

## **Perception et verbalisation de la qualité vocale dans le chant savant occidental de l'adulte : Apport d'un groupe de réflexion pluridisciplinaire**

Nathalie Henrich, Pascal Bezard, Robert Expert, Maëva Garnier, Christian Guerin, Claire Pillot, Sophie Quattrocchi, Bernard Roubeau, Boris Terk

Traduction française de l'article :  
**Perception and verbalisation of voice quality in western lyrical singing:  
Contribution of a multidisciplinary research group**  
publié dans :

K. Maimets-Volk, R. Parncutt, M. Marin & J. Ross (Eds.)  
Proceedings of the third Conference on Interdisciplinary Musicology (CIM07)  
Tallinn, Estonia, 15-19 August 2007, <http://www-gewi.uni-graz.at/cim07/>

### **Introduction**

Dans le domaine du traitement de la parole, la qualité vocale se définit par ce qui différencie deux productions vocales ayant le même contenu lexical. Ces différences peuvent être prosodiques ou acoustiques, liées à des variations de rythme, de hauteur, d'intensité, de timbre ... Elles interviennent à tous les niveaux de segmentation de la parole : au niveau du phonème, par une variation du contenu spectral d'un son tenu, au niveau du mot, en relation avec des variations temporelles et spectrales locales, et au niveau de la phrase, en relation avec des variations temporelles et spectrales plus globales. Comment décrit-on alors la qualité perçue d'une voix ? Ce questionnement s'inscrit dans un problème global de description d'un son : quels sont les modes d'écoute et les termes pour décrire ce que l'on entend ? La perception de la qualité vocale est une notion complexe, fortement subjective et dépendante de l'auditeur. Les modes d'écoute et les termes employés pour sa description varient selon les spécialités : les orthophonistes et phoniâtres, les acteurs, chanteurs et professeurs de chant, les coaches vocaux, et les scientifiques menant des recherches sur la voix, ne partagent pas de méthode d'écoute ni de vocabulaire consensuel pour décrire la qualité perçue d'une voix. Dans leur pratique professionnelle, les phoniâtres et orthophonistes ont développé un vocabulaire pour décrire la qualité de voix pathologiques. Leur écoute est principalement orientée vers la recherche de « défauts » à des fins de diagnostic. Un effort a déjà été fait dans ce domaine pour retenir les termes les plus consensuels et pertinents à la discrimination perceptive des différentes pathologies vocales. Des échelles perceptives d'évaluation de la qualité vocale, telle l'échelle GRBAS « Grade, Roughness, Breathiness, Aesthenia, Strain » (Isshiki & Takeuchi, 1970 ; Hirano, 1981, 1989), ou l'échelle RBS « Roughness, Breathiness, Hoarseness » (Wendler et al., 1986) sont couramment utilisées. Dans le cas de la voix parlée non-pathologique et de la voix chantée, un tel consensus reste à établir. Dans le domaine du chant lyrique en particulier, il existe une terminologie très riche et variée (Guerin 2006; Vennard, 1967 ; Miller, 1986). Parmi la multitude de descripteurs utilisés pour décrire la qualité vocale, certains montrent un bon consensus entre les spécialistes de ce domaine. D'autres, plus métaphoriques ou esthétiques, montrent une grande polysémie ou une signification individuelle, malgré leur usage très fréquent. Les descripteurs utilisés par les spécialistes ne concernent pas seulement le son perçu, mais également le mode de production (Wapnick and Ekholm, 1997; Garnier et al., 2005). L'imitation est souvent utilisée en complément de la description verbale.

Il manque aux experts des divers domaines un langage commun à la pédagogie vocale, à la rééducation de la voix et à l'acoustique musicale pour décrire la qualité perçue d'une voix (Ekholm et al., 1998). La recherche de ce langage commun a amené la formation d'un groupe de recherche pluridisciplinaire auquel participent activement des chercheurs

en acoustique musicale, des orthophonistes, des chanteurs et professeurs de chant. Ce désir de trouver une terminologie unifiée autour d'un objet sensible, et non plus relative à chaque discipline qui gravite autour, n'est pas propre à la voix. D'autres domaines tels que l'œnologie (Guinard and Noble, 1986) ou le textile (Philippe et al., 2001) ont, principalement pour des raisons commerciales, trouvé un consensus pour permettre aux différents acteurs de ces domaines de dialoguer. Le travail présenté ici est le fruit de trois années de réflexion, de discussion et d'écoute de ce groupe de recherche pluridisciplinaire. Les objectifs de ce groupe de recherche sont (1) de clarifier la notion de qualité vocale et ses modes d'écoute, (2) de recenser la terminologie utilisée pour la décrire verbalement et en retenir les termes ou critères consensuels entre les différentes spécialités, (3) de définir précisément ces termes ou critères retenus, et de les illustrer perceptivement par des exemples sonores de référence, dans le but double d'établir un consensus de référence pour dialoguer entre les différentes disciplines et d'« éduquer » de nouveaux auditeurs à l'écoute analytique de voix.

Pour cette étude, nous nous plaçons dans le cadre du chant savant occidental de l'adulte, pour lequel nous cherchons à caractériser la notion de qualité vocale et à décrire les voix perçues à l'aide d'une grille de description basée sur des qualificatifs techniques consensuels et des exemples sonores illustratifs. Après avoir aborder quelques aspects fondamentaux de la perception d'un phénomène sensible, nous discuterons de la verbalisation libre de la qualité vocale, des termes et modes d'écoute qui en émergent. En s'appuyant sur ces observations, nous présenterons une grille d'écoute sélective selon deux axes majeurs. A partir des résultats d'un test perceptif mené sur cette grille, nous explorerons le consensus établi. Nous concluons enfin sur l'intérêt de cette approche et sur l'outil proposé.

## **De la perception à la verbalisation**

Plusieurs aspects fondamentaux de la perception d'un phénomène sensible ont une conséquence directe sur sa description verbale.

### **Une perception catégorielle**

De nombreuses études ont montré que la perception tend avant toute chose à identifier l'objet, afin de le situer par rapport aux catégories mentales déjà construites par le sujet (Castellengo, 1986). Une perception analytique peut avoir lieu seulement dans un deuxième temps (Schaeffer, 1966) et les descripteurs verbaux utilisés sont alors dépendants de la catégorisation initiale de l'objet (Dubois, 1991). Ainsi, nous ne décrirons pas de la même façon une voix parlée ou une voix chantée, et au sein de la catégorie « voix chantée », une voix lyrique ou une voix non-lyrique. Dans ce contexte, le groupe de recherche a dû choisir en premier lieu la catégorie de productions vocales sur laquelle porter son investigation de la qualité vocale. Ce choix s'est porté sur le chant lyrique, ou chant savant occidental de l'adulte.

### **Une perception individuelle**

L'information qualitative ne réside pas dans l'objet, mais résulte de l'interprétation d'indices par le sujet, à travers le filtre de ses représentations mentales. Ainsi, les expériences passées de chaque individu, ses attentes vis à vis de l'objet, et les buts de son écoute (tous dépendants de son domaine d'expertise), vont l'influencer dans sa perception et dans l'attention qu'il va porter à certains indices de l'objet plutôt qu'à d'autres. Il semble donc essentiel de guider la perception, d'orienter l'écoute vers les aspects à partager. Rapidement, le groupe de recherche a donc été amené à élaborer une grille d'écoute sélective, dont l'un des buts principaux est de permettre de diriger la perception vers des indices choisis liés à la qualité vocale.

### **Une perception différentielle**

La perception humaine est différentielle : le sujet ne peut évaluer ou décrire un objet dans l'absolu, mais toujours par comparaison à un autre objet présenté ou à un prototype en mémoire de la catégorie de l'objet. Par conséquent, la description verbale

d'un objet peut s'appuyer sur l'usage de comparatifs, et elle concerne plus souvent les défauts de l'objet (i.e. ses écarts à une norme) que ses qualités (Faure, 2000). Cela est d'autant plus valable dans des domaines esthétiques comme le chant lyrique, où la dimension hédoniste fait partie intégrante de l'objet et ne peut être mise de côté lors de son évaluation. La mise en place d'une méthodologie d'écoute et de description de la qualité vocale doit prendre en compte autant que possible le caractère différentiel de la perception humaine. L'élaboration de référents en mémoire partagés par une communauté peut s'appuyer sur la familiarisation avec des objets sonores prototypiques. C'est dans cet esprit qu'a été constituée une base d'exemples sonores illustratifs des critères de qualité vocale retenus dans cette étude.

### Exprimer sa perception : quels mots pour décrire la qualité perçue d'une voix lyrique ?

La littérature fournit de nombreux glossaires de termes liés à la qualité vocale. Citons par exemple les glossaires proposés par Vennard (1967), par Miller R. (1986, pp. 330-332 dans la version française), par Titze (1995). Chaque spécialiste se construit son vocabulaire pour décrire la voix, et celui-ci n'est que partiellement partagé par les autres spécialistes. Afin de cerner le vocabulaire de chacun et la façon dont il s'organise, un premier travail a consisté à comprendre comment et avec quels mots nous parlons de la qualité vocale. Nous avons donc décrit verbalement et librement la qualité vocale de plusieurs exemples sonores commerciaux et de laboratoire. Un vocabulaire très varié est ressorti de cette première phase exploratoire, qu'il a alors fallu mettre en catégorie. Nous nous sommes alors appuyés sur une analyse psycholinguistique de la verbalisation de la qualité vocale menée parallèlement auprès de professeurs de chant (Garnier et al., 2005). Dans cette étude, des professeurs de chant étaient amenés à s'exprimer librement sur la qualité vocale à l'écoute d'exemples vocaux produits en laboratoire. L'analyse linguistique de leurs discours a fait émerger une partie du lexique associé à la notion de qualité vocale, et a permis d'explorer les différents concepts qui y sont associés. En appliquant une méthodologie psycholinguistique développée pour la catégorisation et l'expression en langue des phénomènes sensibles (Dubois, 1997, 2000), le lexique employé par les professeurs de chant dans cette étude a été réparti en quatre grandes catégories, qui sont présentés sur la Figure 1.

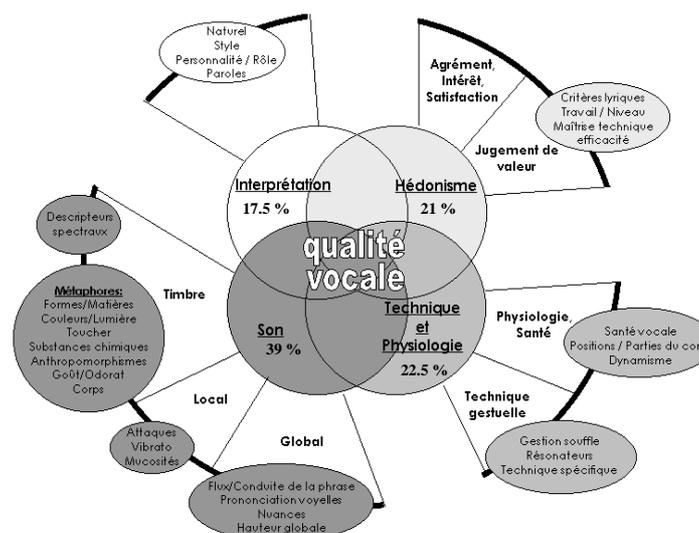


Figure 1: Catégorisation des termes employés en verbalisation libre de la qualité vocale (d'après Garnier et al., 2005). Les termes employés pour décrire la qualité vocale se partagent en quatre catégories principales : référence au son (39%), à la technique et à la physiologie (22.5%), à l'hédonisme (21%), et à l'interprétation (17.5%).

## Vers une écoute dirigée

Ces quatre grandes catégories ont fortement inspirées la mise en place des axes principaux d'une grille d'écoute. Le choix a été fait de laisser de côté les aspects hédonistes liés à l'agrément, au jugement de valeur. Trois axes ont donc été proposés par le groupe de recherche :

1. Perception du geste vocal ou de la technique vocale
2. Perception de la matérialité sonore
3. Perception de l'interprétation

Par la mise en avant de la perception, nous voulons mettre l'accent sur le fait que l'on ne prétend pas décrire le geste vocal ou ses caractéristiques sonores, mais uniquement la perception que nous en avons. Cette perception peut parfois être éloignée de la réalité physiologique ou physique de la production vocale. C'est par exemple le cas des termes « en avant » ou « en arrière » utilisés couramment dans le chant, qui réfèrent à des impressions de placement sans qu'il y ait de lien démontré avec une réalité physiologique (Vurma and Ross, 2003).

Dans un premier temps, l'axe portant sur la perception de l'interprétation a été écarté, et les efforts ont été concentrés sur les deux premiers, et en particulier sur l'axe portant sur la « Perception du geste vocal ou de la technique vocale ». Les Figure 2 et Figure 3 présentent la version (en français) de la grille d'écoute dirigée selon ces deux axes. Chaque axe se divise en pôles sur lesquels l'auditeur peut porter son attention. Les termes descriptifs proposés dans chaque pôle ont été sélectionnés au fil des discussions et des écoutes effectuées par le groupe de recherche, comme étant des termes non ambigus et représentatifs. Les termes redondants ou imprécis ont été écartés.

### Ecoute dirigée selon l'axe « Perception du geste vocal ou de la technique vocale »

<b>DYNAMIQUE INSPIRATOIRE ET EXPIRATOIRE</b>			<b>VIBRATO</b>		
inspiration sonore bruitée <input type="checkbox"/> voisée <input type="checkbox"/>	inspiration silencieuse <input type="checkbox"/>	pauses inspiratoires fréquentes <input type="checkbox"/> rares <input type="checkbox"/>	sans vibrato <input type="checkbox"/> adapté <input type="checkbox"/> inadapté <input type="checkbox"/>	lent <input type="checkbox"/> <input type="checkbox"/> <input type="checkbox"/> <input type="checkbox"/> rapide	retenu <input type="checkbox"/> trémolo <input type="checkbox"/> sans vibrato <input type="checkbox"/> ample <input type="checkbox"/> chevrottement <input type="checkbox"/>
fréquence des pauses inspiratoires : rapprochées <input type="checkbox"/> <input type="checkbox"/> <input type="checkbox"/> <input type="checkbox"/> éloignées équilibre (de la dépense d'air) : équilibrée <input type="checkbox"/> <input type="checkbox"/> <input type="checkbox"/> non-équilibrée			variations maîtrisées <input type="checkbox"/> non-maîtrisées <input type="checkbox"/>		
<b>DYNAMIQUE VIBRATOIRE</b>			<b>mécanisme laryngé</b>		
<b>transitoires d'attaque</b>			adapté <input type="checkbox"/> inadapté <input type="checkbox"/>		
attaque pondérée <input type="checkbox"/>	attaque glottique : modérée <input type="checkbox"/> coup de glotte <input type="checkbox"/>	attaque en M0 <input type="checkbox"/>	M0 <input type="checkbox"/> M1 <input type="checkbox"/> M2 <input type="checkbox"/> M3 <input type="checkbox"/>	maintien <input type="checkbox"/> variations maîtrisées <input type="checkbox"/> non-maîtrisées <input type="checkbox"/>	
attaque soufflée <input type="checkbox"/>		précision fréquentielle de l'attaque juste <input type="checkbox"/> par le haut <input type="checkbox"/> par le bas <input type="checkbox"/>	justesse : juste <input type="checkbox"/> manque de justesse <input type="checkbox"/>		
<b>transitoires de fin</b>			<b>articulation mélodique</b>		
fin pondérée <input type="checkbox"/>	fin glottique : modérée <input type="checkbox"/> accentuée <input type="checkbox"/>	fin en M0 <input type="checkbox"/>	port de voix <input type="checkbox"/> lié <input type="checkbox"/> détaché <input type="checkbox"/>		
fin soufflée <input type="checkbox"/>		précision fréquentielle de la fin juste <input type="checkbox"/> chute vers le bas <input type="checkbox"/>	respect du rythme : dans le rythme <input type="checkbox"/> <input type="checkbox"/> <input type="checkbox"/> <input type="checkbox"/> hors du rythme		
<b>PLACEMENT</b>					
en avant <input type="checkbox"/>	ouvert <input type="checkbox"/>	laryngé <input type="checkbox"/>	nasal: un peu <input type="checkbox"/> très <input type="checkbox"/>	serrée <input type="checkbox"/>	
en arrière <input type="checkbox"/>	couvert <input type="checkbox"/>	pharyngé (engorgé) <input type="checkbox"/>	nasillard <input type="checkbox"/>	soufflée <input type="checkbox"/>	
placement vocalique ou extra-vocalique : proche du "parlé" <input type="checkbox"/> <input type="checkbox"/> <input type="checkbox"/> <input type="checkbox"/> proche du "chanté"					

Figure 2: Grille d'écoute orientée selon l'axe "Perception du geste vocal ou de la technique vocale".

Un premier pôle de cet axe d'écoute concerne la **dynamique inspiratoire et expiratoire**. Deux types d'inspiration sont distingués : l'inspiration silencieuse et l'inspiration sonore. L'inspiration sonore peut être bruitée, quand la prise d'air s'accompagne d'un bruit de souffle, ou voisée, quand la prise d'air s'accompagne d'une vibration glottique associée à un bruit de souffle. La dynamique inspiratoire et expiratoire se traduit également par la fréquence des pauses inspiratoires, qui peuvent être rapprochées ou éloignées en rapport avec la gestion globale du souffle sur la phrase. La

répartition du débit d'air sur la phrase, dénommée équilibre (de la dépense d'air) dans la grille, est également un facteur d'importance.

Un second pôle de cet axe d'écoute porte sur la **dynamique vibratoire**. Il porte en particulier sur les transitoires d'attaque et de fin auxquels notre oreille humaine est très sensible. L'attaque ou la fin du son peuvent être :

- direct, sans bruit en amont (attaque ou fin pondérée)
- précédée d'un bruit de souffle (attaque ou fin soufflée)
- avec contact abrupt entre les cordes vocales (attaque ou fin glottique modérée).  
Quand l'attaque glottique est très prononcée, on parlera de coup de glotte. Dans le cas d'un transitoire de fin, on parlera d'une fin glottique accentuée.

L'attaque directe sur la note (attaque juste) est distinguée d'une attaque par un léger glissando ascendant (attaque par le bas) ou descendant (attaque par le haut). Il en est de même pour le transitoire de fin, même s'il apparaît très singulier de trouver des fins accompagnées d'un léger glissando ascendant (fin par le haut). L'attaque du son peut parfois débiter par l'utilisation du mécanisme MO, mécanisme laryngé<sup>1</sup> synonyme des registres fry et pulse. Il en va de même pour le transitoire de fin.

L'utilisation des mécanismes laryngés est un autre aspect de la dynamique vibratoire. Il y a maintien quand le même mécanisme laryngé est utilisé pour l'intégralité du morceau musical. En cas d'utilisation successive de différents mécanismes laryngés, l'auditeur peut percevoir une bonne gestion des transitions entre mécanismes laryngés (variations maîtrisées) ou une mauvaise gestion des transitions entre mécanismes, auquel cas les passages s'entendent (variations non-maîtrisées). La justesse, l'articulation mélodique et le rythme sont également pris en compte dans la dynamique vibratoire. La mélodie peut être chantée de façon liée, détachée, ou avec un port de voix, c'est-à-dire une continuité dans les variations mélodiques.

Les aspects liés au **vibrato**, à sa présence ou son absence, à la façon dont il est mené sur la phrase musicale interviennent également dans la perception du geste vocal ou de la technique vocale. Le vibrato correspond à une modulation en fréquence et en amplitude de la vibration laryngée, ce qui se traduit par une modulation en fréquence et en intensité du son perçu. La fréquence de modulation peut être basse (vibrato lent) ou élevée (vibrato rapide). Son amplitude, ou excursion en fréquence, peut être faible (vibrato réduit) ou importante (vibrato ample). On peut percevoir un trémolo, c'est-à-dire une modulation très rapide de la fréquence laryngée, pouvant s'accompagner d'instabilités. A l'opposé, on peut percevoir un chevrottement, c'est-à-dire une modulation lente et ample de la fréquence laryngée, pouvant s'accompagner d'instabilités. Les variations de fréquence et d'amplitude du vibrato par rapport à sa gestion globale sur une phrase musicale peuvent être maîtrisées ou non maîtrisées.

La sensation par l'auditeur de **la localisation de la source acoustique, ou placement**, est un aspect important de cet axe. La source acoustique peut être perçue en avant ou en arrière de la tête, sur le larynx (laryngé), dans la gorge (pharyngé) ou dans le nez. Le caractère nasal, pour lequel on perçoit une participation de la partie postérieure des cavités nasales, est distingué du caractère nasillard, pour lequel semblent participer également la partie antérieure des cavités nasales. La voix peut être perçue avec du souffle (soufflée) ou avec une impression de tension laryngée (serrée). Enfin, la notion de couverture du son intervient bien souvent, selon que le son est perçu comme ouvert ou couvert par le chanteur.

Un dernier aspect de ce pôle concerne le placement vocalique, selon que la production vocale du chanteur semble proche de la parole (proche du parlé) ou proche du chant (proche du chanté).

---

<sup>1</sup> voir Roubeau 1991 ou Henrich 2006 pour une définition de la notion de mécanisme laryngé.

## Ecoute dirigée selon l'axe « Perception de la matérialité sonore »

ASPECTS PHONETIQUES			
<b>Segmental :</b>			
voyelles	<i>proches</i> <input type="checkbox"/> <input type="checkbox"/> <input type="checkbox"/> <input type="checkbox"/>	<i>contrastées</i> <input type="checkbox"/> <input type="checkbox"/> <input type="checkbox"/> <input type="checkbox"/>	
	<i>identifiables</i> <input type="checkbox"/> <input type="checkbox"/> <input type="checkbox"/> <input type="checkbox"/>	<i>non-identifiables</i> <input type="checkbox"/> <input type="checkbox"/> <input type="checkbox"/> <input type="checkbox"/>	
consonnes	<i>courtes</i> <input type="checkbox"/> <input type="checkbox"/> <input type="checkbox"/> <input type="checkbox"/>	<i>longues</i> <input type="checkbox"/> <input type="checkbox"/> <input type="checkbox"/> <input type="checkbox"/>	
	<i>prononciation effacée</i> <input type="checkbox"/> <input type="checkbox"/> <input type="checkbox"/> <input type="checkbox"/>	<i>prononciation renforcée</i> <input type="checkbox"/> <input type="checkbox"/> <input type="checkbox"/> <input type="checkbox"/>	
<b>Suprasegmental :</b>			
respect de la phrase	<i>oui</i> <input type="checkbox"/> <input type="checkbox"/> <input type="checkbox"/> <input type="checkbox"/>	<i>non</i> <input type="checkbox"/> <input type="checkbox"/> <input type="checkbox"/> <input type="checkbox"/>	<i>N/A</i> <input type="checkbox"/>
respect des accents	<i>oui</i> <input type="checkbox"/> <input type="checkbox"/> <input type="checkbox"/> <input type="checkbox"/>	<i>non</i> <input type="checkbox"/> <input type="checkbox"/> <input type="checkbox"/> <input type="checkbox"/>	<i>N/A</i> <input type="checkbox"/>
<b>Intelligibilité</b>	<i>oui</i> <input type="checkbox"/> <input type="checkbox"/> <input type="checkbox"/> <input type="checkbox"/>	<i>non</i> <input type="checkbox"/> <input type="checkbox"/> <input type="checkbox"/> <input type="checkbox"/>	<i>N/A</i> <input type="checkbox"/>

VIBRATO		
sans vibrato <input type="checkbox"/>		
adapté <input type="checkbox"/>		
inadapté <input type="checkbox"/>		
<i>lent</i> <input type="checkbox"/> <input type="checkbox"/> <input type="checkbox"/> <input type="checkbox"/>	<i>rapide</i> <input type="checkbox"/> <input type="checkbox"/> <input type="checkbox"/> <input type="checkbox"/>	
sans vibrato <input type="checkbox"/>	retenu <input type="checkbox"/> ample <input type="checkbox"/>	trémolo <input type="checkbox"/> chevrotement <input type="checkbox"/>

COULEUR SONORE	
Timbre :	<i>aigu</i> <input type="checkbox"/> <input type="checkbox"/> <input type="checkbox"/> <input type="checkbox"/>
	<i>grave</i> <input type="checkbox"/> <input type="checkbox"/> <input type="checkbox"/> <input type="checkbox"/>
	<i>timbré</i> <input type="checkbox"/> <input type="checkbox"/> <input type="checkbox"/> <input type="checkbox"/>
	<i>détimbré</i> <input type="checkbox"/> <input type="checkbox"/> <input type="checkbox"/> <input type="checkbox"/>
	<i>équilibré</i> <input type="checkbox"/> <input type="checkbox"/> <input type="checkbox"/> <input type="checkbox"/>
	<i>déséquilibré</i> <input type="checkbox"/> <input type="checkbox"/> <input type="checkbox"/> <input type="checkbox"/>
	<i>homogène</i> <input type="checkbox"/> <input type="checkbox"/> <input type="checkbox"/> <input type="checkbox"/>
	<i>non-homogène</i> <input type="checkbox"/> <input type="checkbox"/> <input type="checkbox"/> <input type="checkbox"/>
sombre <input type="checkbox"/>	clair <input type="checkbox"/>
Autres qualificatifs: .....	

<b>Perception de la fréquence fondamentale :</b>	
<i>adaptée</i> <input type="checkbox"/> <input type="checkbox"/> <input type="checkbox"/> <input type="checkbox"/>	<i>inadaptée</i> (par rapport au tempérament, par rapport à la justesse de l'orchestre) <input type="checkbox"/> <input type="checkbox"/> <input type="checkbox"/> <input type="checkbox"/>
justesse absolue : <i>adaptée</i> <input type="checkbox"/> <input type="checkbox"/> <input type="checkbox"/> <input type="checkbox"/>	<i>inadaptée</i> <input type="checkbox"/> <input type="checkbox"/> <input type="checkbox"/> <input type="checkbox"/>
justesse relative: <i>adaptée</i> <input type="checkbox"/> <input type="checkbox"/> <input type="checkbox"/> <input type="checkbox"/>	<i>inadaptée</i> (justesse de timbre – justesse de hauteur) <input type="checkbox"/> <input type="checkbox"/> <input type="checkbox"/> <input type="checkbox"/>
<b>Sensation d'intensité vocale :</b>	
efficacité : <i>très efficace</i> <input type="checkbox"/> <input type="checkbox"/> <input type="checkbox"/> <input type="checkbox"/>	<i>peu efficace</i> <input type="checkbox"/> <input type="checkbox"/> <input type="checkbox"/> <input type="checkbox"/>
puissance : <i>très puissant</i> <input type="checkbox"/> <input type="checkbox"/> <input type="checkbox"/> <input type="checkbox"/>	<i>peu puissant</i> <input type="checkbox"/> <input type="checkbox"/> <input type="checkbox"/> <input type="checkbox"/>
formant du chanteur : <i>présent</i> <input type="checkbox"/> <input type="checkbox"/> <input type="checkbox"/> <input type="checkbox"/>	<i>absent</i> <input type="checkbox"/> <input type="checkbox"/> <input type="checkbox"/> <input type="checkbox"/>
<b>Rapport entre la dynamique et la hauteur (possibilité dynamique) :</b>	
<i>grave</i> : <i>grande dynamique</i> <input type="checkbox"/> <input type="checkbox"/> <input type="checkbox"/> <input type="checkbox"/>	<i>faible dynamique</i> <input type="checkbox"/> <input type="checkbox"/> <input type="checkbox"/> <input type="checkbox"/>
	<i>N/A</i> <input type="checkbox"/>
<i>aigu</i> : <i>grande dynamique</i> <input type="checkbox"/> <input type="checkbox"/> <input type="checkbox"/> <input type="checkbox"/>	<i>faible dynamique</i> <input type="checkbox"/> <input type="checkbox"/> <input type="checkbox"/> <input type="checkbox"/>
	<i>N/A</i> <input type="checkbox"/>

Figure 3: Grille d'écoute orientée selon l'axe "Perception de la matérialité sonore"

Un premier pôle de cette grille d'écoute orientée sur la perception de la matérialité sonore porte sur les **aspects phonétiques** au niveau segmental et suprasegmental. Au niveau segmental, il permet de mettre l'accent sur la perception du contraste vocalique (voyelles perceptivement proches ou contrastées) et de l'identification vocalique (voyelles facilement identifiables ou non identifiables), sur la perception de la gestion consonantique (consonnes de durées courtes ou longues) et sur la prononciation des consonnes (prononciation effacée ou renforcée). Au niveau suprasegmental, l'accent est mis sur le respect de la phrase et de ses accents. L'intelligibilité de la phrase est également un paramètre pris en compte.

Un second pôle porte sur la **couleur sonore**, principalement le timbre. Les aspects de timbre abordés sont les suivants : aigu/grave, timbré/détimbré, équilibré/déséquilibré au niveau de la répartition spectrale de l'énergie, homogène/non homogène sur la phrase musicale. Les caractères clair et sombre sont également proposés.

Le troisième pôle porte sur les aspects liés à **la dynamique** et à **la hauteur**. Sont abordés dans ce pôle les aspects de justesse absolue et relative, et de leur adéquation par rapport au tempérament et à la justesse de l'orchestre. Les aspects liés à la **sensation d'intensité vocale** sont également abordés, en termes d'efficacité (voix qui porte ou pas), de puissance (voix sonore ou peu sonore), et en rapport avec la présence ou l'absence perçue du formant du chanteur. Les possibilités dynamiques de la voix perçue, en particulier le **rapport entre la dynamique et la hauteur**, sont également mentionnées par rapport au grave et à l'aigu de la tessiture.

### Exploration du consensus

Les grilles d'écoute présentées dans la partie précédente ont été établies pour permettre une écoute orientée de la qualité d'une voix, en vue de pouvoir partager une description consensuelle entre spécialistes de différents domaines. L'intérêt d'utiliser une telle grille et le caractère consensuel de la description par rapport à la grille ont été testés sur un groupe de 18 auditeurs (âge moyen 38 ans (+/- 11)), dont 8 musiciens professionnels, 8 musiciens amateurs, et 2 non musiciens. Seuls 3 sujets ne pratiquaient pas le chant en amateur ou professionnel. 13 étaient familiers avec la technique lyrique, soit qu'ils la pratiquent régulièrement en tant que chanteur, ou qu'ils écoutent souvent ce type de

musique. Tous sauf un verbalisaient la qualité vocale, 10 épisodiquement et 7 très souvent.

### Déroulement du test perceptif

Le test s'est déroulé dans un auditorium, par petits groupes de 5 à 8 sujets. Le test était composé de trois parties :

1. écoute et verbalisation libre des exemples sonores
2. présentation des travaux du groupe de réflexion sur la qualité vocale, et description des deux grilles d'écoute. La grille d'écoute sur la perception du geste vocal ou de la technique vocale a été illustrée perceptivement à partir d'une base d'exemples sonores prototypiques élaborée par le groupe.
3. écoute des exemples sonores de la partie 1 du test, et caractérisation à partir des deux grilles d'écoute. Les sujets devaient cocher les cases qui leur semblaient appropriées. Quand un caractère leur semblait apparaître de façon intermittente sur la phrase musicale, la mention pouvait en être faite dans la case à cocher.

A la fin du test, une fiche était à remplir par le sujet. Elle portait sur les compétences et connaissances musicales, ainsi que sur les impressions par rapport au test et à l'utilité de la grille d'écoute.

Six exemples sonores étaient proposés, correspondant à deux chanteurs masculins interprétant chacun trois exemples de qualité vocale différente : un exemple qualifié de référence et deux variantes. Le premier chanteur (B1) interprétait une phrase musicale en français qui avait été élaborée pour un précédent travail sur la qualité vocale (Sotiropoulos, 2004). L'enregistrement de ce chanteur a été effectué préalablement à ce test, et les variations de qualité vocale ont été choisies à partir d'éléments de la grille d'écoute portant sur la perception du geste vocal ou de la technique vocale. Le second chanteur (B2) interprétait une phrase musicale en latin, i.e. les premières mesures de l'Ave Maria de Gounod, enregistrée lors d'une étude précédente sur la qualité vocale (Henrich, 2001).

Deux modes d'écoute étaient testés. L'exemple qualifié de référence était présenté en premier et de façon isolée aux sujets, tandis que les variantes étaient présentées en référence au premier exemple (écoute de paires d'exemples sonores). L'exemple à décrire était répété autant de fois que nécessaire.

### Impression des sujets sur la grille d'écoute

Dans la fiche à remplir par les sujets en fin de test, la question suivante était posée : *Pensez-vous, suite à ce test, qu'une telle grille d'écoute sélective puisse vous aider dans la perception et la verbalisation de la qualité vocale ? oui, beaucoup ; oui, un peu ; non, très peu ; non, pas du tout*

83% des sujets ont considéré que la grille pourrait leur apporter une aide : 39% des sujets ont sélectionné *oui, beaucoup*, et 44% *oui, un peu*. Deux sujets (11%) étaient sans avis, et un sujet (6%) considérait que la grille ne lui apporterait que très peu.

La question suivante concernait le consensus : *Pensez-vous, suite à ce test, qu'une grille d'écoute sélective puisse permettre un dialogue plus consensuel sur la qualité vocale entre les différents spécialistes de la voix ? oui, sûrement ; oui, peut-être ; non*

Tous les sujets ont pensé qu'une telle grille d'écoute permettrait un dialogue plus consensuel : 6 sujets (33%) pour *oui peut-être* et 12 sujets (67%) pour *oui sûrement*.

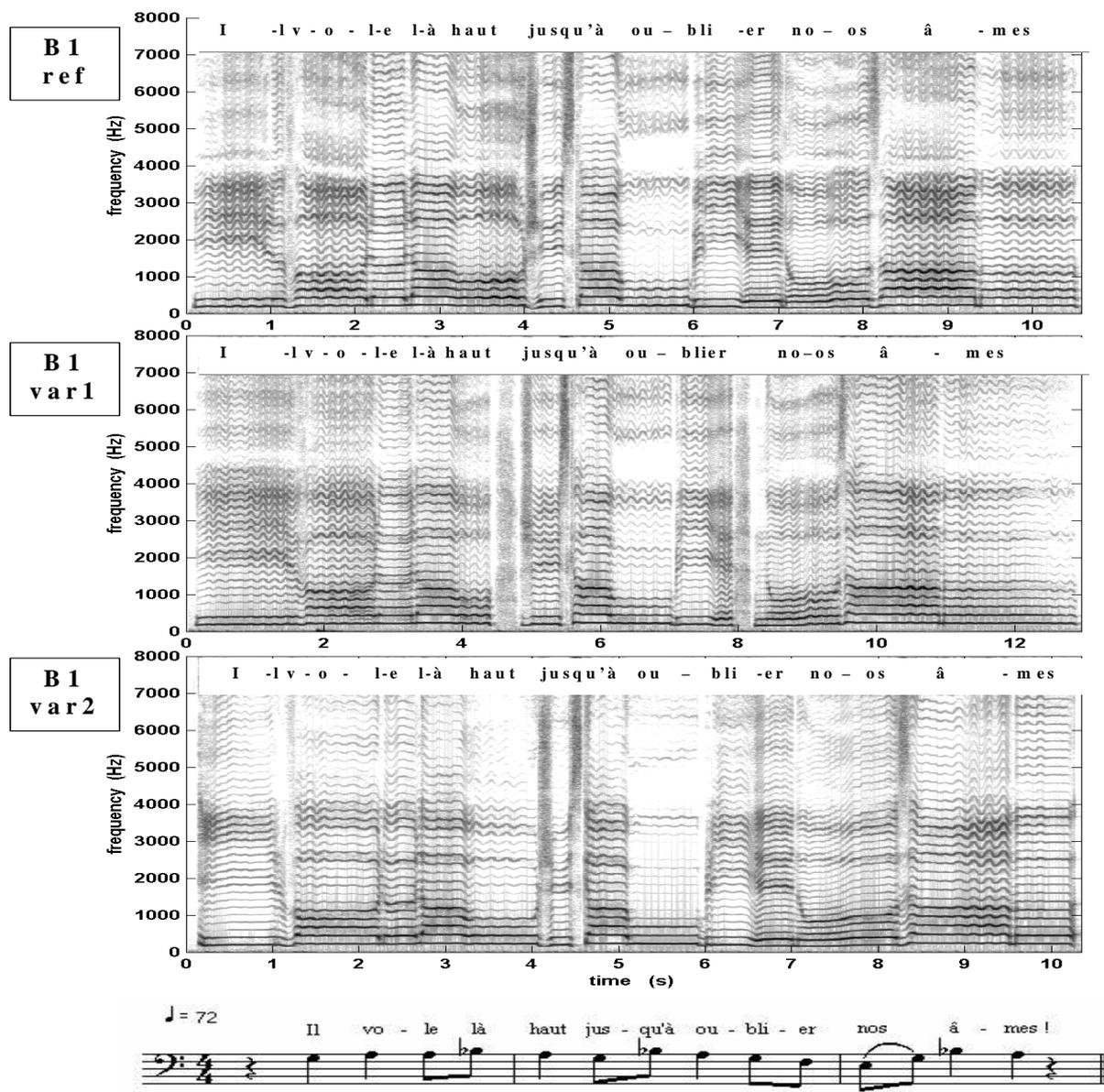


Figure 4: Phrase musicale "Il vole là-haut jusqu'à oublier nos âmes" chanté par un baryton avec trois qualités vocales différentes.

**Détection des caractéristiques saillantes**

Dans un premier temps, il nous a semblé intéressant de regarder si des locuteurs pourraient percevoir des caractéristiques de qualité vocale choisies au préalable et interprétées par le chanteur. Le premier chanteur a donc enregistré, en plus de sa production normale (la référence, cf. Figure 4 B1 ref), deux exemples dans lequel il lui était imposé de varier divers paramètres :

- variation 1 (cf. Figure 4, B1 var1) : des inspirations sonores bruitées, avec des pauses inspiratoires fréquentes et un déséquilibre marqué de la dépense d'air, des transitoires d'attaque et de fin soufflés, une production vocale en M1, une articulation mélodique détachée et hors du rythme, un placement soufflé.
- variation 2 (cf. Figure 4, B1 var2) : des inspirations silencieuses, avec des pauses inspiratoires rares, sans vibrato, avec des attaques en coup de glotte et des fins glottiques accentuées, une production vocale en M1, une articulation mélodique avec port de voix et dans le rythme, un placement laryngé et serré.

Ces caractéristiques ont bien été perçues et verbalisées à travers la grille par une majorité des auditeurs. Les résultats sont présentés sur la Figure 5.

A l'écoute de l'exemple 1, 89% des sujets considèrent que les inspirations sont sonores et bruitées, et que les pauses inspiratoires sont fréquentes (avis partagé par 83% des sujets). Le déséquilibre de la dépense d'air est remarqué (89%). Les transitoires d'attaque sont plutôt perçus comme glottiques modérés (44%) que soufflés (33%). Par contre, une majorité des sujets (55%) perçoivent les fins soufflées. L'utilisation du mécanisme laryngé M1 est bien perçue (78%). Les caractéristiques de l'articulation mélodique ne sont pas décrites de façon consensuelle entre les sujets : 39% perçoivent la mélodie produite de façon liée et 39% la perçoivent comme produite de façon détachée. 67% des sujets perçoivent que le chanteur semble relativement hors du rythme. Le caractère soufflé n'est pas perçu de façon saillante par les sujets : seuls 39% le mentionnent.

A l'écoute de l'exemple 2, 67% des sujets perçoivent des inspirations silencieuses, accompagnées de pauses inspiratoires rares (78%). La moitié seulement des sujets (50%) notent l'absence de vibrato. Ceci pourrait s'expliquer par le fait que le chanteur utilise un vibrato en toute fin de sa phrase, et que donc ce vibrato, même exprimé très brièvement, peut être perçu et pris en compte par les auditeurs. Les sujets qui n'ont pas choisi d'exprimer l'absence de vibrato, ont tous mentionné un vibrato inadapté (44%), et/ou un vibrato retenu (50%). Les attaques en coup de glotte sont perçues de façon quasi unanime (94%), et les fins glottiques accentuées sont également bien perçues (67%). L'utilisation du mécanisme laryngé M1 est bien perçue (83%), ainsi que le port de voix (78%). L'aspect rythmique n'est pas perçu de façon consensuelle, et la plupart des sujets ne détectent pas de placement laryngé (33%). Ils perçoivent un placement serré (67%) et en avant (56%).

### Discussion sur la consensualité pour la description de la qualité vocale

Les réponses des sujets sont illustrées graphiquement sur la Figure 5 pour le chanteur B1 et sur la Figure 6 pour le chanteur B2. Ces figures présentent, pour chaque paramètre de la grille, le nombre de sujets (ramené en pourcentage du nombre total de sujets) qui ont choisi de cocher ce paramètre. On peut considérer qu'il y a un consensus entre les sujets sur la description quand plus de la moitié d'entre eux (plus de 50%) ont coché la même case. Dans le cas de choix multiples (échelle de 1 à 5 proposée pour certains paramètres), la proportion pour chaque choix est donnée. Dans ce cas, la variation des couleurs et leurs proportions respectives renseignent sur le caractère partagé ou non de la description.

En analysant l'ensemble des réponses des sujets pour les six écoutes, nous pouvons voir apparaître des paramètres qui semblent utilisés de façon consensuelle par les sujets, tandis que d'autres montrent un désaccord entre sujets.

La **dynamique inspiratoire et expiratoire** est décrite de façon consensuelle par les sujets. Au sein de ce pôle, seul l'équilibre de la dépense d'air n'est pas toujours perçue de façon consensuelle, en particulier pour l'exemple B1 var2, et B2 var2.

Au sein du pôle sur la **dynamique vibratoire**, la description de la perception des transitoires d'attaque et de fin est souvent partagée par les sujets, et d'autant plus s'il n'y a pas de variation du caractère de ces transitoires sur la phrase (exemples B1ref et B2ref).

Le mécanisme laryngé utilisé est souvent décrit de façon consensuelle pour ces exemples, mais les aspects adapté/inadapté ne sont pas consensuels. Il est également intéressant de noter que des sujets qui choisissent d'exprimer le maintien du mécanisme laryngé mentionnent également parfois des variations. Il semblerait donc que la notion de variations maîtrisées ou non du mécanisme laryngé n'ait pas été comprise par ces sujets. Dans les exemples présentés, il n'y avait pas de variation du mécanisme laryngé, ces chanteurs chantant toujours en M1.

La description de la justesse est généralement partagée, excepté dans le cas des exemples B1 var2 et B2 var2. C'est également le cas de l'articulation mélodique qui est généralement décrite de façon consensuelle sauf pour l'exemple B1 var1.

Le rythme est un paramètre dont la description varie beaucoup entre les sujets. Tous y sont sensibles, puisqu'ils le sélectionnent de façon quasi-unanime à chaque écoute (entre 78% et 100%). Néanmoins, dans de nombreux cas, l'adéquation au rythme est perçue de façon très différente (exemples B1 var2, B2 var1 et var2). L'absence d'accompagnement musicale pourrait expliquer ce désaccord.

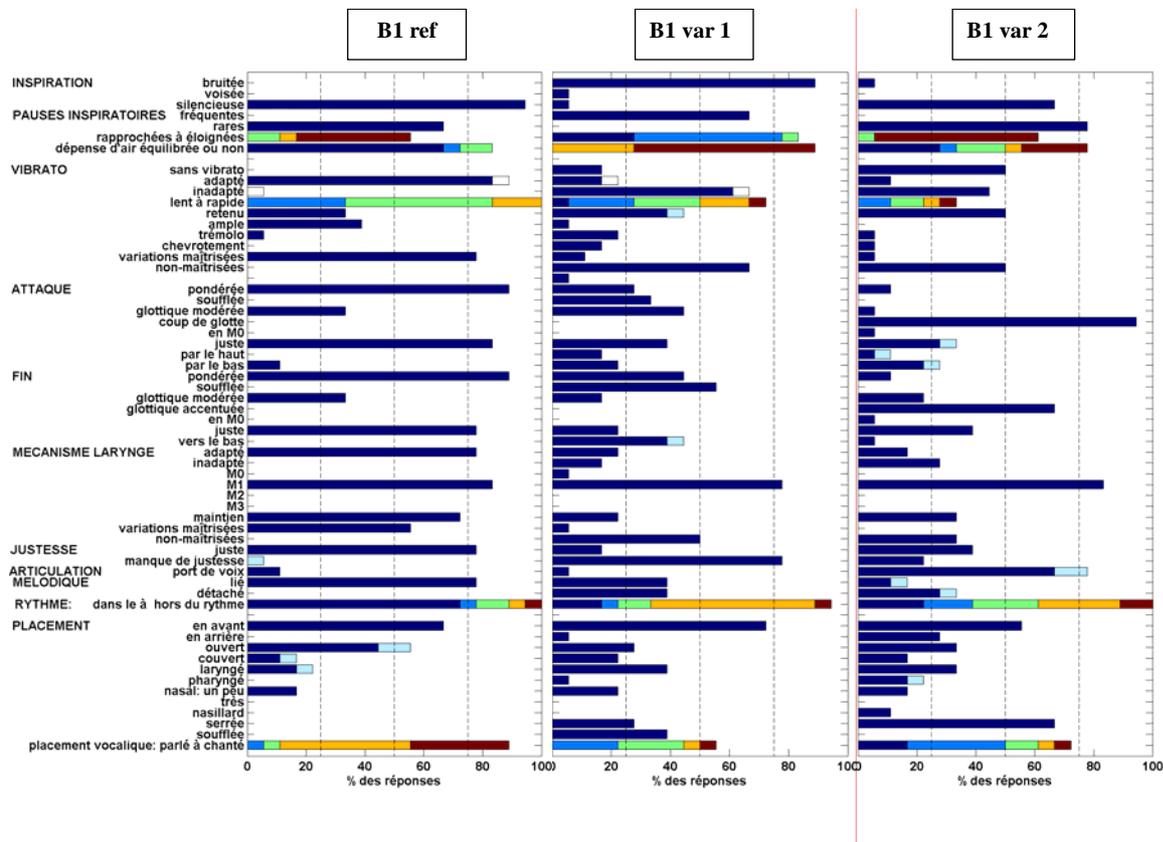


Figure 5: Résultat de la description des 3 exemples chantés par le chanteur B1 par 18 sujets selon l'axe perceptif sur le geste vocal ou la technique vocale. Pour chaque paramètre de la grille d'écoute, les barres horizontales bleu foncé présentent le pourcentage des sujets qui ont mentionné ce paramètre. Les compléments en bleu très clair correspondent aux cas où le sujet perçoit le caractère par intermittence. Les compléments transparents correspondent aux cas où le sujet considère que le paramètre ne peut pas être jugé dans le contexte proposé (n/a). Les barres de plusieurs couleurs présentent les réponses sur des échelles bipolaires à 5 points : 1. bleu foncé, 2. bleu clair, 3. vert, 4. orange et 5. rouge foncé).

La description du **vibrato** n'est pas souvent unanime entre les sujets. Si le caractère adapté ou inadapté du vibrato semble partagé (sauf pour l'exemple B1 var2 qui est perçu soit comme sans vibrato, soit avec un vibrato inadapté), la rapidité du vibrato est souvent perçue de façon très différente entre les sujets. Le caractère ample ou retenu est également rarement consensuel. La perception de la maîtrise ou non des variations de vibrato est partagée.

La description de la **perception du placement** n'est pas partagée de façon claire par les sujets. Dans le cas d'une caractéristique prépondérante (en avant ou serrée), on peut observer un bon accord dans la description des sujets (exemples B1 ref, B1 var1, var2 et B2 var2). Néanmoins, bien souvent, les descriptions des sujets sont multiples, et parfois divergentes (comme c'est souvent le cas pour le caractère ouvert/couvert). Cette difficulté de trouver un consensus sur la description du placement avait déjà apparu de façon très nette dans les discussions et verbalisation du groupe de recherche. Cela pourrait s'expliquer par le fait que le placement n'est pas une notion clairement définie. Une étude plus approfondie de ce point nous paraît nécessaire.

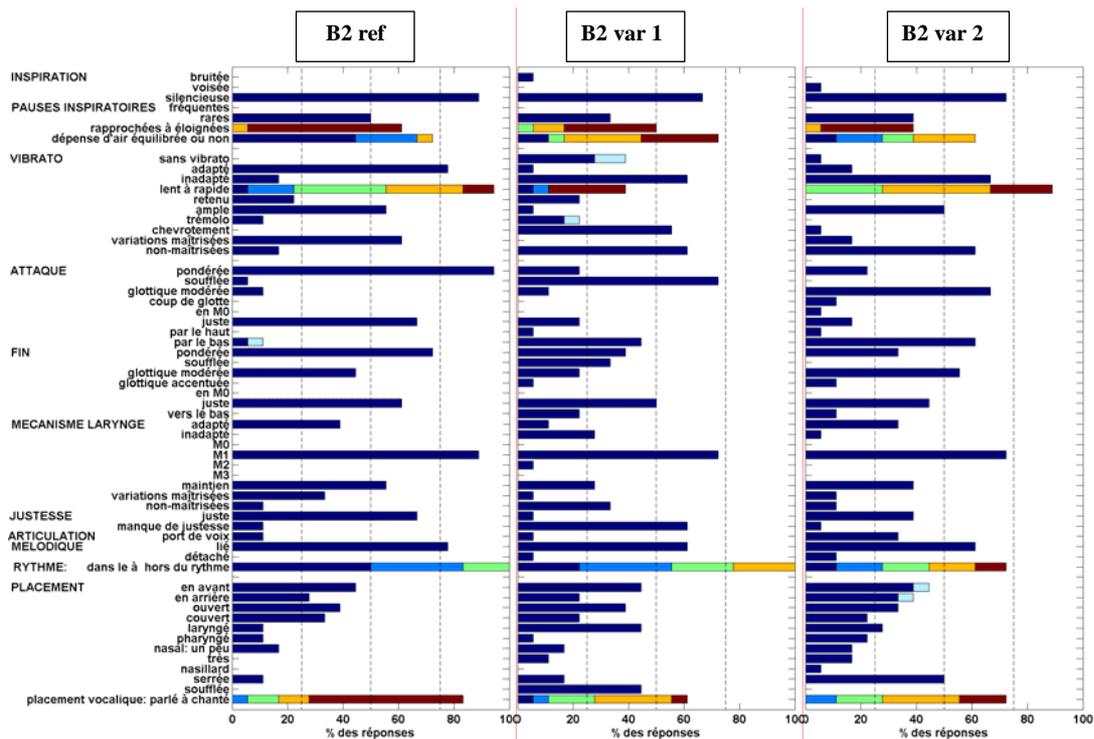


Figure 6 : Résultat de la description effectuée par 18 sujets selon l'axe perceptif sur le geste vocal ou la technique vocale, pour les trois exemples chantés par le chanteur B2. Se reporter à la Figure 5 pour plus de détails.

## Conclusion

Dans cette étude, nous avons présenté le résultat des réflexions d'un groupe de recherche pluridisciplinaire sur la perception de la qualité vocale et sa verbalisation dans le chant lyrique occidental. Dans sa recherche d'un langage de description commun à plusieurs domaines d'expertise, ce groupe a été amené à élaborer une grille d'écoute sélective permettant de décrire la perception de la qualité vocale au travers de deux axes : perception du geste vocal ou de la technique vocale, et perception de la matérialité sonore. L'usage de cette grille a été testé sur une population de 18 sujets venant d'horizons disciplinaires différents. L'utilité d'une telle approche semble unanimement reconnu par les sujets. L'analyse de leurs réponses par rapport au premier axe montre qu'un bon consensus est obtenu sur la perception des dynamiques respiratoire et vibratoire. Par contre la description de la perception du vibrato et du placement vocal montre une grande diversité entre les sujets, ce qui pourrait s'expliquer par des référents en mémoire non partagés, différents modes d'écoute, ou des notions mal définies, et appelle à des recherches futures.

Outre le fait que cette grille d'écoute permet de progresser dans la recherche d'une terminologie consensuelle, elle constitue un outil pédagogique intéressant. D'une part, elle permet de se familiariser avec la catégorisation des différents paramètres de la qualité vocale, et d'autre part, elle guide l'identification et l'évaluation perceptuelle de ces paramètres lors de l'écoute d'une voix. Elle peut enfin constituer un précieux support de discussion entre les experts des différentes disciplines gravitant autour de la voix.

## Références bibliographiques

- Castellengo, M. (1986). Les sources acoustiques, in D Mercier (Ed.). Le livre des techniques du son, tome 1, Paris, Editions Fréquences, pp. 45-80.  
 Dubois, D. (1991). Prototypes et typicalité. In Sémantique et cognition.  
 Dubois, D. (1997). Catégorisation et cognition : de la perception au discours. Collectif, Kimé Ed., Paris.

- Dubois, D. (2000). Categories as acts of meaning: the case in olfaction and audition. *Cognitive Science Quarterly*, vol.1, 35-68.
- Ekholm E., Papagiannis G.C. and Chagnon F.P. (1998). *Relating objective measurements to expert evaluation of voice quality in western classical singing: Critical perceptual parameters*. *Journal of Voice* 12(2): 182-196.
- Faure, A. (2000). Des sons aux mots : comment parle-t-on du timbre musical ? Thèse de doctorat. EHES.
- Garnier, M., Henrich, N., Dubois, D., Castellengo, M., Poitevineau, J., and Sotiropoulos, D. (2005). Etude de la qualité vocale dans le chant lyrique. *Scolia* 20, pp.151-169.
- Guerin, C. (2006) Glossaire du chant lyrique. Website : <http://chanteur.net/glossair.htm>
- Guinard, J.X., Noble, A.C. (1986). Proposition d'une terminologie pour une description analytique de l'arôme des vins. *Sci. Aliment* 6, pp.657-662.
- Henrich N. (2001). Etude de la source glottique en voix parlée et chantée : modélisation et estimation, mesures acoustiques et électroglottographiques, perception. PhD thesis, Université Paris 6, 2001.
- Henrich N. (2006). Mirroring the voice from Garcia to the present day: Some insights into singing voice registers, *Logopedics Phoniatrics Vocology*, vol. 31, pp. 3-14.
- Isshiki N., Takeuchi (1970). Factor analysis of hoarseness, *Stud. Phonol.*, 5, 37-44.
- Hirano M. (1981). *Clinical examination of voice*, Springer Verlag, New York.
- Hirano M. (1989). Objective evaluation of human voice: clinical aspects. *Folia Phoniatr.*, 41, 89-144.
- Miller, R. (1986). *The structure of singing: system and art in vocal technique*. Ed Schirmer/Thomson Learning.
- Philippe, F., Schacher, L., Adolphe, D., Dacremont, C. (2001) Développement d'une méthodologie d'analyse sensorielle tactile des textiles. *International Journal of Clothing* 15, pp. 268-. 275.
- Roubeau B. (1993) Mécanismes vibratoires laryngés et contrôle neuro-musculaire de la fréquence fondamentale, PhD thesis, Université Paris-Orsay, Orsay.
- Schaeffer, P. (1966). *Traité des objets musicaux*. Ed du seuil, Paris, pp. 712.
- Sotiropoulos, D. (2004). Analyse acoustique et catégorisation d'un ensemble de qualités vocales pertinent pour la description de voix lyriques masculines. Mémoire DEA ATIAM.
- Titze, I. R. (1995). Definitions and nomenclature related to voice quality. in *Vocal Fold Physiology*, edited by Fujimura O. and Hirano M. (Singular, San Diego), pp. 335-342.
- Vennard, W. (1967). *Singing, the mechanism and the technic*. Ed. Carl Fisher.
- Vurma A, Ross J. (2003). *The perception of 'forward' and 'backward placement' of the singing voice*. *Logoped Phoniatr Vocol*. Vol. 28(1):19-28.
- Wapnick J. and Ekholm E. (1997). *Expert consensus in solo voice performance evaluation*. *Journal of Voice* 11(4): 429-436.
- Wendler J., Rauhut A., Krüger H. (1986). Classification of voice qualities. *J. Phonet.*, 14, 483-488