



HAL
open science

Towards a Common Terminology to Describe Voice Quality in Western Lyrical Singing: Contribution of a Multidisciplinary Research Group

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

► To cite this version:

Nathalie Henrich Bernardoni, Pascal Bezard, Robert Expert, Maëva Garnier, Christian Guerin, et al.. Towards a Common Terminology to Describe Voice Quality in Western Lyrical Singing: Contribution of a Multidisciplinary Research Group. *Journal of interdisciplinary music studies*, 2008, 2 (1&2), pp.71-93. hal-00297248

HAL Id: hal-00297248

<https://hal.science/hal-00297248>

Submitted on 15 Jul 2008

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

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

Towards a Common Terminology to Describe Voice Quality in Western Lyrical Singing: Contribution of a Multidisciplinary Research Group

Nathalie Henrich¹, Pascal Bezard, Robert Expert, Maëva Garnier², Christian Guerin,
Claire Pillot³, Sophie Quattrocchi, Bernard Roubeau⁴, Boris Terk

¹Department of Speech and Cognition, GIPSA-lab, CNRS/INPG/UJF/Univ. Stendhal

²IJLRA-Equipe LAM, CNRS/Univ. Paris 6/Minist. de la culture

³Laboratoire de Phonétique et Phonologie UMR 7018, Université Paris III Sorbonne nouvelle

⁴Service d'ORL et de Chirurgie Cervico-faciale, Hôpital Tenon

Background in performance. In the field of lyrical singing, an extensive terminology is dedicated to voice quality description. Among the many terms, some are used with consistent meaning by virtually all voice specialists, whereas others, which are more metaphorical or aesthetic, have multiple meanings despite their frequent use. The descriptors used by voice specialists deal not only with perceived sound, but also with the production of sound.

Background in acoustics. Acousticians do not have a specific vocabulary for describing vocal sounds. They often make use of terms related to timbre. Many studies conducted on the determination of physical criteria for voice-quality description imply a listening focused on voice spectral content and transient phenomena.

Aims. Perception of voice quality is subjective and depends on the listener's own experiences and expectations. However, a consensus on its verbal description can be found, in a similar way that a technical vocabulary exists for wine-tasting. Our aim is to elaborate a common terminology for voice-quality description in voice pedagogy, voice therapy and musical acoustics.

Main contribution. This paper presents a three-year study conducted by a research group composed of musical acousticians, speech therapists, singers, singing teachers and choir directors. Three main perceptual angles have been considered: perception of vocal gesture or vocal technique, perception of sound, and perception of performance. The listening sheet related to perception of vocal gesture or vocal technique is presented here, and its relevance is perceptually tested. Descriptive terms and illustrative sound examples are given, which have been selected by their consensuality within the research group. The listening sheet related to perception of sound is also briefly described.

Implications. The proposed listening sheets facilitate the perceptual and verbal description of voice quality in singing. They allow the listener to concentrate on a given aspect of voice quality, and provide voice professionals with a consensual terminology for expressing singing voice-quality perception. They may also be used as a tool for vocal pedagogy and aural training.

Keywords: Voice quality, verbal description, listening training, listening sheet

Introduction

In the field of speech processing, voice quality is defined by what differs between two vocal productions with identical lexical content. These differences can be prosodic or acoustic, related to variations of rhythm, pitch, intensity, spectral content, etc. They can be observed at all levels of speech segmentation: at the phoneme level, by a variation of a sustained-sound spectral content (e.g. dependent on vowel, register, aperiodicities); at the syllabic level, with regards to local variations in timing and spectrum (e.g. during attack and release transients); and at the level of the phrase or sentence, in relation to global variations in timing and spectra (e.g. intonation, rhythm, articulation). It is a perceptual notion: as Kreiman and Gerratt (2000: 73) write, “the acoustic signal itself does not possess quality, it evokes it in the listener”.

How can we describe the perceived quality of a voice? This question introduces a more general problem of sound description: what are the listening modes and the terms to be used to describe what we hear? Voice quality perception is complex, highly subjective and listener dependent. The listening modes and the descriptive terms vary between fields: physicians, speech therapists, actors, singers, voice teachers, voice coaches, and voice scientists share neither the same listening mode nor a common vocabulary to describe the perceived quality of a voice.

In their daily practice, physicians and speech therapists have developed a common language to describe the quality of pathological voices. Their listening is mainly oriented towards finding “defects” for diagnosis. In this field, much effort has been expended to retain the more consensual and adequate terms in order to discriminate perceptually among different voice pathologies. Perceptual evaluation scales of voice quality are commonly used, such as the GRBAS (“Grade, Roughness, Breathiness, Aesthenia, Strain”: Isshiki and Takeuchi, 1970; Hirano, 1981, 1989), the RBS scale (“Roughness, Breathiness, Hoarseness”: Wendler, Rauhut, and Krüger, 1986), or the SVEC (“Stockholm Voice Evaluation Consensus Model”: Hammarberg, 1986, 1997). In the field of phonetics, various classifications of voice quality have been suggested, one of the most commonly-used being the “Vocal Profile Analysis” (VPA) scheme proposed by Laver and his colleagues (Laver, 1980, 2000). In the field of speech pathology and therapy as well as in the field of phonetics, the proposed approaches of perceptual evaluation have all resulted from long-term collaborative projects among voice experts.

In the field of Western lyrical singingⁱ, attempts have also been made to evaluate voice quality on multiple-criteria rating scales (for instance, a 12-criteria rating scale proposed by Wapnick and Ekholm, 1997, Ekholm, Papagiannis and Chagnon, 1998). It resulted in very poor interjudge agreements. Voice experts from different backgrounds (voice pedagogy, voice therapy, and musical acoustics) seem to lack a common language to describe singing-voice quality (Ekholm, Papagiannis, and Chagnon, 1998), which may be due to the great variety of terms found in the literature (Vennard, 1967; Miller, 1986; Guerin, 2006; Garnier, Henrich, Castellengo, Sotiropoulos, and Dubois et al., 2007). Among the many descriptors used to describe voice quality, some are common to specialists whereas others, more metaphorical or

aesthetic, are characterised by either multiple or individually-defined meanings despite their frequent use (Garnier et al., 2007). The terms used by specialists deal not only with perceived sound, but also with the mode of production (Wapnick and Ekholm, 1997; Garnier et al., 2005, 2007). Imitations are also often used as a complement to verbal description.

The search for such a common language has led to the development of a multidisciplinary research group in France, in which musical acoustics researchers, voice therapists, singers and singing teachers all actively participate. The desire to find a unified terminology for a sensory object, one that would transcend disciplinary boundaries, is certainly not confined to voice. Other disciplinary fields, such as oenology (Guinard and Noble, 1986) and textile engineering (Philippe, Schacher, Adolphe, and Dacremont, 2001) have found a consensus for allowing discussions between the different specialists in these fields, mainly for commercial reasons. In this paper, we present the result of a three-year study conducted by this multidisciplinary research group. Within the framework of adult Western lyrical singing, the aims of the research group were:

1. to clarify the notion of voice quality and modes of listening,
2. to elaborate a consensual voice-quality terminology, illustrated by sound examples, in order to:
 - a. establish a consensual base to verbal exchanges between disciplines, and
 - b. train new listeners to analytical listening to voices.

Some fundamental aspects of the perception of a sensory process will first be addressed. Free verbalisation about voice-quality will be discussed, together with the terms and listening modes that emerge. On the basis of these observations, two listening sheets will be presented, and the consensus will be assessed by a listening test on the first one. In conclusion, the relevance of this approach and the proposed tool for perceptual evaluation of voice quality in Western lyrical singing will be discussed.

From perception to verbalisation

Several fundamental aspects of the perception of a sensory process have a direct consequence on its verbal description; these aspects are detailed below.

A categorical perception

Many studies have shown that perception tends first to identify an object, in order to place it within the listener's existing mental categories (Castellengo, 1986). Analytical perception may occur thereafter (Schaeffer, 1966), and the verbal descriptors are then dependent on the initial object categorisation (Dubois, 1991, 1997, 2000). For instance, we do not describe a spoken or singing voice in the same way, nor do we use the same description for a lyrical and non-lyrical singing voice. Within this

framework, the research group first had to choose the vocal-production category with which to investigate voice quality. The choice was made to work on Western adult lyrical singing (in French: *chant savant occidental de l'adulte*; the literal English translation of which is *adult Western learned singing*).

An individual perception

Qualitative listener assessments involve interpretation through the filter of a listener's mental representation. Therefore, the past experiences of each listener, their expectations and listening aims (which depend on their areas of expertise) will influence their perception and the cues to which they would pay attention. It seems necessary to guide perception, to direct the listening to shared aspects. Rapidly, the research group was led to elaborate a listening sheetⁱⁱ, for which one of the main goals was to guide perception to selected cues related to voice quality.

A differential perception

Human perception is differential: no evaluation or description is absolute. Rather, it involves comparison with another presented object or with a remembered prototype of the object category. As a consequence, verbal description of a sensory object can take advantage of comparatives, and it often involves the object's defects (or its differences from standards) rather than its positive qualities (Faure, 2000). This is even stronger in aesthetic fields such as lyrical singing, where the personal preference is part of the object and cannot be held apart during evaluation. During the elaboration of a methodology for listening and description of voice quality, the differential aspect of human perception has to be taken into account. The elaboration of shared memory reference can benefit from the training with prototypic sound objects. Therefore, the research group recorded a database of reference sound examples, which perceptually illustrates the selected voice-quality criteria.

What words best express the perceived quality of a lyrical voice?

Several glossaries are provided in the literature (e.g. Vennard, 1967; Miller, 1986; Titze, 1995), and illustrate the variability and redundancy of terms in use. Each specialist has his/her own vocabulary to speak about voice, and this vocabulary is only partly shared with the other specialists. The language (English, French, etc.) is of much importance, as a direct translation from one language to another (e.g. French to English) may not be appropriate.

To gather each expert's vocabulary and the way it is organised, a preliminary study was conducted to determine how and with which words we speak about voice quality in French. Each expert gave an unconstrained verbal description of a set of several commercial and experimental sound examples. From this exploratory phase came a rich vocabulary that was organised into categories. We based the categorisation

process on a psycholinguistic study of voice-quality verbalisation that was conducted on singing teachers (Garnier et al., 2005, 2007). In that study, French singing teachers were asked to speak freely about voice quality while listening to sound examples recorded in a laboratory. The linguistic analysis of their discourse has provided a great part of the French lexica related to the notion of voice quality and corresponding concepts. The lexica used by singing teachers have been separated into four main categories: sound (39% of the terms), technique and physiology (22.5%), pleasure and value judgments (21%), and performance (17.5%). These four categories together participate in the perception of voice quality, as a perceptual judgment integrates different criteria. They overlap to some extent, as a given criterion can sometimes be considered in different ways. For example, vibrato can be an indicator of vocal technique, an acoustical parameter of fundamental-frequency modulation, or an indicator of musical style.

Toward an oriented listening

The four main categories (sound, technique and physiology, pleasure and value judgments, performance) have inspired the choice of the major aspects of the listening sheet. Three perceptual angles have been proposed by the research group:

1. Perception of vocal gesture or vocal technique
2. Perception of sound
3. Perception of performance

These three angles recall the authors' three main fields of expertise (1. voice therapy, 2. acoustics and 3. music performance). The aspects related to pleasure judgment have not been considered, since they correspond to very subjective and variable judgments. Indeed, we aim at finding terms and criteria on which specialists from different disciplinary fields would agree to describe a voice. The aspects related to value judgements have been shown to be relatively consistent within experts of Western lyrical singing (Garnier et al., 2005, 2007), as they are based on objective criteria as well as on shared cultural conventions. Therefore they have been included in the perception of some sound and vocal technique criteria.

By giving pre-eminence to perception, we wish to make clear that we make no claim to describe vocal gesture or its acoustical characteristics, but only the perceptions that we have of these. Indeed, perception can sometimes be far away from physiological or physical realities of vocal production. The terms 'forward' and 'backward' commonly used in singing are good examples of terms that refer to placement feelings without any demonstrated link to a physiological reality (Vurma and Ross, 2003). Our approach is based on the assumption that there may be no one-to-one relationship between aspects of the physical world and the perception we have of them.

As a first step, the research group has concentrated its efforts on the elaboration and testing of the first two perceptual angles. The exploration of the third calls for further work. Figures 1 and 2 present the French version of the listening sheets for these two angles. Each listening sheet is mapped onto several areas to which labels are attached.

The listeners can focus their attention on selected areas. The descriptive terms used for each label were selected during discussions and listening tests done by the research group as the less ambiguous and most representative terms. Synonymous and imprecise terms have been discarded.

**Listening oriented from the first angle:
“perception of vocal gesture or vocal technique”**

The first listening area deals with the **dynamics of inhalation and exhalation**. Two kinds of inhalation are distinguished: sonorous and silent inhalation. Sonorous inhalation can be breathy, when air breathing involves turbulence, or voiced, when breathing has both turbulence sound and a glottal vibration. The dynamics of inhaling and exhaling are also characterised by breathing pauses, which can be frequent or infrequent. The airflow management during the phrase is important: breathing pauses can be found close together or far apart. The air consumption can be well distributed along the phrase (balanced air consumption) or not (unbalanced). In this context, the balance of air consumption is meant as a dynamic criterion over the whole musical phrase.

A second area relates to **vibratory dynamics**. This concerns the attack and final transients, to which the ear is very sensitive. A sound attack or sound end can be produced silently, with no audible noise (balanced). It can be associated with a breath noise (breathy) or with an abrupt vocal-fold contact (glottal). When the contact is marked, it could characterise a strong glottal attack (glottal stop) or glottal end. An attack at the final quiescent pitch (true) is set apart from an attack with slight upward glide (i.e. starting from a lower pitch) or downward glide. The attack can be produced in laryngeal mechanism M0iii, synonymous with vocal fry or pulse registers. The final transients can be described similarly (true, downward glide, in M0), though a final sound with slight upward glide has only rarely been observed. The use of laryngeal mechanismⁱⁱⁱ is another aspect of the vibratory dynamics. Sometimes, the same mechanism is used through the whole sentence (maintained). When different laryngeal mechanisms are used in the sentence, the listener may perceive a good control of the transition phases between mechanisms (controlled variations) or a poor one, for which transitions can be heard (uncontrolled variations). Pitch accuracy, melodic articulation, and rhythm are also considered with vibratory dynamics. The melodic line can be sung *legato*, *staccato*, or with a *portamento*, which is a continuous slide in the melodic variations.

A third area deals with **vibrato**, its presence or absence, and the way it is used in a musical phrase. Vibrato corresponds to a frequency and amplitude modulation of the laryngeal vibration, which induces pitch and loudness modulations in the perceived sound. The modulation frequency can be low (slow vibrato) or high (fast vibrato). Its amplitude, or frequency extent, can be small (restrained vibrato) or large (ample vibrato). A fast laryngeal-frequency modulation (tremolo) or a slow and ample one (quiver) can be perceived. Both cases may be associated with instabilities. The frequency and amplitude variations of vibrato can be well or poorly controlled over the musical phrase.

DYNAMIQUE INSPIRATOIRE ET EXPIRATOIRE (DYNAMICS OF INHALATION AND EXHALATION)		
inspiration sonore (sonorous inhalation) bruitée <input type="checkbox"/> voisée <input type="checkbox"/> (noisy) (voiced)	inspiration silencieuse <input type="checkbox"/> (silent inhalation)	pauses inspiratoires (breathing pauses) fréquentes (frequent) <input type="checkbox"/> rares (infrequent) <input type="checkbox"/>
fréquence des pauses inspiratoires : <i>rapprochées</i> <input type="checkbox"/> <i>éloignées</i> <input type="checkbox"/> (frequency of breathing pauses) (<i>close together</i>) (<i>far apart</i>) équilibre de la dépense d'air : <i>équilibrée</i> <input type="checkbox"/> <i>non-équilibrée</i> <input type="checkbox"/> (air consumption) : (<i>balanced</i>) (<i>unbalanced</i>)		

VIBRATO		
sans vibrato (no vibrato) <input type="checkbox"/> adapté (adequate) <input type="checkbox"/> inadapté (inadequate) <input type="checkbox"/>		
lent fâpide (slow)	(fast)	retenu <input type="checkbox"/> (restrained) ample <input type="checkbox"/> (ample)
		trémolo <input type="checkbox"/> (tremolo) chevrottement <input type="checkbox"/> (quiver)
variations maîtrisées <input type="checkbox"/> non -maîtrisées <input type="checkbox"/> (controlled) (uncontrolled)		

DYNAMIQUE VIBRATOIRE (VIBRATORY DYNAMICS)		
transitoires d'attaque (attack transients)		
attaque pondérée <input type="checkbox"/> (balanced attack)	attaque glottique (glottal attack) : modérée <input type="checkbox"/> (moderate) coup de glotte <input type="checkbox"/> (glottal stop)	attaque en M0 (attack in M0) <input type="checkbox"/> précision fréquentielle de l'attaque : (frequency accuracy of attack) juste <input type="checkbox"/> par le haut <input type="checkbox"/> par le bas <input type="checkbox"/> (true) (downward glide) (upward glide)
transitoires de fin (final transients)		
fin pondérée <input type="checkbox"/> (balanced end)	fin glottique (glottal end) : modérée <input type="checkbox"/> (moderate) accentuée <input type="checkbox"/> (glottal strong)	fin en M0 (end in M0) <input type="checkbox"/> précision fréquentielle de la fin : (frequency accuracy of end) juste <input type="checkbox"/> chute vers le bas <input type="checkbox"/> (true) (downward glide)

mécanisme laryngé (laryngeal mechanism)		adapté (adequate) <input type="checkbox"/> inadapté (inadequate) <input type="checkbox"/>	
M0 <input type="checkbox"/>	M1 <input type="checkbox"/>	M2 <input type="checkbox"/>	M3 <input type="checkbox"/>
maintien (maintained) <input type="checkbox"/> variations maîtrisées <input type="checkbox"/> non -maîtrisées <input type="checkbox"/> (controlled) (uncontrolled)			
justesse : juste <input type="checkbox"/> manque de justesse <input type="checkbox"/> (pitch accuracy) (in tune) (out of tune)			
melodic articulation		port de voix <input type="checkbox"/> (portamento)	lié (legato) <input type="checkbox"/> détaché <input type="checkbox"/> (staccato)
respect du rythme (respect of rhythm) : <i>dans le rythme</i> <input type="checkbox"/> <i>hors du rythme</i> <input type="checkbox"/> (in the rhythm) (out of rhythm)			

PLACEMENT				
en avant <input type="checkbox"/> (forward)	ouvert <input type="checkbox"/> (open)	laryngé <input type="checkbox"/> (laryngeal)	nasal : un peu <input type="checkbox"/> très <input type="checkbox"/> (a little) (a lot)	serrée ? <input type="checkbox"/> (pressed)
en arrière <input type="checkbox"/> (backward)	ouvert <input type="checkbox"/> (covered)	pharyngé <input type="checkbox"/> (pharyngeal)	nasillard <input type="checkbox"/> (twang)	soufflée ? <input type="checkbox"/> (breathy)
placement vocalique ou extra-vocalique : <i>proche du "parlé" ?????</i> <input type="checkbox"/> <i>proche du "chanté"</i> <input type="checkbox"/> (vocalic placement) (<i>close to speech</i>) (<i>close to singing</i>)				

Figure 1. Listening sheet from the first angle “Perception of vocal gesture or vocal technique”. A suggested English translation of French terms is given in bold. Illustrative sound examples were recorded by the research group for most of the criteria. They can be heard online at <http://www.musicstudies.org/>

The listener’s assessment of **acoustic source localisation, or ‘placement’**, is another important aspect of this perceptual angle. The acoustic source can be perceived as ‘forward’ or ‘backward’ in the head, in the larynx (laryngeal), in the throat (pharyngeal), or in the nose (nasal). The nasality, for which a contribution of posterior nasal cavities is perceived, is set apart from the twang quality, for which anterior nasal cavities also seem to contribute. The voice can be perceived as breathy, or as giving an impression of laryngeal tension (pressed). Covering is also part of the placement assessment (open or covered sound). Vocalic placement is mentioned, depending on whether the singer’s vocal production seems closer to speech or to singing. If importance is given to consonants and if vowels are contrasted, voice quality may be viewed as ‘close to speech’. On the contrary, if vowels are less contrasted and if the duration of vowels is extended in comparison to that of consonants, voice quality may be viewed as ‘close to singing’.

ASPECTS PHONETIQUES (PHONETIC ASPECTS)	
Segmental :	
voyelles (vowels)	proches (close) <input type="checkbox"/> <input type="checkbox"/> <input type="checkbox"/> <input type="checkbox"/> contrastées (contrasted) <input type="checkbox"/> <input type="checkbox"/> <input type="checkbox"/> <input type="checkbox"/> identifiables (recognisable) <input type="checkbox"/> <input type="checkbox"/> <input type="checkbox"/> <input type="checkbox"/> non-identifiables (non recognisable) <input type="checkbox"/> <input type="checkbox"/> <input type="checkbox"/> <input type="checkbox"/>
consonnes (consonants)	courtes (short) <input type="checkbox"/> <input type="checkbox"/> <input type="checkbox"/> <input type="checkbox"/> longues (long) <input type="checkbox"/> <input type="checkbox"/> <input type="checkbox"/> <input type="checkbox"/> prononciation efficace (unstressed pronunciation) <input type="checkbox"/> <input type="checkbox"/> <input type="checkbox"/> <input type="checkbox"/> prononciation renforcée (stressed pronunciation) <input type="checkbox"/> <input type="checkbox"/> <input type="checkbox"/> <input type="checkbox"/>
Suprasegmental :	
respect de la phrase (compliance with phrase)	oui (yes) <input type="checkbox"/> <input type="checkbox"/> <input type="checkbox"/> <input type="checkbox"/> non (no) <input type="checkbox"/> <input type="checkbox"/> <input type="checkbox"/> <input type="checkbox"/> N/A <input type="checkbox"/>
respect des accents (compliance with accents)	oui (yes) <input type="checkbox"/> <input type="checkbox"/> <input type="checkbox"/> <input type="checkbox"/> non (no) <input type="checkbox"/> <input type="checkbox"/> <input type="checkbox"/> <input type="checkbox"/> N/A <input type="checkbox"/>
Intelligibilité (intelligibility)	oui (yes) <input type="checkbox"/> <input type="checkbox"/> <input type="checkbox"/> <input type="checkbox"/> non (no) <input type="checkbox"/> <input type="checkbox"/> <input type="checkbox"/> <input type="checkbox"/> N/A <input type="checkbox"/>

COULEUR SONORE (SOUND COLOUR)	
Timbre :	
aigu (high-pitched)	<input type="checkbox"/> <input type="checkbox"/> <input type="checkbox"/> <input type="checkbox"/> grave (low-pitched) <input type="checkbox"/> <input type="checkbox"/> <input type="checkbox"/> <input type="checkbox"/>
timbré (with ring)	<input type="checkbox"/> <input type="checkbox"/> <input type="checkbox"/> <input type="checkbox"/> détimbré (without ring) <input type="checkbox"/> <input type="checkbox"/> <input type="checkbox"/> <input type="checkbox"/>
équilibré (balanced)	<input type="checkbox"/> <input type="checkbox"/> <input type="checkbox"/> <input type="checkbox"/> déséquilibré (unbalanced) <input type="checkbox"/> <input type="checkbox"/> <input type="checkbox"/> <input type="checkbox"/>
homogène (homogeneous)	<input type="checkbox"/> <input type="checkbox"/> <input type="checkbox"/> <input type="checkbox"/> non-homogène (inhomogeneous) <input type="checkbox"/> <input type="checkbox"/> <input type="checkbox"/> <input type="checkbox"/>
sombre (dark)	<input type="checkbox"/> clair (light) <input type="checkbox"/>
Autres qualificatifs (Other terms) :	

Perception de la fréquence fondamentale (PITCH PERCEPTION) :	
adaptee (adequate)	<input type="checkbox"/> <input type="checkbox"/> <input type="checkbox"/> <input type="checkbox"/> inadaptee (inadequate) <input type="checkbox"/> <input type="checkbox"/> <input type="checkbox"/> <input type="checkbox"/> (par rapport au tempérament, par rapport à la justesse de l'orchestre)
justesse absolue (perfect pitch)	: adaptee (adequate) <input type="checkbox"/> <input type="checkbox"/> <input type="checkbox"/> <input type="checkbox"/> inadaptee (inadequate) <input type="checkbox"/> <input type="checkbox"/> <input type="checkbox"/> <input type="checkbox"/>
justesse relative (relative pitch)	: adaptee (adequate) <input type="checkbox"/> <input type="checkbox"/> <input type="checkbox"/> <input type="checkbox"/> inadaptee (inadequate) <input type="checkbox"/> <input type="checkbox"/> <input type="checkbox"/> <input type="checkbox"/> (justesse de timbre –justesse de hauteur)
Sensation d'intensité vocale (LOUDNESS ASSESSMENT) :	
efficacité (efficiency)	: très efficace (very efficient) <input type="checkbox"/> <input type="checkbox"/> <input type="checkbox"/> <input type="checkbox"/> peu efficace (inefficient) <input type="checkbox"/> <input type="checkbox"/> <input type="checkbox"/> <input type="checkbox"/>
puissance (power)	: très puissant (very powerful) <input type="checkbox"/> <input type="checkbox"/> <input type="checkbox"/> <input type="checkbox"/> peu puissant (weak) <input type="checkbox"/> <input type="checkbox"/> <input type="checkbox"/> <input type="checkbox"/>
formant du chanteur (singing formant)	: présent (presence) <input type="checkbox"/> <input type="checkbox"/> <input type="checkbox"/> <input type="checkbox"/> absent (absence) <input type="checkbox"/> <input type="checkbox"/> <input type="checkbox"/> <input type="checkbox"/>
Rapport entre la dynamique et la hauteur (possibilité dynamique) :	
(VOICE RANGE POSSIBILITIES)	
grave (low pitches)	: grande dynamique (great dynamic) <input type="checkbox"/> <input type="checkbox"/> <input type="checkbox"/> <input type="checkbox"/> faible dynamique (weak dynamic) <input type="checkbox"/> <input type="checkbox"/> <input type="checkbox"/> <input type="checkbox"/> N/A <input type="checkbox"/>
aigu (high pitches)	: grande dynamique (great dynamic) <input type="checkbox"/> <input type="checkbox"/> <input type="checkbox"/> <input type="checkbox"/> faible dynamique (weak dynamic) <input type="checkbox"/> <input type="checkbox"/> <input type="checkbox"/> <input type="checkbox"/> N/A <input type="checkbox"/>

Figure 2. Listening sheet from the second angle “Perception of sound”. A suggested English translation of French terms is given bold.

Listening oriented from the second angle: “perception of sound”

The first area of this listening sheet deals with **phonetic aspects** at the segmental and suprasegmental levels. It should be noted that this part of the listening sheet may not be relevant to a listener who is not familiar with the corresponding language. At the segmental level, the stress is put on perception of vocalic contrast (vowels articulated with a similar mouth shape to preserve a timbre homogeneity or contrast), vocalic identification (vowels easily recognisable or not), perception of consonant control (short or long consonants) and consonant pronunciation (unstressed or stressed consonants). At the suprasegmental level, the compliance with phrase and accents is considered. More generally, sentence intelligibility is taken into account.

A second area concerns **sound colour**, mainly timbre aspects: high- or low-pitched, *timbré/détimbré* (with/without ring), balanced/unbalanced in respect to energy spectral distribution, homogeneous/inhomogeneous on the musical sentence. The dark and light characters are also considered.

A third area is related to **sound intensity and pitch**. The pitch can be perceived as absolute (perfect pitch) or, more commonly, as relative to the tuning of the accompaniment. Aspects related to loudness concern power (powerful or weak voice), efficiency (very efficient/inefficient, i.e. whether the perceived vocal effort seems to be well related or not to the perceived loudness), and the perceived presence or absence of a singing formant. The voice range possibilities, in particular the relation

between vocal intensity and pitch, are considered at both the low- and high-pitched part of the singer's tessitura.

Perceptual test of use of the listening sheets

The relevance of the listening sheets and the description consensual properties were tested on two groups of listeners. The first group was composed of 18 listeners (mean age 38 (+/- 11) years old), and included eight professional musicians, eight amateurs, and two non-musicians. 13 of these listeners were familiar with lyrical technique, either by regular practice as a singer, or by frequent listening to this musical style. All but one were familiar with voice quality description, ten listeners describing voice quality occasionally and seven very often. This group of listeners will be referred to as "untrained listeners" in the following sections, since they were not familiar with the listening sheet prior to the test. The second group was composed of six of the authors, who regularly attended the meetings during the three years. They participated in the listening test to explore the degree of consensus among the group. This second group of listeners will be referred to as "trained listeners", since they were already familiar with the listening sheet, and had agreed on the terms and their meaning during the group meetings.

Description of the perceptual test

The perceptual test took place in a meeting room with small groups of 5 to 8 listeners. The sound examples were played using a stereo high-quality sound reproduction system (Haliaetus^{iv}).

For the untrained listeners, the procedure was divided into three parts:

1. Listening and free verbalisation of sound examples.
2. Presentation of the research group work on voice quality, and description of two main listening sheets. The verbal presentation of the first angle (perception of vocal gesture or vocal technique) was complemented by perceptual illustration using prototypic sound examples recorded by the group.^v
3. Replaying of the same sound examples from the Step 1. The subjects were asked to mark the parameters that seemed relevant to them in the listening sheet. When a quality seemed to occur occasionally in the musical phrase, this particularity could be mentioned in the listening sheet by writing an "i" (for intermittent) in the checkbox.

At the end of the test, the subjects filled out a form about their musical skills and knowledge, their feelings about the test and the relevance of using such a listening sheet.

The trained listeners had only to complete part 3 of this listening test.

Six sound examples were selected for the test, corresponding to two professional baritone singers performing a reference example and two variations. The first singer (B1) sang a melody in French composed for the purpose of a previous study (Sotiropoulos, 2004). The reference example (see Figure 3, B1 ref) uses his normal phonation. For this study, the singer performed the following two variations which were chosen to illustrate listening-sheet elements (first angle):

- variation 1 (Figure 3, B1 var1): noisy inhalations, frequent breath intakes and noticeably unbalanced air supply, breathy attack and final transients, voice production in M1iii, staccato melodic articulation and out of rhythm, breathy placement.
- variation 2 (Figure 3, B1 var2): silent inhalations with infrequent breath pauses, without vibrato, with glottal stops and strong glottal final transients, voice production in M1, portamento melodic articulation and in the rhythm, laryngeal and pressed placement.

The time-frequency analyses of the three selected examples are displayed in Figure 3. They visually illustrate acoustical properties related to the vibrato behaviour in the phrase, the attack and final transients, the spectral energy distribution, the breathing pauses, the rhythmic and melodic aspects.

The second singer (B2) sang a melody in Latin – the opening measures of Gounod's *Ave Maria* – recorded during a previous study on voice quality (Henrich, 2001), as illustrated in Figure 4.

Two listening modes were tested. First, the reference example was presented alone. Second, the variations were presented in comparison with the reference example (listening of sound examples in pairs). The example was then repeated as many times as necessary.

Untrained listeners' feelings about the listening sheet

On the form, the following question was asked: *Do you think that, after this test, such a listening sheet could help you in the perception and verbalisation of voice quality? yes, a lot; yes, a little; no, not much; no, not at all*

83% of the subjects considered that this listening sheet could be helpful: 39% chose yes, a lot, and 44% yes, a few. Two subjects (11%) had no opinion, and one subject (6%) thought that the listening sheet would not be of much help.

The next question was about consensus: *Do you think, after this test, that such a listening sheet could provide a more consensual dialogue on voice quality among the different voice specialists? yes, certainly; yes, possibly; no*

All the subjects considered that such a listening sheet could lead to a more consensual dialogue: six subjects (33%) chose yes, possibly and 12 subjects (67%) yes, certainly.

Discussion on the consensus for description of voice quality

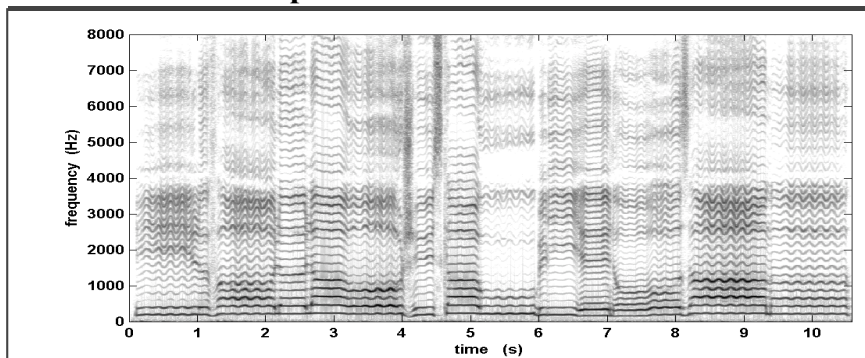
The listeners' answers are illustrated in Figures 5-9. For each category of parameters on the first-angle listening sheet and for the two groups of listeners, these figures present the percentage of listeners who marked the given parameters. We consider that a consensus on description is noticeable among the listeners when a majority of them (more than 50%) mark the same box. In the 5-point bipolar scales, the proportion of each choice, given by the gray scale, indicates the degree of agreement among the listeners.

By analysing the listeners' answers, agreement was found on some parameters, while others showed disagreement.

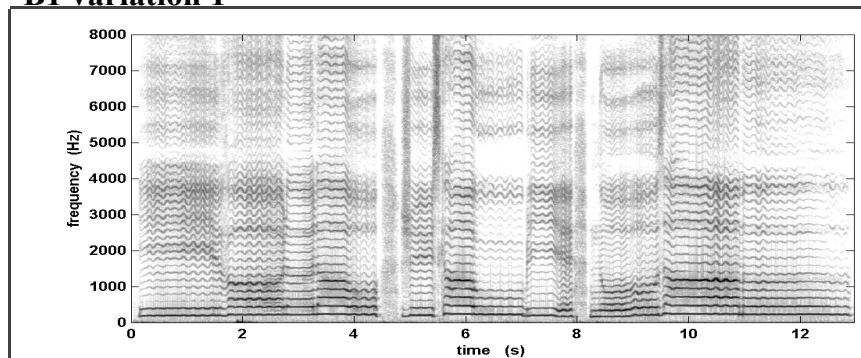
As illustrated in Figure 5, the **dynamics of inhalation and exhalation** is described in a rather consensual manner by all the listeners, whether trained or not. In this area, the listeners did not always agree on the perception of air supply balance (e.g., examples B1 var2 and B2 var2).

In the **vibratory dynamics** area (see Figures 6 and 7), listeners often agreed on the description of attack and final transients' perception, and this agreement was stronger when the type of transient was not varied along the sentence (see Figure 6, examples B1 ref and B2 ref). The answers are less scattered in the case of trained listeners. The perception of laryngeal mechanism was consensual for these examples (see Figure 7). On the contrary, judgment of adequate/inadequate character was not consensual among untrained listeners. Interestingly, the listeners who considered that the laryngeal mechanism was maintained throughout the phrase also sometimes noted variations. It seems therefore that the notion of controlled or uncontrolled laryngeal mechanism variations has not been understood by the untrained listeners. Moreover, in the selected examples, the laryngeal mechanism was not varied, as the singers were always singing in M1. The listeners generally shared the perception of whether the music was in or out of tune, except for examples B1 var2 and B2 var2. The tuning aspect for these two examples was not clearly described by the trained listeners. This is also the case for melodic articulation, which was described in a consensual way, except for example B1 var1. Rhythm is a parameter for which description differs considerably among listeners, be they trained or not. Almost all listeners noted it at each listening (between 78% and 100% for untrained listeners). Yet, in many cases, the rhythmic adequacy was perceived very differently (examples B1 var2, B2 var1 and var2). The absence of a musical accompaniment may explain this disagreement.

B1 reference example



B1 variation 1



B1 variation 2

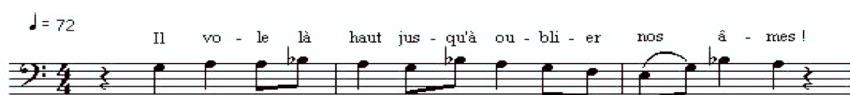
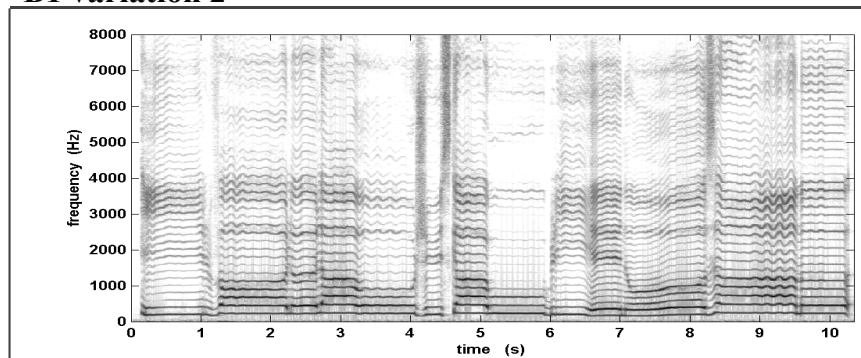
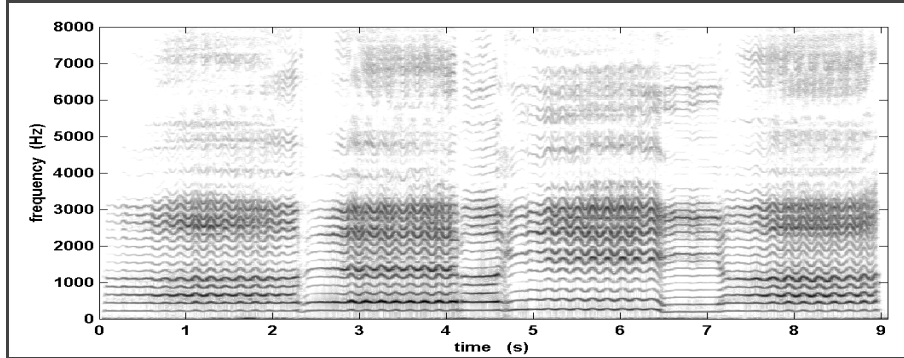
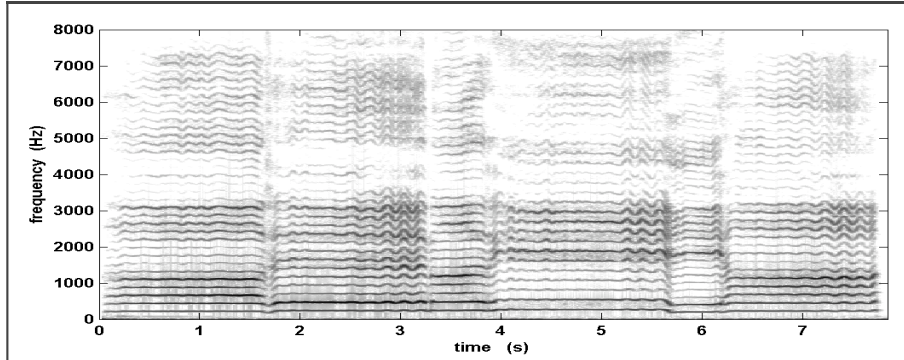


Figure 3. Musical phrase sung by baritone B1 with three different voice qualities. The sung phrase “Il vole là-haut jusqu’à oublier nos âmes” was designed for the acoustic corpus on voice quality (Sotiropoulos 2004). The notation is given in the bottom panel.

B2 reference example



B2 variation 1



B2 variation 2

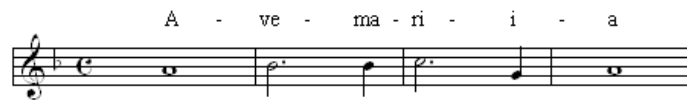
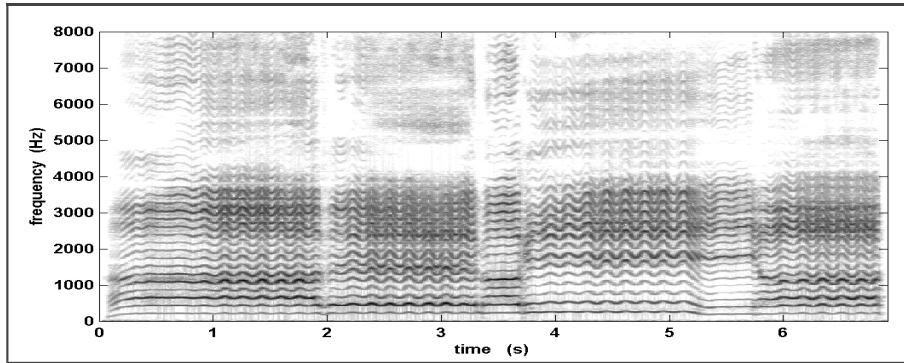
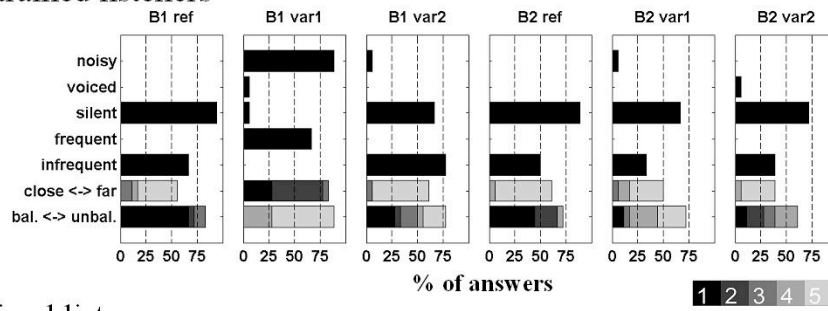


Figure 4. Musical phrase sung by baritone B2 with three different voice qualities. The sung phrase is: “Ave Maria” (opening measures from Gounod’s Ave Maria).

Description of **vibrato** was not consensual among listeners, but improved with training (see Figure 8). Agreement was observed on vibrato adequacy or inadequacy (except for the previously-mentioned example B1 var2, which was either perceived with no vibrato or with an inadequate one). The vibrato frequency was always perceived very differently by the untrained listeners. The ample or restrained characters were also barely consensual between them, whereas they were consensual among the trained listeners. The listeners shared a common perception of the control of vibrato variations.

As with vibrato, the listeners did not agree clearly on ‘**placement**’ perception, and this was only slightly improved by the training (see Figure 9). A good agreement was found in the case of a salient quality (e.g. forward or pressed in examples B1 ref, B1 var1, B1 var2 and B2 var2). However, very often the listeners noted different qualities, sometimes opposite ones. This difficulty to find a consensus on placement description was already observed during the discussions and previous inner verbalisation tests conducted by the research group. One reason may be that the notion of placement has no clear meaning. This observation calls for further research in this area.

A. untrained listeners



B. trained listeners

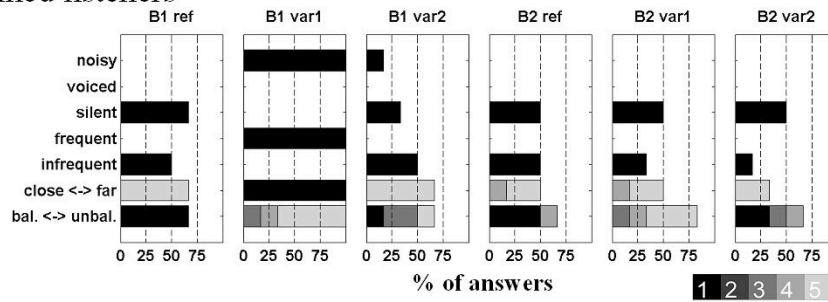
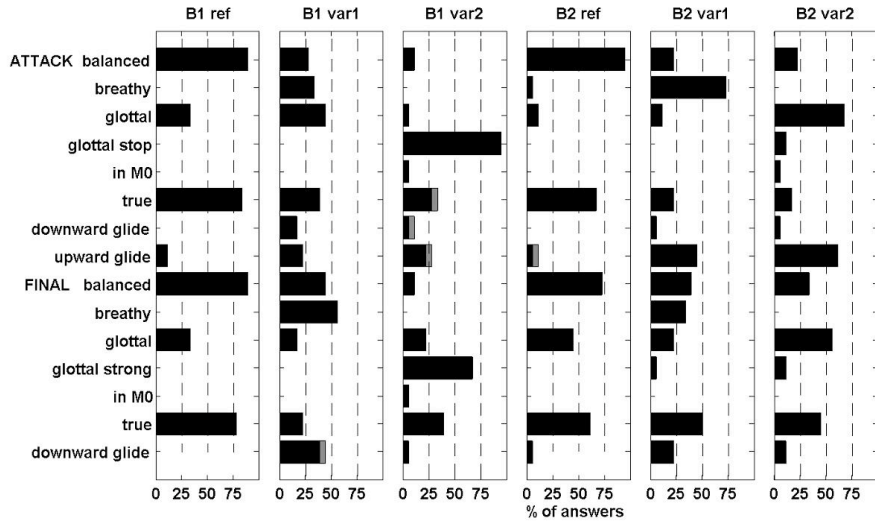


Figure 5. Results for the “dynamics of inhalation and exhalation” area. The upper panels present the answers of the group of 18 untrained listeners, and the lower the answers of the 6 trained listeners. For each listening parameter on the y-axis, the horizontal bars present the percentage of listeners who have indicated it (it was not compulsory to evaluate all criteria). The 5-point bipolar scale answers are presented using a gray scale (from left to right: 1-black to 5-light gray).

A. untrained listeners



B. trained listeners

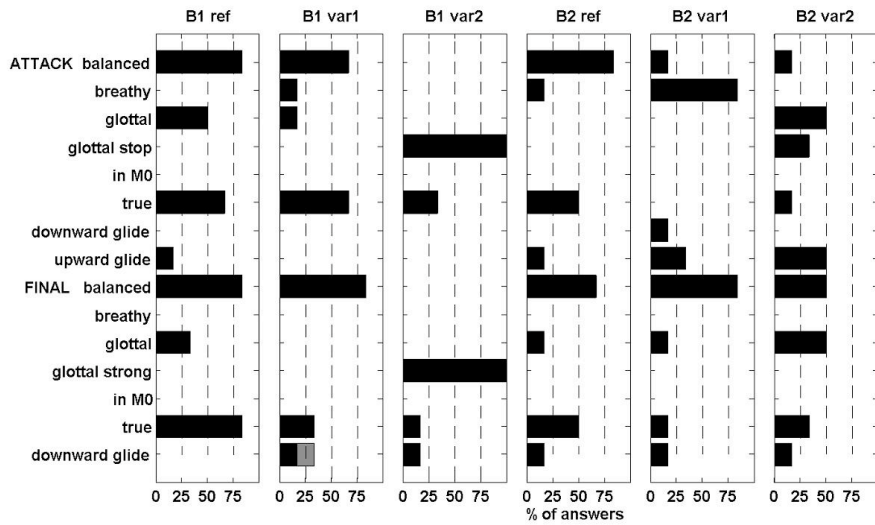
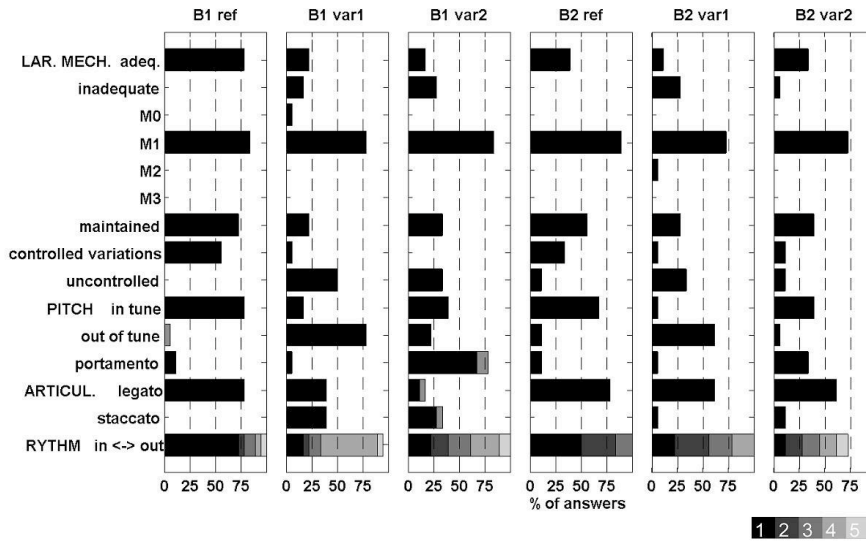


Figure 6. Results for the “vibratory dynamics” area (attack and final transients). The upper panels present the answers of the group of 18 untrained listeners, and the lower the answers of the 6 trained listeners. For each listening parameter on the y-axis, the horizontal bars present the percentage of listeners who have indicated it (it was not compulsory to evaluate all criteria). The gray complements correspond to the cases for which the quality was only occasionally perceived.

A. untrained listeners



B. trained listeners

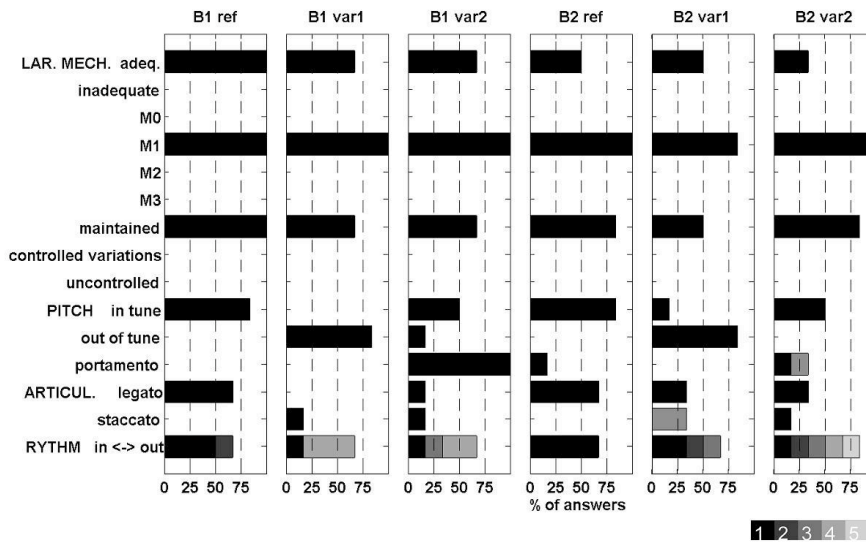
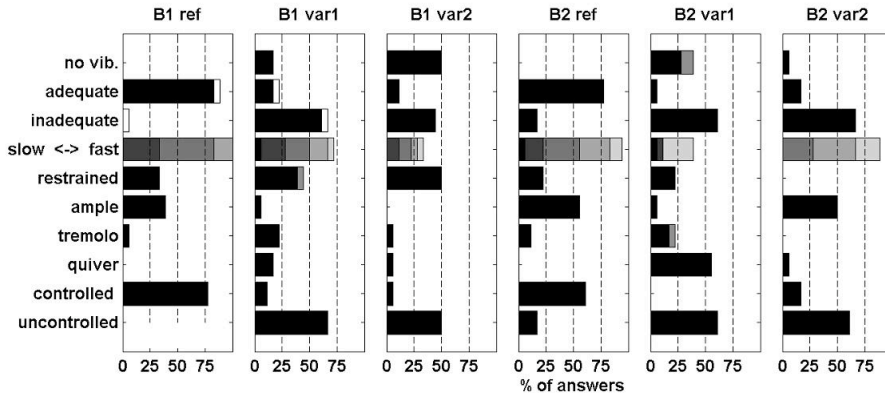


Figure 7. Results for the “vibratory dynamics” area (laryngeal mechanisms, pitch, articulation, rhythm). The upper panels present the answers of the group of 18 untrained listeners, and the lower the answers of the 6 trained listeners. For each listening parameter on the y-axis, the horizontal bars present the percentage of listeners who have indicated it (it was not compulsory to evaluate all criteria). The gray complements correspond to the cases for which the quality was only occasionally perceived. The 5-point bipolar scale answers are presented using a gray scale (from left to right: 1-black to 5-light gray).

A. untrained listeners



B. trained listeners

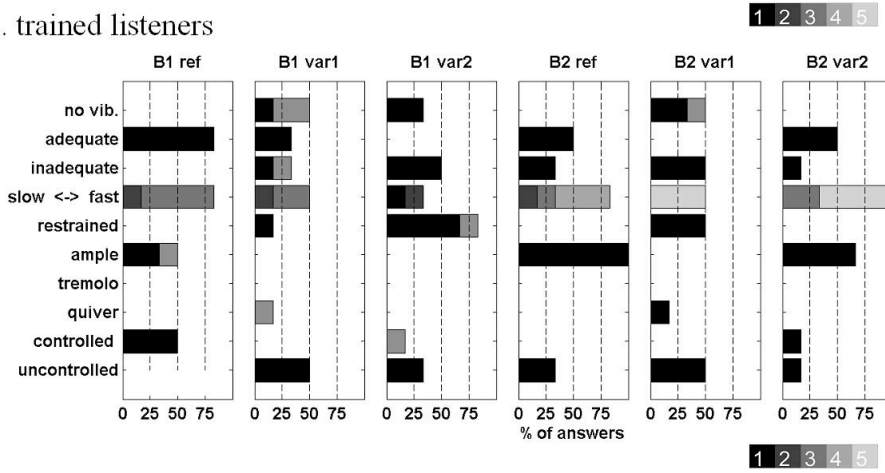
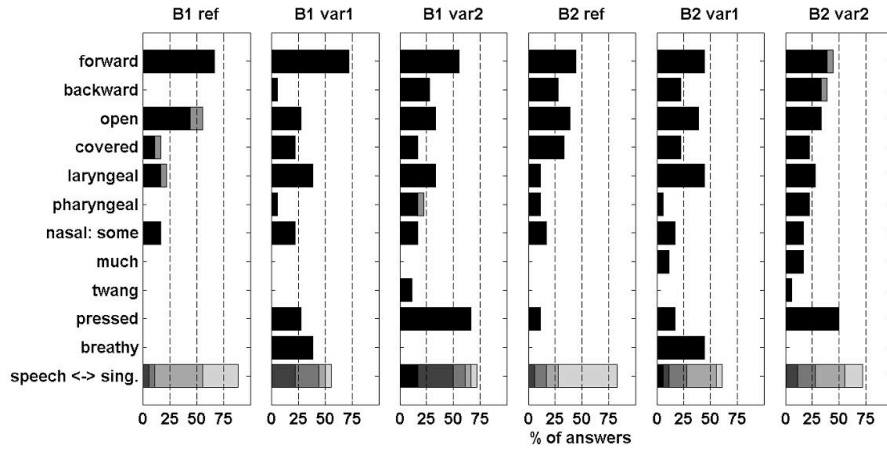


Figure 8. Results for the “vibrato” area. The upper panels present the answers of the group of 18 untrained listeners, and the lower the answers of the 6 trained listeners. For each listening parameter on the y-axis, the horizontal bars present the percentage of listeners who have indicated it (it was not compulsory to evaluate all criteria). The gray complements correspond to the cases for which the quality was only occasionally perceived. The white complements correspond to the cases for which the parameter is judged to be inapplicable (n/a). The 5-point bipolar scale answers are presented using a gray scale (from left to right: 1-black to 5-light gray).

A. untrained listeners



B. trained listeners

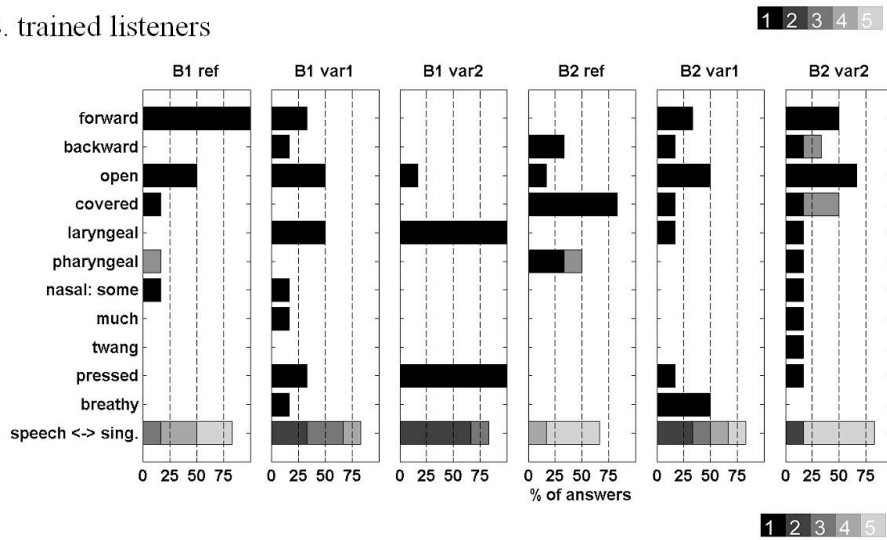


Figure 9. Results for the “placement” area. The upper panels present the answers of the group of 18 untrained listeners, and the lower the answers of the 6 trained listeners. For each listening parameter on the y-axis, the horizontal bars present the percentage of listeners who have indicated it (it was not compulsory to evaluate all criteria). The gray complements correspond to the cases for which the quality was only occasionally perceived. The 5-point bipolar scale answers are presented using a gray scale (from left to right: 1-black to 5-light gray).

Perception of salient characteristics

We wished to determine whether listeners would perceive salient voice qualities selected from the listening sheet. Therefore, singer B1 was asked to perform specific characteristics (see the description of the perceptual test). These characteristics were perceived and verbalised through the listening sheet by a majority of subjects. After listening to the first variation (B1 var1), 89% of the untrained listeners considered that breath intakes were sonorous and noisy. 83% of them said that breath pauses were frequent. The imbalance of breath supply was noticed by 89%. The trained listeners unanimously made the same judgment. The attack transients were perceived more as glottal (44%) than breathy (33%) by untrained listeners, and trained ones considered them balanced. The majority of untrained listeners perceived breathy final transients (55%), whereas the final transients were considered balanced by trained listeners. The use of laryngeal mechanism M1 was well perceived (78% for untrained listeners).

The characteristics of melodic articulation were not consensual among the untrained listeners: 39% perceived a *legato* melody and 39% a *staccato* one. The trained listeners did not even mark these parameters. A majority of listeners perceived that the singer seemed relatively out of rhythm. The breathy placement was not saliently perceived by the listeners, trained or not.

After listening to the second variation (B1 var2), 67% of the untrained listeners perceived silent inhalations and 78% rare breath pauses. These characteristics were perceived less by the trained listeners. Only half of the untrained listeners (50%), and even fewer of the trained listeners, mentioned the lack of vibrato. This could be explained by the fact that the singer did in fact sing with vibrato at the end of his phrase (see Figure 3, bottom panel). This vibrato, even expressed briefly, may have been perceived and so taken into account by listeners. The listeners who did not mark the lack of vibrato all mentioned an inadequate vibrato (44% for untrained listeners, and 50% for trained ones), and/or a restrained vibrato (50% for untrained listeners, and 83% for trained ones). The glottal stops were unanimously perceived. The strong glottal ends were also perceived (67% for untrained listeners, and 100% for trained ones). The use of laryngeal mechanism M1 was perceived (83% for untrained listeners, and 100% for trained ones), together with the portamento (78% for untrained listeners, and 100% for trained ones). The rhythm was not perceived in a consensual manner by untrained listeners. Most of the untrained listeners did not detect a laryngeal placement (33%). They perceived a pressed voice (67%), placed forward (56%). All the trained listeners perceived the laryngeal and pressed placements.

Discussion and prospects

The perception of salient characteristics over the musical phrase was not consensual among untrained listeners, whereas a better agreement was reached among members of the research group. In addition, ratings were globally less scattered within a given perceptual area among members of the research group than among untrained listeners.

This result shows that an oriented-listening is possible and that a consensus is reachable on some criteria among listeners coming from different fields of expertise. It seems possible to improve the interjudge agreement by learning and practice. To this end, a glossary coupled with a library of prototypic sound examples is currently being developed.

The agreement on vibrato and vocal placement achieved during the research group meetings is not reflected by the results of the trained listeners in the perceptual test, nor did the untrained listeners agree much on these criteria. Several explanations for this variability of judgment can be proposed. First, prototypic sound samples were selected in an intuitive way, although their choice resulted from an agreement within the research group. They may not be the best possible representation, despite the singers' attempts to mimic a given voice quality. Secondly, some selected terms or scales may not be relevant to the listener: ambiguity of a term, lack of antonyms on a semantic scale, use of a checkbox to evaluate a continuous quantity or, on the contrary, use of a scale to evaluate a "binary" criterion. The more scattered answers observed for untrained listeners seem to argue in favor of this. In addition, some scales may not have been appropriate to the temporal unit of perception of some criteria. Finally, Garnier et al. (2007) have reported that the perception of vocal placement seems to be very different from the perception of other aspects of voice quality, since it may be perceived at the motor level, through the listeners' own knowledge of voice production and through their experience of the internal vibratory sensations which accompanies it. Thus, it is possible that the listeners did not agree on vocal placement because they may achieve a perceived vocal placement with a different articulatory position. This could explain why the perception of this particular aspect is more individual and variable than others.

There are several ways in which the listening sheet could be improved. It may be more appropriate to place the criterion 'close to speech/close to singing' in the phonetic area, in relation to vowel articulation. The terms 'breathy' and 'pressed' could be moved to the area "vibratory dynamics" or to "dynamics of inhalation and exhalation," as they are more related to breathing management than to sound localisation. Some terms, which seem to be redundant for the listeners, could be either suppressed or better defined. The following pairs of criteria may be considered: '(in)frequent breathing pauses' and 'close/distant pauses', 'laryngeal' and 'pressed', 'close to speech / close to singing' and 'vowels close / contrasted'. Finally, some additional descriptors may default and could be added to the listening sheet, such as the notion of "instability".

The labels and evaluation mode of criteria which have not been consensual in this study should further be explored in order to ensure that the interjudge disagreement did not come from the listening sheet itself. We could look at the binary or gradual characteristics of these criteria, at their potential ambiguity, and at their relevance from a temporal point of view. With regard to this latter point, the method for taking into account the variation of a given parameter could be improved.

This lack of consensus on some characteristics may also come from the fact that the listening sheet presents all criteria on the same level, whereas some of them may be

very salient to the listener and others very “neutral”. In this respect, interjudge agreement may considerably be improved if listeners only evaluate salient characteristics, and if they are able to hierarchize their importance. They could, for instance, number the order in which they fill up the listening sheet.

Lastly, a better consensus may be achieved if the listening sheet could better take into account the listening mode. Indeed, listeners had to evaluate a single voice and then compare two different voices using the same listening sheet. Even when listeners are given a reference voice sample, it is not clear whether their evaluation of another voice sample of the same singer really corresponds to the comparison with the reference, to a prototype in their mind of what should be a lyrical voice, or to their idea of what should be the normal voice of this singer (speech therapists and singing teachers often adopt this listening mode). This problem has been debated a lot within the research group, and no ideal solution has yet been found. In the case of comparative listening, a preposition or suffix could be added to the labels of gradual criteria, such as ‘clearer’, ‘more frequent pauses’, with the label “reference voice” in the middle of rating scales. Another possibility could be to ask the listeners to evaluate both the reference voice sample and the other one on a same listening sheet with “absolute” labels (as has been proposed in this study) using two different colours of pen. We could then evaluate the perceptual difference between these two voices by the distance between their ratings.

Conclusion

A three-year study, conducted by a multidisciplinary research group working on perception and verbalisation of voice quality in Western lyric singing, has established a listening sheet to describe voice-quality perception. Three perceptual angles are considered: perception of vocal gesture or vocal technique, perception of sound, and perception of performance. In a first step, the relevance of the first two aspects has been tested on 18 listeners with different disciplinary backgrounds and unfamiliar with the listening sheet (untrained listeners). Six members of the research group also participated in the perceptual test (trained listeners). The relevance was unanimously approved by the 18 untrained listeners. The analysis of their answers regarding the first aspect shows a good consensus on perception of respiratory and vibratory dynamics. However, no clear inter-listener agreement has been observed concerning vibrato and vocal placement perceptions. Unshared references in memory, different listening modes or the vagueness of the definition of vocal placement could explain this result, which calls for further research. When a voice-quality characteristic is salient and stable over the musical phrase, its perception seems to be slightly improved by the training. One main difference between the untrained group of 18 listeners and the trained group of 6 listeners was that their answers were less scattered within a given area.

In addition to contributing to the search for a consensual terminology, the proposed listening sheet already constitutes an interesting pedagogical tool. On the one hand, it is a training tool for learning to categorise different voice-quality parameters. On the

other hand, it guides the identification and perceptual evaluation of these parameters during voice listening. Finally, it may constitute a very useful discussion aid for experts from the different voice disciplines. Nevertheless, further work is needed to improve the listening sheet and to complete it with the third aspect on perception of performance. Several perspectives have been suggested with regard to the selection and categorisation of descriptive terms, to the evaluation mode, and to the consideration of the listening mode.

Acknowledgments

We are grateful to the listeners who volunteered to participate in the listening test. We acknowledge the Laboratoire d'Acoustique Musicale for regularly hosting the research group meetings.

References

- Castellengo, M. (1986). Les sources acoustiques. In: D. Mercier (Ed.), *Le livre des techniques du son*, Tome 1 (pp. 45-80). Paris: Editions Fréquences.
- Dubois, D. (1991). Les catégories sémantiques naturelles: Prototype et typicalité. In: D. Dubois (Ed.), *Sémantique et cognition: Catégories, prototypes, typicalité* (pp. 15-27). Paris: Éditions du CNRS.
- Dubois, D. (1997). *Catégorisation et cognition: de la perception au discours*. Paris: Collectif, Kimé Ed.
- _____. (2000). Categories as acts of meaning: the case in olfaction and audition. *Cognitive Science Quarterly*, 1: 35-68.
- Eckholm, 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. EHESS.
- 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: 151-169.
- Garnier, M., Henrich, N., Castellengo, M., Sotiropoulos, D., and Dubois, D. (2007). Characterisation of voice quality in Western lyrical singing: from teachers' judgements to acoustic descriptions. *Journal of Interdisciplinary Music Studies*, 1(2): 62-91.
- Guerin, C. (2006). *Glossaire du chant lyrique*. Retrieved July 2007 from <http://chanteur.net/glossair.htm>.
- Guinard, J. X., and Noble, A. C. (1986). Proposition d'une terminologie pour une description analytique de l'arôme des vins. *Sci. Aliment*, 6: 657-662.
- Hammarberg, B. (1997). Perceptual evaluation of dysphonic voices. In: *Proc. Larynx'97*, Marseille, 16-18 June, pp. 11-19.
- 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.
- _____. (2006). Mirroring the voice from Garcia to the present day: Some insights into singing voice registers. *Logoped Phoniatr Vocol*, 31: 3-14.
- Hirano, M. (1981). *Clinical examination of voice*. Wien, New York: Springer-Verlag.

- _____. (1989). Objective evaluation of human voice: Clinical aspects. *Folia Phoniatr*, 41: 89-144.
- Isshiki, N., and Takeuchi, Y. (1970). Factor analysis of hoarseness. *Studia Phonologica*, 5: 37-44.
- Kreiman, J. and Gerratt, B. (2000). Measuring vocal quality. In: R. D. Kent, M. J. Ball (Eds.), *Voice Quality Measurement* (pp. 73-101). San Diego: Singular Publishing Group.
- Laver, J. (1980). *The phonetic description of voice quality*. Cambridge: Cambridge University Press.
- _____. (2000). Phonetic evaluation of voice quality. In: R. D. Kent, M. J. Ball (Eds.), *Voice Quality Measurement* (pp. 37-48). San Diego: Singular Publishing Group.
- Miller, R. (1986). *The Structure of singing: system and art in vocal technique*. New York: Schirmer.
- Philippe, F., Schacher, L., Adolphe, D., and Dacremont, C. (2003). The sensory panel applied to textile goods – a new marketing tool. *Journal of Fashion Marketing and Management* 7(3): 235-248.
- Roubeau, B. (1993). *Mécanismes vibratoires laryngés et contrôle neuro-musculaire de la fréquence fondamentale*. PhD thesis. Université Paris-Orsay.
- Roubeau, B., Henrich, N., and Castellengo, M. (in press). Laryngeal vibratory mechanisms: The notion of vocal register revisited. Accepted for publication in *Journal of Voice*.
- Schaeffer, P. (1966). *Traité des objets musicaux*. Paris: Ed. du Seuil.
- Sotiropoulos, D. (2004). *Analyse acoustique et catégorisation d'un ensemble de qualités vocales pertinente pour la description de voix lyriques masculines*. Mémoire DEA ATIAM. Université Paris 6 Ircam.
- Titze, I. R. (1995). Definitions and nomenclature related to voice quality. In: O. Fujimura and M. Hirano (Eds.), *Vocal Fold Physiology* (pp. 335-342). San Diego, CA: Singular Publishing Group.
- Vennard, W. (1967). *Singing; the mechanism and the technic*. (revised & enlarged edition), New York: Carl Fischer.
- Vurma, A. and Ross, J. (2003). The perception of 'forward' and 'backward placement' of the singing voice. *Logoped Phoniatr Vocol*, 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., and Krüger, H. (1986). Classification of voice qualities. *Journal of Phonetics*, 14: 483-488.

ⁱ Western singing with a classical technique.

ⁱⁱ The term "listening sheet" has been chosen in reference to the "tasting sheets" commonly used in wine tasting, because of the similarity in the approach.

ⁱⁱⁱ The laryngeal mechanisms correspond to different vocal fold configurations. As far as vibratory dynamics is concerned, we chose to use this notion instead of that of vocal register. We refer the reader to Roubeau, 1993; Henrich, 2006; or Roubeau, Henrich, and Castellengo, (in press) for a definition and a detailed description of the vibratory mechanisms. Laryngeal M0 is synonymous of vocal fry or pulse register. Laryngeal mechanism M1 corresponds to the mechanism used for producing modal, chest and male head voice. Falsetto and female head voice are examples of production in laryngeal mechanism M2. Laryngeal mechanism M3 is synonymous of whistle register.

^{iv} <http://www.haliaetus.com/>

^v The corresponding sound examples can be heard online at <http://www.musicstudies.org/>