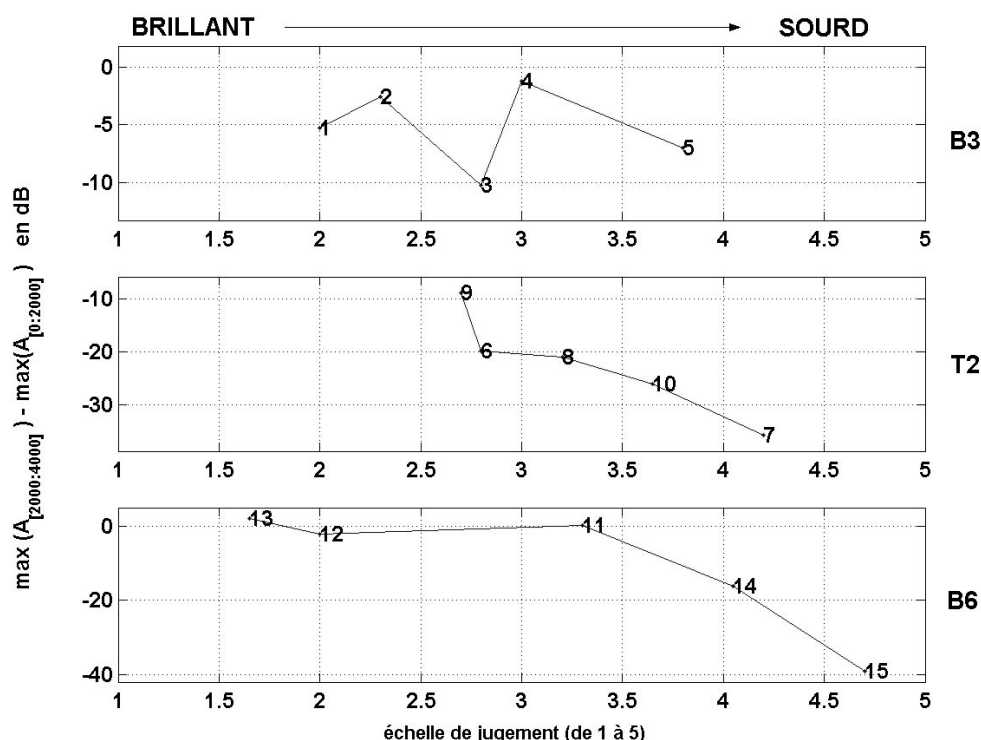


Figure 27 : Emergence spectrale de la zone du formant du chanteur (entre 2000 et 4000 Hz) par rapport à la zone des premiers formants (entre 0 et 2000 Hz), en fonction du caractère brillant ou sourd de la voix. Cas de la première occurrence de la voyelle [i] dans la phrase chantée.

La différence d'amplitude entre les deux zones formantiques [0 2000 Hz] et [2000 4000Hz] est bien représentative de la qualité de brillance pour T2 et B6. Les productions de B3, plus complexes, demanderaient une analyse plus poussée.

4) Recherche de corrélats pour le timbrage

Nous avons l'intention d'analyser le spectre par bande de fréquence pour étudier la balance spectrale. Le temps nous a manqué pour le faire. Cependant, en traçant le spectre moyenné sur une même voyelle des différentes voix d'un même chanteur, on voit nettement se séparer les voix qualifiées de timbrées et celles décrites comme détimbrées.

Les figures 28, 29 et 30 présentent les **spectres moyennés sur l'ensemble de la phrase musicale** pour les 15 voix. Les productions vocales qui ont été jugées plutôt « timbrées » (score de jugement supérieur à 2.5) sont présentées en trait gras sur les figures et les productions qui ont été jugées plutôt « détimbrées » (score de jugement inférieur à 2.5) sont présentées en trait fin. Dans le cas du chanteur B3, la voix 3 n'a été jugée ni très timbrée, ni très détimbrée (score de 2.7) et elle est donc présentée en gras pointillé sur la figure 1.

Ces figures illustrent les différences spectrales que l'on peut observer entre une voix jugée « timbrée » et une voix jugée « détimbrée » si l'on s'intéresse à la phrase musicale dans sa globalité. Comme on s'y attendrait intuitivement, une voix jugée détimbrée présente une forte atténuation spectrale dans la région 1000-4000 Hz, pouvant également s'étendre au-delà (comme dans le cas du chanteur B6).

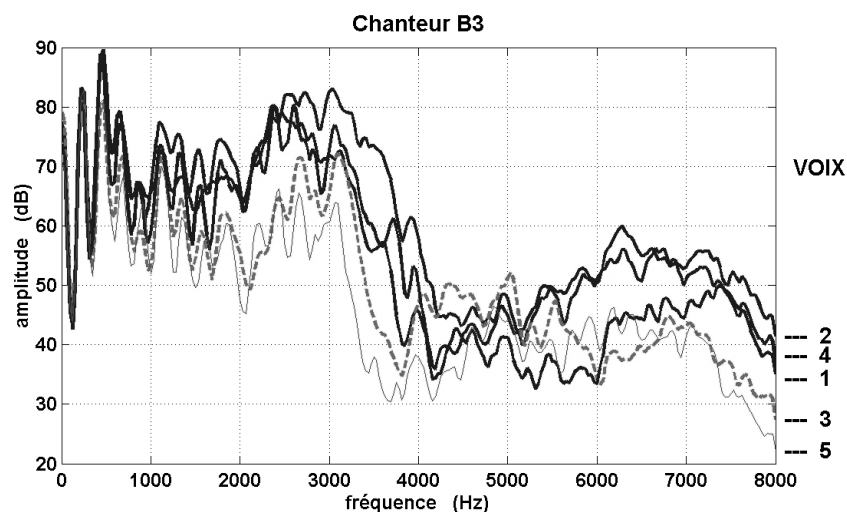


Figure 28 : Spectre moyenné sur l'ensemble de la phrase musicale, dans le cas du chanteur B3. En trait épais plein : voix jugée timbrée (>2.5), en trait épais pointillé, voix jugée ni timbrée, ni détimbrée (~ 2.5); en trait fin, voix jugée détimbrée (<2.5).

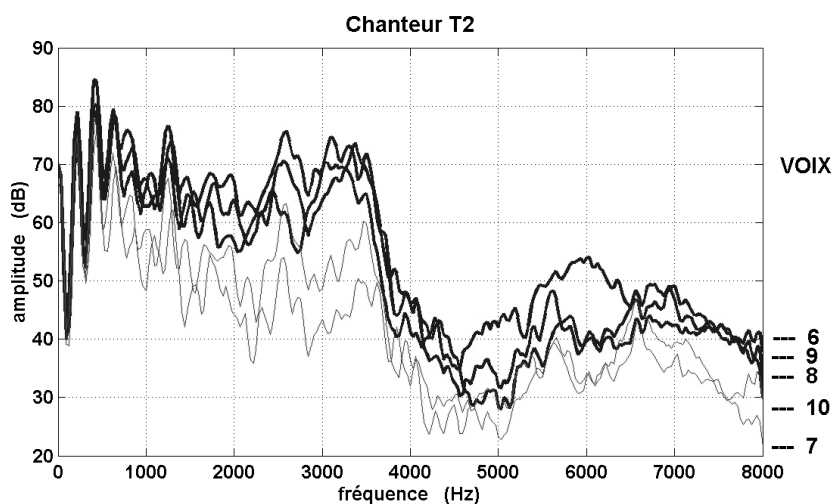


Figure 29 : Spectre moyenné sur l'ensemble de la phrase musicale, dans le cas du chanteur T2. En trait épais : voix jugée timbrée; en trait fin, voix jugée détimbrée (<2.5).

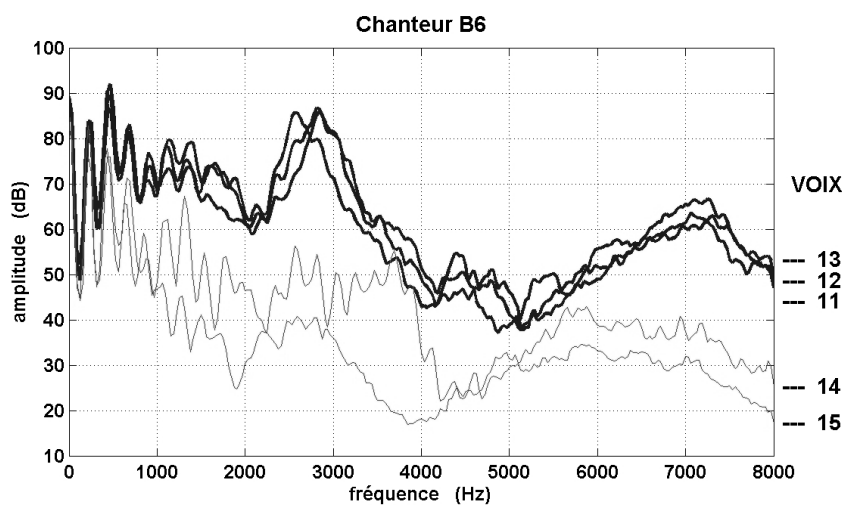


Figure 30 : Spectre moyenné sur l'ensemble de la phrase musicale, dans le cas du chanteur T2. En trait épais : voix jugée timbrée; en trait fin, voix jugée détimbrée (<2.5).

5) Corrélation entre l'air, la brillance et le timbrage.

Par les calculs statistiques de corrélation entre échelles et dans le discours des experts, nous avons mis en évidence un lien entre la brillance et le timbrage, et entre le détimbrage et la présence d'air sur la voix.

Ici, nous avons donc voulu examiner l'influence du **rapport d'amplitude entre les premiers formant et le formant du chanteur** pour le timbrage et l'air sur la voix, de la même façon qu'on l'avait fait pour l'étude de la brillance.

Voici représenté sur les figures qui suivent la répartition des voix selon leur évaluation par les chanteurs sur l'échelle Timbré/Détimbré et sur la différence d'amplitude entre les deux zones spectrales.

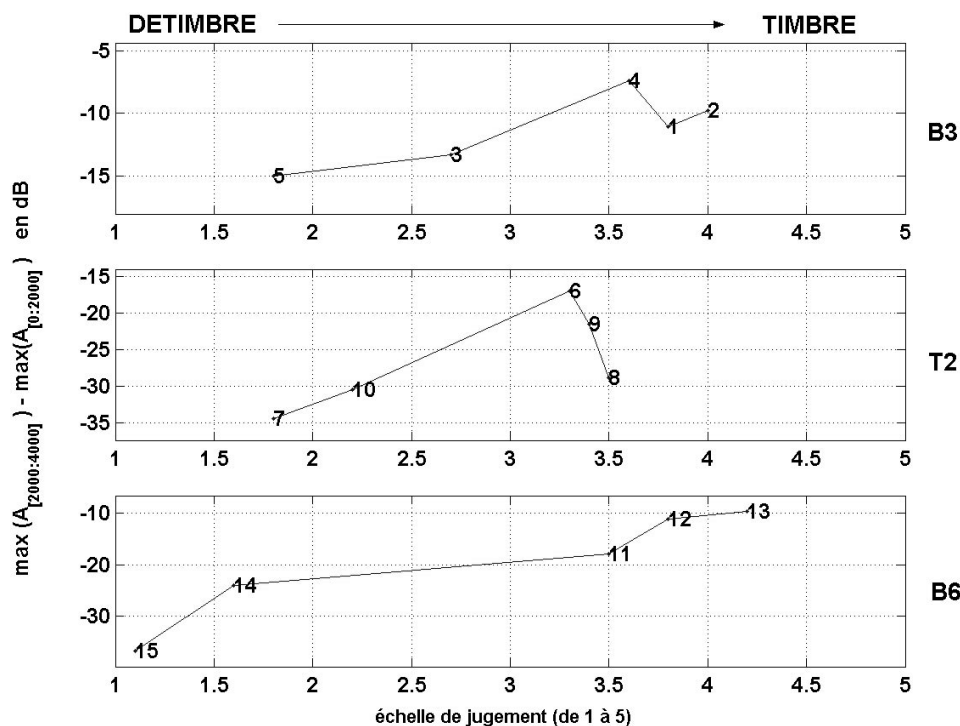


Figure 31: Emergence spectrale de la zone du formant du chanteur (entre 2000 et 4000 Hz) par rapport à la zone des premiers formants (entre 0 et 2000 Hz), en fonction du caractère timbré ou détimbré de la voix. Cas de la première occurrence de la voyelle [a] dans la phrase chantée.

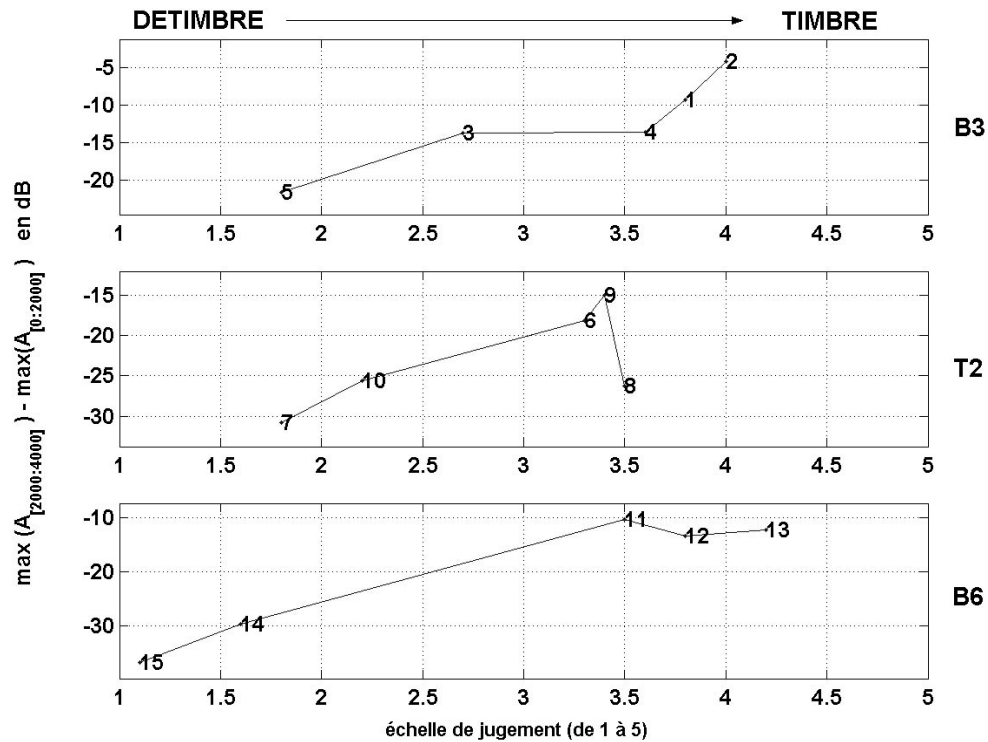


Figure 33: Emergence spectrale de la zone du formant du chanteur (entre 2000 et 4000 Hz) par rapport à la zone des premiers formants (entre 0 et 2000 Hz), en fonction du caractère timbré ou détimbré de la voix. Cas de la deuxième occurrence de la voyelle [a] dans la phrase chantée.

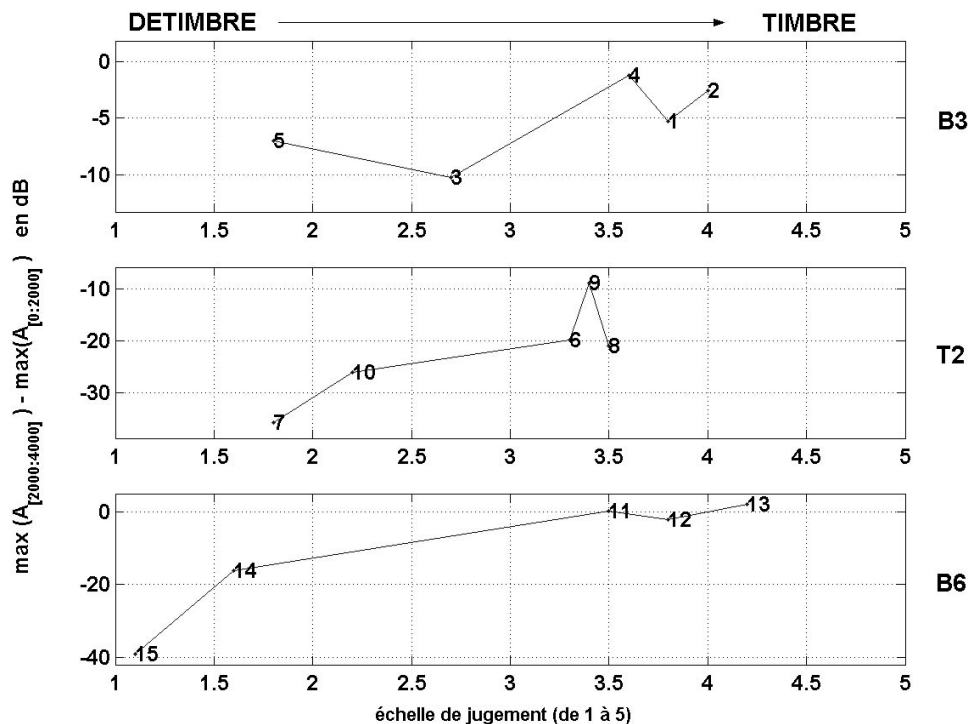


Figure 34: Emergence spectrale de la zone du formant du chanteur (entre 2000 et 4000 Hz) par rapport à la zone des premiers formants (entre 0 et 2000 Hz), en fonction du caractère timbré ou détimbré de la voix. Cas de la voyelle [i] dans la phrase chantée.

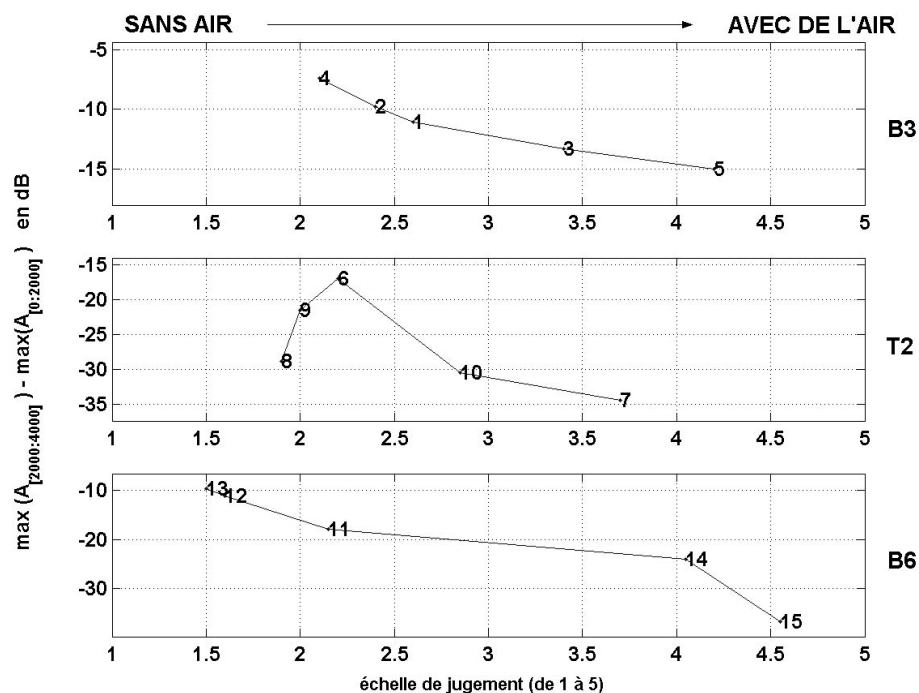


Figure 35: Emergence spectrale de la zone du formant du chanteur (entre 2000 et 4000 Hz) par rapport à la zone des premiers formants (entre 0 et 2000 Hz), en fonction de la présence ou non d'air sur la voix. Cas de la première occurrence de la voyelle [a] dans la phrase chantée.

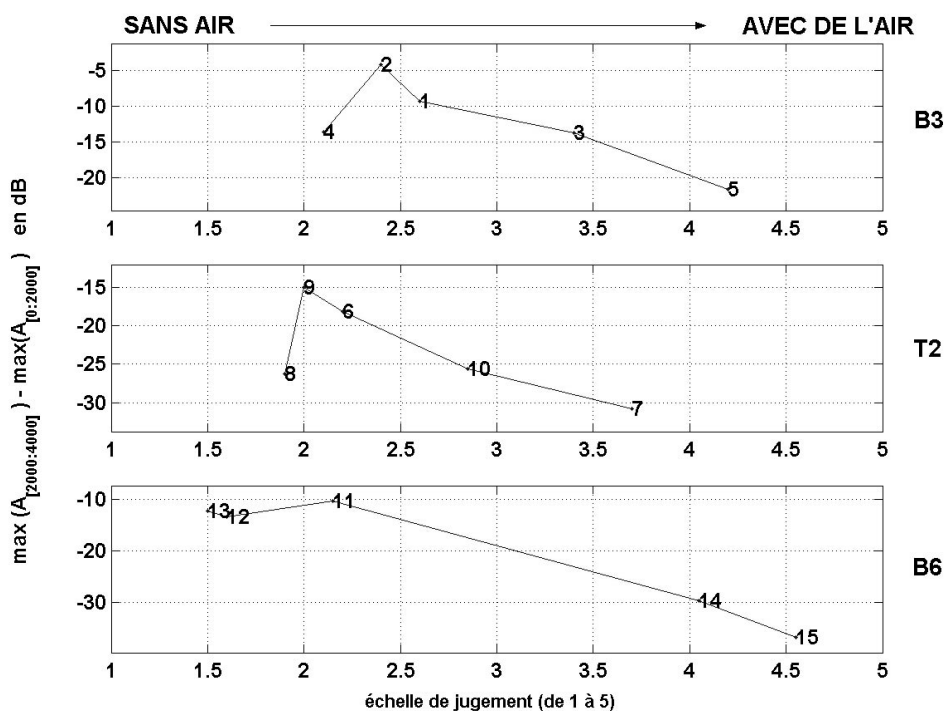


Figure 36: Emergence spectrale de la zone du formant du chanteur (entre 2000 et 4000 Hz) par rapport à la zone des premiers formants (entre 0 et 2000 Hz), en fonction de la présence ou non d'air sur la voix. Cas de la deuxième occurrence de la voyelle [a] dans la phrase chantée.

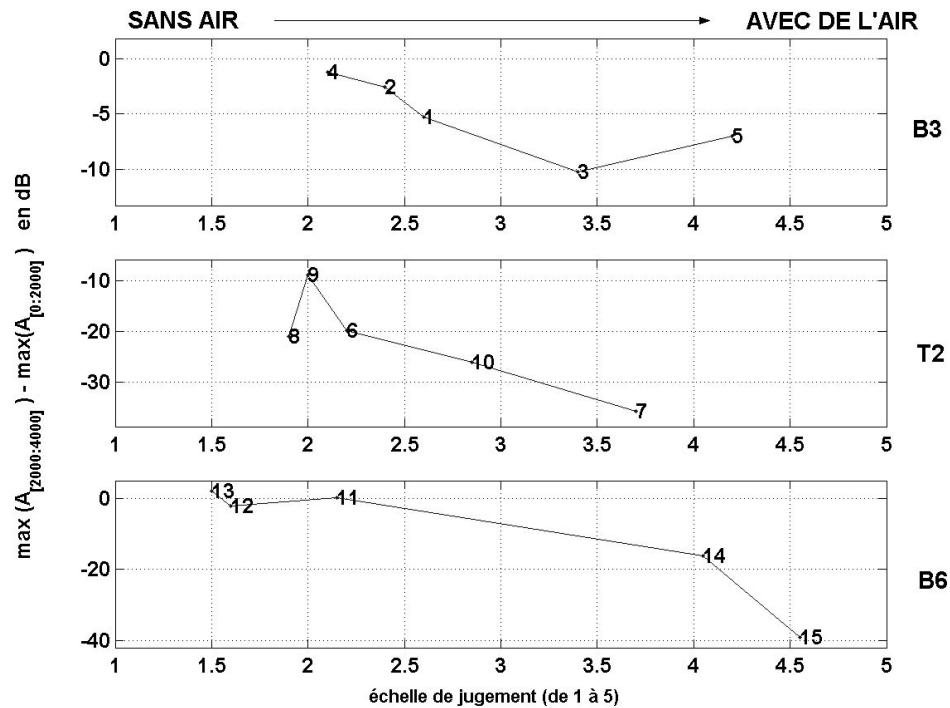


Figure 37: Emergence spectrale de la zone du formant du chanteur (entre 2000 et 4000 Hz) par rapport à la zone des premiers formants (entre 0 et 2000 Hz), en fonction de la présence ou non d'air sur la voix. Cas de la voyelle [i] dans la phrase chantée.

On voit que le renforcement du chanteur par rapport aux premiers formants de la voix correspond à la fois à la qualité de brillance, au timbrage d'une voix. Et à l'inverse, la présence de souffle comme le détimbrage sont caractérisés pas l'importance des premiers formants de la voix au détriment de la partie haute du spectre, donnant cette impression de pauvreté au son.

Conclusion

A venir

Bibliographie

Ouvrages

- « L'enquête et ses méthodes : l'entretien »
Alain Blanchet, Anne Gotman
Collection Sociologie 128 . N°19 . Nathan Université
1992
- « L'enquête et ses méthodes : le questionnaire »
François de Singly
Collection Sociologie 128 . N°18 . Nathan Université
1992
- « dictionnaire de linguistique »
J Dubois, L Guespin, M Giacomo, C et JB Marcellesi, JP Mevel
Larousse
1973
- « Catégorisation et cognition : de la perception au discours »
sous la direction de Danièle Dubois
Edition Kimé
1997
- « Expériences individuelles et savoirs collectifs : Questions de méthodes »
Cahier du LCPE N°3
1999
- « Langages, cognitions, pratiques, ergonomie »
Cahier du LCPE
2003
- « arbres, classes, distances... »
Cahier du LCPE N°6
2002
- « Normes catégorielles et listes lexicales »
Cahier du LCPE N°5
2002
- « Trouver sa voix »
Louis Jacques Rondeleux
Edition du seuil
1977
- « Vocal Fold Physiology : Voice quality control »
Singular publishing group, Inc
1995
- "Emotion recognition in Human-computer interaction"
IEEE Signal Processing Magazine
January 2001
- "La musique et les sciences cognitives"
S McAdams, I Delège
Pierre Madage, éditeur
1989
- CD + livre « écoute et découverte de la voix »
M Asselineau, E Bérel
Editions Fuseau
1990
- « Singing : the mechanism and the technique »
William Vennard
Editions Carl Fischer

1967

- “Le geste technique. Réflexions méthodologiques et anthropologiques »
chapitre : Geste, mouvement, action. Analyse lexicale et sémantique des concepts
Danièle Dubois, Nathalie Klumpp, Mary-Annick Morel
TIP (Technologies / Idéologies / Pratiques) volume XIV n°2. Revue d’anthropologie des connaissances.
Editions Eres
2002
- Neisser U (1987) Concepts and conceptual development: ecological and intellectual factors in categorization. Cambridge: Cambridge University Press
- Rastier F (1991). *Sémantique et recherches cognitives*, Paris: PUF

Thèses et rapports de DEA

- « Simulation des ambiances sonores urbaines : intégration des aspects qualitatifs »
Thèse de Manon Raimbault
2002
- « Etude de la source glottique en voix parlée et chantée : modélisation et estimation, mesures acoustiques et électroglottographiques, perception »
Thèse de Nathalie Henrich
2001
- « Mise en discours d’un parcours olfactif »
Rapport de stage de DEA de Gaëlle Delepaut
2002
- « Des sons aux mots : Comment parle-t-on du timbre musical »
Thèse de Anne Faure
- « Perception de la voix parlée : cohérence du timbre du locuteur »
Thèse de Blas Payri
2000
- “Les registres et passages dans la voix chantée”
Thèse de Bertrand Chuberre
2000

Articles

- « The human voice »
J Sundberg
Comprehensive Human Physiology, Vol.1
1996
- “Otho-rhino-laryngologie”
Instantanés médicaux
Encyclopédie médico-chirurgicale
- « Analysis, synthesis and perception of voice quality variations among female and male talkers »
D H Klatt, L C Klatt
Journal of acoustical society of america
1990
- “Vocal quality factors ; analysis, synthesis and perception”
D G Childers, C K Lee
Journal of acoustical society of america
1991
- “Perceptual evaluation”
Sören Fex

- Journal of voice Vol 6 , N°2, pp. 155-158
1992
- “Perceptual aspect of singing”
J Sundberg
Journal of voice Vol 8 , N°2, pp. 106-122
1994
 - “A method for describing different styles of singing”
J Sundberg, M Thalen
 - “Prosodic Parameters of Perceived Emotions in Vocal Servers Voices”
Ehrette T., Chateau N., d'Alessandro C. and Maffiolo V
Speech Prosody, 259-262
2002
 - "Predicting the Perceptive Judgment of Voices in a Telecom Context: Selection of Acoustic Parameters”
Ehrette T., Chateau N., d'Alessandro C. and Maffiolo V
Eurospeech, In Press
2003