journal of interdisciplinary music studies
fall 2007, volume 1, issue 2, art. #071204, pp. 62-91

Characterisation of Voice Quality in Western Lyrical Singing:
from Teachers' Judgements to Acoustic Descriptions

Batı Müziğindeki Lirik Şarkı İcrasında Ses Niteliğinin Tanımlanması:
Eğitmen Yargılarından Akustik Betimleyicilere

Maëva Garnier[1], Nathalie Henrich[2], Michèle Castellengo[1], David Sotiropoulos[1], Danièle Dubois[1]

[1]LAM-IJLRA (CNRS UMR 7604 , Université Paris VI, Ministère de la Culture)
[2]GIPSA-Lab (CNRS UMR5216 - INPG, Univ. Stendhal, Univ. Joseph Fourier)

Abstract. This pilot study aims at defining the notion of voice quality in Western lyrical singing and seeks significant and objective criteria to characterise it, from both cognitive and acoustic points of view. We have chosen an approach based on the semantic analysis of experts' discourses, as we assume that the description of the acoustic signal itself cannot fully account for the characterisation of voice quality and that this notion rather relies on the listeners' cognitive representations, which allow them to elaborate a meaningful judgement from acoustic properties.

Therefore we started this study with a listening test, conducted with 11 singing teachers who freely described the voice quality of 18 western lyrical extracts, recorded from three different male singers. The linguistic analysis of these verbalisations brought information about the specific lexical resources (in French) involved in discourses, from which inferences can be made regarding the different conceptions of voice quality and the listening modes of the lyrical expert field. For these listeners, voice quality appeared to be a series of cues allowing to identify sound production, the singer's identity and his emotional state rather than just a set of acoustic characteristics perceived and processed for themselves.

Next, an acoustic database was recorded with three lyrical male singers, who produced different voice qualities illustrating graduations of some verbal descriptors. These terms were selected for their relevance and from the

Özet. Bu pilot çalışma Batı müziğindeki lirik şarkı icrasını tanımlamayı amaçlar ve bunu hem bilişsel hem de akustik açılardan tanımlayacak önemli ve bilimsel ölçütleri araştırır. Akustik sinyalin kendisinin tek başına betimlenmesinin, ses niteliğinin tanımlanması için tam olarak açıklayıcı olamayacağını ve bu nosyonun daha çok dinleyicilerin akustik özelliklerden anlamlı yargılara ulaşmasını sağlayan kendi bilişsel temsillerine dayandığını düşündüğümüz için uzmanların söylemlerinin semantik analizine dayanan bir yaklaşımı benimsedik.

Bu nedenle bu çalışmaya üç farklı erkek şarkıcıdan kaydedilen batı müziğinden 18 lirik müzik parçasını serbestçe betimleyen 11 şan eğitmeni ile gerçekleştirilen bir dinleme testiyle başladık. Bu sözel ifadelerin dilbilimsel analizi, sesin niteliğine dair farklı kavramları ve uzmanlık alanına dair dinleme biçimlerini dikkate alan çıkarsamaların yapılabildiği söylemler içinde, belirli sözlüksel kaynaklar (Fransızca) hakkında malumat vermiştir. Bu dinleyicilere göre ses niteliği, algılanan ve kendi başlarına işleme tabi tutulan akustik özelliklerin bir kümesinden ziyade ses üretimini, şarkıcının kimliğini ve onun duygusal durumunu tanımlamaya yarayan bir dizi ipucu gibidir.

Bir sonraki adımda bazı sözel betimleyicilerin derecelerini temsil eden farklı ses nitelikleriyle icra eden üç erkek şarkıcıyla akustik bir veritabanı kaydedilmiştir. Bu koşullar ilgilerine göre ve daha önceki dilbilimsel ana-

•*Correspondence:* Maeva Garnier, IJLRA, 11 rue de Lourmel, 75015 PARIS.
e-mail: garnier@lam.jussieu.fr

previous linguistic analyses. The acoustic analyses of this database allowed us to put forward different sound descriptors which may be cues for listeners to perceive some aspects of voice quality, and which may account for the semantic overlap of several criteria for evaluating voice quality.

lizlerden seçilmiştir. Bu veritabanının akustik analizi, dinleyicilerin ses niteliğini algılamasıyla ilgili ipuçlarını verebilecek ve ses niteliğini değerlendirmeye dair çeşitli ölçütlerin semantik çakışmalarını açıklayabilecek farklı ses betimleyicilerini öne sürmemizi sağlamıştır.

Keywords: Voice quality, perception, cognitive representation, semantics, timbre, singing, verbal description, acoustic description

Anahtar kelimeler: Ses niteliği, algılama, bilişsel temsil, semantik, tını, sözel betimleme, akustik betimleme

1 Introduction

Let us start with a few questions: by the use of which lexical forms[1] and criteria would one describe the differences between the voices of Maria Callas and Barbara Hendricks, of a pop singer and a lyrical one, of a mezzo-soprano and a counter tenor? Which terms would be useful for a singing teacher to guide a student in improving voice quality? How could one describe the voice quality modifications that a singer can use to perform alternately several roles and several emotions? To answer these questions, we should come back to the concept and characterisation of voice quality in the scientific literature.

Voice quality is often considered from the laryngeal point of view. Several studies combining physiology and acoustics explore how perception and acoustic characteristics of voice are affected by modifications in the vocal-fold vibratory movement, in the glottal flow or contact area pulse shape. Several verbal descriptors of phonation (“hyperfonctionnal”, “pressed”, “breathy”, ‘laryngealised”, “whispered”, “creaky”) or registers (“modal”, “fry”, “chest”, “head”, “falsetto”, “mixte”) are commonly used in the scientific literature and have been related to the variation of parameters such as subglottal pressure, closed quotient, speed quotient, normalised amplitude quotient, peak to peak amplitude, difference between the first two harmonics of the glottal flow, maximum flow declination rate, etc. (Chilers and Lee 1991, Klatt and Klatt 1990, Laver 1980, Henrich et al. 2005, Castellengo et al. 2004). The verbal descriptors “graded”, “rough”, “breathy”, “aesthenic”, and “strained” (GRBAS) are also considered as reference criteria for the perceptual evaluation of the pathological voice, and have also been related to physical descriptors of aperiodicities in the vocal folds vibration (jitter, shimmer, tremor, harmonic to noise ratio, voice breaks, etc.) (Bhuta et al. 2004). Vibrato also appears to be an essential dimension of singing voice quality and the perception of “tremolo”, “wobble” or vibrato adequacy has been related to the pitch modulation rate, extent and delay until onset (Ekholm et al. 1998, Rothman et al. 1990).

[1] We deliberatedly use here the word « lexical forms » and latter « word » or « term » inasmuch (in linguistics) those two latter words make different presuppositions on the representations they refer to (Dubois & Giboreau, 2006). We contrast « term » used in expert languages and « word » in common sense use of (eventually the same) linguistic form.

Yet, the radiated sound at the singer's lips does not relate directly to the sound produced by the laryngeal vibration, but results from its filtering through the vocal tract shaped by the action of articulators. Therefore, voice quality is not only influenced by the glottal-source characteristics, but also by the vocal-tract ones. Many studies have taken an interest in characterising voice quality from the spectral content of the lip-radiated voice sound, in relation to articulatory positions or vocal-tract acoustic properties. In particular, the shift of the first formant has been related to jaw aperture (Lindblom et al. 1971) and to the perception of vocal effort (Liénard and Di Benedetto 1999), and the second formant to the backward/forward movement of the tongue (Lindblom et al. 1971). Vowel nasalisation has been related to the apparition of pole-zeros, also called "anti-formants", in the spectrum (Feng and Castelli 1996). Finally, the presence of singer's formant, which corresponds to a spectral enhancement around 3 kHz, especially in the lyrical voice, has been attributed by Sundberg 1972 to the clustering of the third, fourth and fifth formants. It has been related to the perception of "resonance", "efficiency", and "projection" of a voice (Ekholm et al. 1998, Pillot 2004, Pinczower and Oates 2005). Several parameters, such as the singing power ratio (Omori et al. 1996), the α coefficient (Sundberg and Nordenberg 2006), the low-band and high-band spectral tilts (Castellanos et al. 1996), or the difference in energy between different frequency bands, have been proposed to quantify the spectral balance from a long term average spectrum (LTAS). A few studies have already explored the difference in the first formants frequencies and in spectral balance beteween several singing techniques ("Open throat", "singing forward", "covering' . See Mitchell 2005, Vurma and Ross 2002, Hertegard et al. 1990, Chuberre 2000), singing voice qualities ("resonance/ring", "color/warmth", "clarity/focus", "throaty", "twang", "yawn". See Ekholm et al. 1998, Bjorkner 2006, Steinhauer 1992) and singing styles ("lyrical", "musical theatre", "belting". See Stones et al. 2002, Sundberg et al. 1993)

These first two approaches are based on the coupling between physiological and acoustic parameters. Within this framework, voice quality is defined as a production characteristic of voice, and the produced sound as its consequence. The acoustic characteristics of voice are then considered as cues for the listener, which allow him to identify the production mode and its eventual dysfunctions (voice pedagogy, diagnosis of pathological voices).

Other psychoacoustical studies have taken an interest in relating spectrum characteristics to verbal descriptors of voice quality (De Krom 1995; Husson 1962; Gobl et al. 1992). In this case, the concept of voice quality is no longer considered as a production mode but as a set of acoustical properties made audible and studied in the field of perceptual experience, in a way similar to that used for musical timbre (Von Bismark 1974). However, the listener's perceptual implication (i.e. his subjectivity) is often reduced to the functioning of his auditory system, which allows him to "acquire" and process "information". The reference to this concept of information implicitly postulates that voice quality could accurately be described by the value of acoustic parameters, and that perceptual judgments result from a direct bottom-up processing of the signal "up to" its interpretation and evaluation. This epistemological positioning comes within the framework of traditional psychophysics and cognitivist paradigms, which consist in varying isolated acoustic parameters and correlating them to sound quality judgments drawn by closed questions or by semantic rating scales (Gerratt et al. 1993, Gobl et al. 2003, Bloothooft 1988). The lexical terms used in the rating scales much often derive from the common sense or from the experimenter's own criteria, considered as obvious, without expert linguistic questioning about their meaning and without assessing their relevance to the listener or consensual shared meaning between them.

However, many studies have demonstrated that the perception of some basic characteristics of complex acoustic phenomena, such as pitch or intensity of environmental sounds, depends on higher level categorisation processes (Maffiolo, et al., 1998 ; Guastavino, 2007) . Moreover, even if the listeners could be as highly capable as sensors in separating and evaluating precisely some acoustic voice characteristics – which may be the case in laboratory experiments or for training sessions of voice listening – this kind of listening mode is only a specific case. Ordinarily, the listeners also use other listening modes, which aim not only at describing sound for itself (Castellengo et al. 2005), but rather at teaching a vocal technique (for a singing teacher), at diagnosing a pathology (for a speech therapist), at equalising sound (for a sound engineer), at processing the audio signal (for an acoustician), or at sharing a value judgment (for an opera critic). The notion of voice quality may be polysemic because of its double status: it is altogether a descriptive characteristic of the voice considered as an acoustic phenomenon, and a cue pointing to some knowledge targeted by a listener while listening: the mode of production, the speaker's or singer's identity, the performance adequacy to a musical style, etc.

In this framework, studies have explored to which extent voice quality can be used to recognise a person (Morange 2005, Furui 1986, Lavner et al. 2000), his/her gender (Hanson and Chuang 1999 , Mendoza et al. 1996), his/her emotional states (Gobl and Chasaide 2003 , Pittam et al. 1990), or whether invariant parameters of the acoustic signal could be correlated to personality or inferred attitude of an individual (Ladd et al. 1985, Maffiolo and Chateau 2003). All these studies share the common point of considering the acoustic properties of the audio signal as cues to the speaker's identity, state and intentions. Voice quality is no longer considered as an object existing by itself, independently from any communicative context, and it extends from voice to speech. It is no longer limited to a set of spectral and static characteristics, but it extends to other local temporal aspects (start transients, sound modulation), or to dynamical ones at the sentence level (rhythm, nuances, phrasing, prosody, etc.). Following the change of stimuli, from sustained vowels to complete sentences, the concept of voice quality is no longer similar to musical timbre ("what differs between two sounds with same pitch and intensity": Von Helmholtz 1974), but it becomes a suprasegmental phonetic aspect ("what distinguishes a series of identical speech or sung segments").

In this study, we consider this last broadened conception of voice quality as an evaluative judgment, adequately described as a set of acoustic, perceptual and semantic properties of a voice. To relate these different fields of description to each other, we have chosen to start from the listeners' point of view in order to first explore from a precise linguistic analysis the meaning and consensus of the verbal descriptors and criteria which are used by singing teachers to describe Western classical voices. Only in a second step, we will try to find whether there exist some acoustical parameters accounting for their perception of these relevant and semantically clarified criteria. In the following part (Part 2), we will detail this semantic and cognitive approach (suggested to be named semiophysics by Dubois, 2006), as well as the epistemological framework and the theoretical background on which it is based. In Part 3, the first step of this approach will be presented, consisting in a psycholinguistic study of the singing teachers' verbal descriptions of voice quality. In Part 4, the first results of the second acoustical step will be reported and discussed.

2 Developing a Semantic and Cognitive Approach

In this study, we assume that "information" on voice quality is not in the acoustic signal itself but in the listener's cognitive representations, which allow him to give meaning to the acoustic signal according to his knowledge, his aims and the listening context. We believe that it is necessary to start from the listener's point of view in order to reach acoustic parameters that are relevant as semantic cues to the perception and judgments of voice quality.

Perception and judgement of voice quality is not only considered as a bottom-up process of a pre-existent information but also as a top-down process regulated by previous knowledge in memory, which leads a listener to select some cues rather than others in the acoustic signal, depending on their relevance to his aims and expectations (Dubois 1997). The listener's expertise does not only influence his auditory capacities or the criteria and specific terms which he can use to describe a voice, but it is also a key factor in his selection processes of meaningful properties "in" the acoustic signal.

In addition, the top-down cognitive processes which constrain the listener's judgment depend not only on his expertise, but also on the kind of voice heard. The cognitive theories of categorisation (Harnad 1987; Dubois 2000; Dubois 1991) defend the idea that the past experiences and knowledge of an individual are not "stored" in memory in an exhaustive and unsorted way, but are organised in categories of individual and collective representation, which are structured in accordance with criteria of similitude and dissimilitude. These criteria are diverse, as they are not necessarily limited to a single physical dimension and as they can integrate several sense modalities as well as semantic features from acquired knowledge. The semantic categories are then much more structured by similarities, around typical representations and with different hierarchic levels, than by a logical and regular structure based on physical parameters. They are constantly in construction because of the memorisation and the assimilation of new sensory experiences and knowledge. The categorical nature of perception implies that a meaningful object cannot be qualified before being identified and categorised (Castellengo 1986). Thus, the evaluation of its quality corresponds to the estimate of its differences from a prototype of the category to which it has been linked by the listener. This prototype, as well as the general organisation of the representations, depends not only on individual sensory experiences, but also on conventions and knowledge shared by a cultural group or an expert field. Therefore all these aspects have an influence on voice quality judgements as they induce some expectations or demands which are different according to the category to which the listener assimilates the heard voice.

Contrary to "binary" tasks of sensory detection or perceptive discrimination, the evaluation of an object quality cannot avoid using language. Free verbalisation or semantic scales are relevant to study the notion of quality inasmuch as language participates to the building of cognitive representations and to the elaboration of relevant and consensual descriptors of the perceived object. Actually, the act of naming an object is equivalent to linking it to a meaningful category (Dubois 2000). Therefore language influences our perception of world as it shapes concepts through which we can perceive it. Carrying out a scientific approach of voice-quality perception from verbal data recquires to develop a semantic theory which clarifies the relation between a word, its meaning and its referent (i.e. the "thing" it refers to). In fact, there is rarely an univocal relation between one word and one thing, generally speaking as well as in acoustics. A word often has different meanings in different discursive contexts (Cheminée 2005). As a result, studying voice quality requires apprehending the listeners discourse as an investigation object itself which allows, thanks to a psycholinguistic analysis, the identification of the listeners' cognitive representations, and not conceiving of it only as a flow from which key-words could easily be extracted. Every discourse demonstrates through its coherence an individual appropriation of language, which is a collective object. From a cognitive point of view rather than from a purely linguistic one, every corpus can then be considered as the expression of the subject's cognitive representations about the sensory experience and knowleddge he/she is talking about (Rastier 1991, 1997).

For all these reasons, this study does not start from prior assumption about relevant acoustic parameters, nor examines the influence of their variation on the perception of imposed criteria. On the contrary, it first aims at identifying the criteria which are relevant to the description of a given category of voice (Western-classical singing) for a given category of listeners (French singing teachers) before searching acoustical parameters which may account for the perception of these criteria. Such an approach is not new and has already been adopted in previous studies (Maffiolo and Chateau 2003, Rioux 2003, Wapnick et Ekholm 1997, Henrich et al. 2007, Busson 2002, Morange et al. 2007). In most of these studies, lexical forms are only listed, and some of them are selected for their frequency of occurrence and apparent consensus. On the contrary, few others of these studies go one step further in this approach, in examining in more detail the lexical meaning and consensus of these terms[1], as well as the listening modes adopted by experts listeners. In the same way, and for this same purpose, a detailed linguistic analysis of the lexical forms structure and of their discursive context has been conducted in this study, on the basis of established theories in psycholinguistics (Rastier 1991, 1997, Dubois 1991, 2007, Rosch 1973). It will be presented in the following part.

3 Semantic and Psycholinguistic Exploration of Voice Quality

Material and method

Stimuli: Selections from a database recorded by Henrich 2001 were used to compose the acoustic corpus of this listening test (Garnier 2003). It consisted of three male singers: two bass-baritones and a tenor, who were recorded in a soundproof booth with a high-quality stereo pair of Neuman cardioid microphones placed 50 cm away from the singer's mouth and linked to a DAT (DA30). The singers were asked to sing the first bars of Gounod's *Ave Maria* in a normal way, then to perform the same musical phrase five times while freely modifying voice quality. The selected database for the listening test comprises a total of 18 musical phrases: three "normal" productions and 15 "modified" ones. All these extracts were produced at the same pitch, and were then normalised in intensity.

Subjects: 11 singing teachers of prestigious music schools attended the listening test. Each had more than ten years of experience in vocal pedagogy.

Interviews and discourse analysis: The expert's spontaneous discourses were recorded during the experiment, and literally transcribed thereafter. It represented about 11 hours of interview. Technics of discourse analysis usually used by psycho- and socio-linguists (Cheminee 2005, Morange et al. 2007, Maxim et al. 2007, Dubois 2006, Achard 1987) were applied on these transcriptions at different levels of the psycholinguistic study:

- at a pure lexical level (listing, frequency of occurrence, grammatical structure of the lexical forms, lexical fields, etc.)
- then at a semantic level, by considering the lexical forms in their disursive context (agreement between listeners, identification of different use of a same lexical form, exploration of lexical meaning and semantic relationship between lexical forms, from the analysis of rephrasing, logical marks, juxtapositions, etc.).
- lastly at a cognitive level (subject's implication and confidence in his own judgement examined through the analysis of individual marks and discourse modalisation etc.). The relationship between the cognitive representations of a subject and some linguistic marks in his/her discourse have been established by recent pyscholinguistic theories (Rastier 1991, 1997, Dubois 1991, 2007, Rosch 1973) and constitutes the object of a promising ongoing research.

This article does not deal with a theoretical work in linguistics but especially aims at presenting a productive coupling of psycholinguistics and acoustics in the understanding of voice quality. Therefore we have chosen here not to present all the irksome steps of the analysis process, tables of words and quotations of the corpus, but only their synthesis and main results. More details can be found in Garnier, 2003.

Lexical resources and selection of relevant expressions to build semantic scales

At a first and basic level of the analysis, every expression used to describe voice quality was listed and decomposed in a table into its different grammatical units (noums, adjectives, verbs, quantifiers, prepositions, etc.)

First, we noticed a very rich and variable lexicon used by singing teachers to describe voice quality. For about 11 hours of interview, we counted more than 600 expressions, if we do not restrict this listing to the adjective class, and if we account for the multiple meanings that some verbal descriptors (for ex: "libre" (*free*)) may have when they qualify different objects (for ex: "vibrato" or "tempo").

These different expressions can be sorted by lexical fields. Voice qualifications refer not only to acoustics (39% : attacks, spectral characteristics, intensity, etc.) but also to voice production (16% : articulatory movements, breathing control, involvement of resonance cavities , etc.), to vocal health (6.5% : dynamism, suppleness, tensions, etc.), to music (17% : performance, music style, pitch tuning, etc.) or else to value judgments (21% : aesthetics, technical level, (dis)agreement, etc.). Among the different expressions available to describe sound quality, we notice that the experts often use onomatopoeia ("Ring", "Crr", "crac-crac" , "humpf",…) or sung imitations instead of lexical forms. We also observe that many qualifiers of timbre are metaphorical and refer to other sensory modalities ("rond" (*round*), "râpeux" (*rough*), "acide" (*sour*), "blanc" (*white*), "chaud" (*hot*), etc.) although they seem to be well integrated to the terminology of western lyrical singing.

This abundance of lexical resources to describe voice quality essentially comes from this huge number of metaphorical qualifiers specific to each listener. In fact, more than two third of the listed expressions do not appear more than two times in the whole corpus. On the contrary, about thirty expressions particularly emerge for their high occurrence in the experts discourse. These ones are technical terms as well as metaphors shared by this expert community ("brillant" (*bright*), "clair" (*light*), "détimbré", "antérieur" (*anterior*), "bâillé" (*yawned*), "naturel" (*natural*), "soufflé" (*breathy*), "dans le masque" (*in the mask*), "ouvert" (*open*), "rond" (*round*), "nasal", "renforcement des harmoniques aigües" (*reinforcement of the high harmonics*), "léger" (light/weightless), "couvert" (*covered*), "métallique" (*metallic*), "soutenu" (*sustained*), "équilibré" (*balanced*), "engorgé" (*engorged*), "efficace" (*efficient*), "en appui laryngé" (*with laryngeal strain*), "serré" (*tight*), vibrato "lent" (*low*), "rapide" (*fast*), "régulier" (*regular*)). Despite their diversity, these lexical forms constitute a terminology, as many of them are known and used by all the experts of that field. Besides, singing teachers often add comments like "comme on dit dans le jargon lyrique" (*as we say in lyrical jargon*) showing that they are aware of the specificity of several lexical ressources to their expert field. Thus, we can envisage using these terms to conduct evaluation tests of voice quality, after verifying whether there is an established consensus on their meaning(s) and what is the validity of associating some of them as antonyms of a semantic rating scale.

Meaning(s), consensus and validity of the expressions composing semantic scales

In a second step of the analysis, we examined the inter-teacher variability in the use of these most frequent expressions, as well as the intra-teacher variability while describing different voices. We noticed that there is a good consensus about the meaning of several qualifiers ("nasal", "timbré", "détimbré", "brillant" (*bright*), "sourd" (*dull*), "clair" (*light*), "sombre" (*dark*), "présence de vibrato" (*presence of vibrato*)) in the free discourse of singing teachers, as all the singing teachers spontaneously use these same terms while describing a same voice.

In a third step, we listed the individual marks in the verbal corpus (I, me, your, one, them, the singer, etc. See David 1997), as well as marks of discourse modalisation (maybe, conditional form, a bit, etc. See Cheminee 2005). Despite the previous consensus observed among the singing teachers, we noticed that a lot of their judgements are introduced by the impersonal form "on" (*one*) instead of the first person, as well as much more marks of modalisation than usual for an expert discourse (Vogel 1991, Cheminee 2005). The disengagement or precaution that experts demonstrate towards their discourse might come from the fact that Western-operatic singing is a quite normative field and that judgements about such voices rather calls upon collective than individual representation and experience. It is also possible that these experts are not used to describe voice quality in such a detailed and analytic way, as teaching of vocal technique consists not only in giving verbal advice but also in producing voice examples which can be imitated by the student. A limited standard lexicon could also explain this observations as well as the use of personal images, onomatopoeia or imitations to describe voice quality in more detail. This would also corroborate previous observations about the way musicians describe the sound quality of their own instrument (Faure 2000; Rioux 2000; Busson 2002). A recent study conducted on our verbal corpus has also brought some arguments in that line, since it has shown, through the analysisis of meta-discursive marks, how these singing teachers are building their terminology during the interview, while searching the most appropriate terms to describe their perception and while trying to establish a consensus on their meaning with the experimenter (Maxim et al. 2007).

In a fourth step, we went back to the listed expressions used to describe voice quality and considered their forward and backward discursive context, in the goal to clarify their meaning and to identify the semantic relationship between them (semantic equivalence or overlap, antonymy, causality, etc.). Therefore we paid a particular attention to logical connectors (because, as, so, therefore, etc.), negation markers, comparative forms (more, less, etc.) or similarity connectors ("comme" (*like*), "pareil que" (*similar to*), "du genre" (*kind of*), etc.). We also studied explicit definitions of terms that the experts gave spontaneously or when they were asked by the experimenter for further precision. On the same way, we examined every case of rephrasing (introduced by connectors such as "c'est a dire" (*that is to say*), "je veux dire" (*I mean*), "enfin" (*rather*), etc. See Cheminee 2005). This metalinguistic procedure is of particular interest for the semantic analysis as it allows the utterer to precise or reproduce what he has just said on another way, and brings information to the linguist about semantic equivalences. Expressions come also very often together in a same utterance, only separated by a comma (for ex : "c'est nasal, avec beaucoup d'harmoniques aigues" (*it's nasal, with a lot of high harmonics*)). In that case, it is not possible for the linguist to interpret whether this coma corresponds to a logical link, a rephrasing or a simple description of different coexisting voice properties. Therefore we examined these juxtapositions separately, and distinguished terms which go almost always together from terms which appear in the same utterance only few times. We first reported all these associations of lexical forms in tables, before synthetising them into semantic networks (Garnier 2003). We will now briefly present the main results of this semantic analysis for the most frequent and consensual terms found in the first lexical step of the linguistic analysis.

When observing the discursive contexts where the term "souffle" is used, we notice that it presents two different meanings. In a first case, it refers to the neutral and natural notion of expired air, which makes the vocal folds vibrate. In a second meaning, it is negatively connoted as an equivalent of "air outflow" and corresponds to an incomplete vocal fold closing during the vibration cycle, which induces some noisy characteristics in the produced sound.

The "sourd"(*dull*)/"brillant"(*bright*) semantic pair of descriptors is also complex in cour corpus, as brightness is not only opposed to dullness but also to roundness ("rondeur" in French). In a first accepted meaning, brightness is a positive characteristic of sound and refers to the presence of timbre, especially in the singing formant area, and to its efficiency. In a second meaning, brightness is rather negatively connoted and corresponds to an excess of "clarté" (*light character*), of "metal" and finally very close to the character "strident" (*piercing*).

The interviewed singing teachers also demonstrate three conceptions of the term "clair" (*light*). In a first meaning, the "light" quality is very close to "brillant" (*bright*) an "timbré" ones and refers to the reinforcement of the high harmonics of the spectrum, contrary to "bouché" (*stuffed up*) or "grossi" (*bigger*). A second accepted meaning of the term "clair" (*light*) is semantically close to those of "détimbré" and "blanche" (*white*). In that case it is rather negatively connoted, related to a lack of vocal technique and to a poor spectral content, contrary to its first meaning. Lastly, a third conception connects the qualities "clair" (*light*) and "sombre" (*dark*) to vowel articulation, respectively more anterior or more posterior than in a speech context.

The nasality ("nasalité" in French) is another property of voice quality which presents a lot of different meanings. In the discourse of singing teachers, the adjective "nasal" (*nasal*) and "nasillard" (*twangy*) are very often used indifferently to describe the same voice quality. However, when they are asked for more precision, they all agree that these two descriptors are not synonyms and they can offer a precise description of their distinction: for some of them, the term "nasillard" (*twangy*) corresponds to a more important and exaggerated degree of nasality than the term "nasal" which refers to a moderated or even natural nasality. For others, the character "nasillard" is very negatively connoted and is associated to the reinforcement of high harmonics of the spectrum at the expense of low harmonics, whereas the character "nasal" is a neutral, natural or even positive aspect of voice quality, related to nose relaxation, to throat and rhino-pharynx opening, and very close to the notion of "rondeur" (*roundness*). The polysemy of nasality also comes from the different ways nose is thought to play a role in the production of that quality. For some of the interviewed teachers, nasality is considered as a phonetic concept and refers to the coupling of the nasal and oral cavities thanks to the lowering of the velum, which transforms "oral" vowels into "nasal" ones. For others, nasality corresponds to a vocal technique which consists in deviating a larger fraction of the expired air flow in the nose, especially to produce high pitches in tenor or belting voice (See Cross 2007 for more information about this vocal technic). This conception of nasality rather corresponds to the notion of "nasonnement" (*hyperrhinophony*) in logopedics. Lastly, nasality is for some of the interviewed teachers no more related to the velum but refers to the localization of vibratory feelings "vers l'avant" (*forward*), "dans le nez" (*in the nose*) or "dans le masque" (*in the mask*) and is associated to the reinforcement of the energy in the high frequencies.

The polysemy or the non exact antonymy of some lexical forms may induce some problems if it is disregarded when building semantic scales. Therefore it is necessary to take care about it, but that precaution is not yet sufficient to ensure the relevance of these scales to the listener. In the discourse of singing teachers, we have indeed noticed that not all the description criteria of voice quality correspond to the same temporal unit. The experts especially pay attention to the attacks and the ends of sound, which are both very local indices. They are also able to give a global judgment about other properties such as vocal placement or spectral content on the whole musical phrase. They also pay heed to time-domain variations of voice quality, particularly to variations of intensity, vocal placement and vowel articulation, which bring them cues to musical phrasing or to the singer's technical level. Differences or ratios between different phrase units allow them to perceive vibrato regularity, rhythm fluctuations or pitch accuracy. Thus, it is important, when setting up a perceptual test, that the chosen sounds have a relevant duration in relation to the criteria evaluated by the listener. In addition, some scales evaluating the variation of a criterion in some cases be more relevant than a scale evaluating the criterion itself. In the same way, we have observed that the description of voice quality is sometimes "binary" (absence vs. presence of a property) and sometimes gradual (more or less present, greater or less). Therefore it is essential when building semantic scales to take into account each criterion's specificity.

Cognitive representations and perception modes of voice quality

In a fifth and last step, we listed in the corpus the different nominal forms standing for evaluation criteria of voice quality (cf. Figure 1). First, we notice that experts are able to describe very analytically the acoustic characteristics of a voice: they can detail the quality of several important dimensions of sound (intensity, pitch, vibrato, attacks and ends of sound, spectral content) and note the absence or the presence of several acoustic properties (air overflow, singer's formant). Through the analysis of causality markers in their discourse, and to the backward and forward discursive context of every lexical form, it appears very clearly that the teachers' analytic sound description does not represent their final aim. Rather, it allows them access to another information level about voice quality. In particular, acoustic characteristics give them some cues to musical aspects: intensity variations are interpreted as nuances, note durations and pauses as phrasing. Likewise, some acoustic particularities such as the presence of vibrato or singer's formant are cultural codes shared by the singer and the listener about musical style (lyrical, baroque, music-hall, rock, etc.)

The singing voice presents a particularity in comparison to instrumental sound: it is not only expressive through its musical timbre or through its acoustical variations, but also through the performing of a text. The interviewed singing teachers are indeed sensitive to auditory cues to "diction", to vowel articulation, to [r] rolling, to the native language of the speaker or to the respect or not of the prosodic characteristics of the utterance. Voice is also a particular music instrument inasmuch as it is intimately dependent on his interpreter's physiology and it is only partly visible by the listener. This can explain why acoustic characteristics of voice are very often used by singing teachers in our corpus in order to deduce the production mode and to infer the psychological and physiological characteristics of the singer who is producing this sound. A lot of voice quality descriptions indeed concerns the singer's age, his voice classification (baritone, tenor, etc.), his personality (pretentious, introverted, etc.) or the lyrical role he could play with such a voice quality (an old priest, an apothecary, etc.), the timbre similarity with some famous people or else the emotions or the attitudes which are perceived by the listener in the singer's voice (stressed, bored, depressed, etc.). In the same way, the interviewed singing teachers try to find out, from acoustical characteristics of voice, the suppleness, the tonicity or the fatigue conditions of the larynx, a vocal pathology or tensions which may lead to it. Lastly, they are particularly interested in understanding with which vocal technique or which physiological adjustments sound has been produced, that may come from their teaching and rehabilitation practice. They especially focus on breathing control and qualify the criteria of "soutien" (*breathing support*), "souffle" (*breath*) and "pression". They also pay attention to the vocal "placement". This one is qualified thanks to prepositions ("dans« (*in*), "vers", "du coté de" (*towards*)) and to body parts, in particular the mouth, the lips, the larynx and the resonators forward from the face. Perception of vocal placement is different from the other aspects of voice quality. Although this perception is still made possible by acoustic cues, the listener is not aware of that and seems to perceive directly the placement at the motor level. In the singing teachers' discourse, we notice a disappearance of logical links between sound descriptions and vocal placement descriptions. When we ask them to give a more precise definition of these descriptors, they then describe internal vibratory feelings of the singer (which they experience themselves when they are singing), or they first imitate the voice they heard before describing the feelings they felt or the gestures they made. Thus, it seems that perception of this aspect of voice quality involves, for these expert listeners, their own knowledge about voice production and about the motor control of vocal gesture. This observation is very similar to the motor theory which considers that speech perception is based on articulatory representations and that acoustic characteristics are only cues for the listener to access the articulatory information (Liberman et al. 1985).

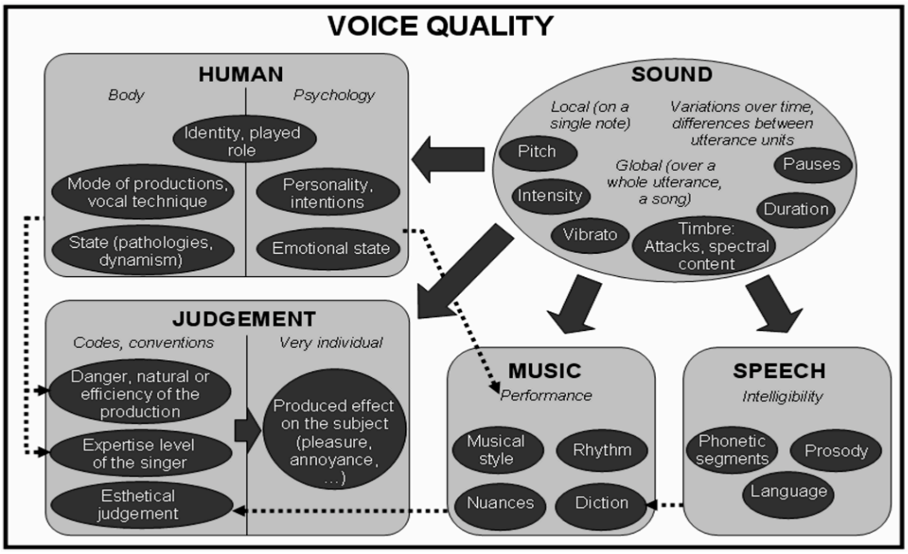

Figure 1. Criteria used by singing teachers to evaluate voice quality of lyrical singing. Voice quality is not restricted to a set of analytic characteristics of sound. These characteristics are used by the listener as cues to musical performance, to speech intelligibility, to the singer's personality or to the production mode of the sound. Then, all these criteria participate to the building of a global judgment of value and agreement.

Finally, singing teachers gather and confront these different cues to speech intelligibility, to musical performance and to the human producing this sound, so as to express value judgments. The value judgments especially deal with the efficiency or the danger of such a production mode, the vocal expertise of the singer (beginner, amateur, professional, etc.) or the adequacy of the performance to the musical style (in or out of tune, value judgments about vibrato, rhythm fluctuations, etc.). Thus, singing teachers characterize voice quality more by its negative aspects than by its positive ones. In other words, they note the absence of expected properties and the unsuitable presence of others. We could think that this hearing mode is specific to singing teachers and comes from their work which consists in correcting their student and making them progress. However, similar observations have been reported in previous study about musical timbre and for different listeners' categories (Faure 2000). Finally, listeners express their agreement ("c'est joli" (*it is beautiful*), "je n'aime pas" (*I do not like*)) from the synthesis of these different criteria. They more generally describe how the heard voice affects them ("agaçante" (*annoying*), "relaxante" (*relaxing*), etc…). This hedonistic judgment is often the first in their discourse. An introspection is then required to make explicit all the different criteria which are responsible for this global feeling.

4 Acoustic Exploration of Voice Quality

Next step in the semio-physic approach is to look for acoustic properties processed as relevant cues for the perception and evaluation of different voice qualities. A preliminar acoustic exploration will now be presented.

The results of the psycholinguistic analyses led up to the selection of the following verbal descriptors relevant for further acoustic analyses: "clair" (*light*), "nasale", "métallique" (*metallic*), "soufflée" (*breathy*), "brillante" (*bright*), "timbrée", "bâillée" (*yawned*), "sourde" (*dull*), "détimbrée", anterior ("antérieure"), posterior ("postérieure"), open ("ouverte"), covered ("couverte"). This selection was motivated by the following questions:

- Would antonymous voice qualities in the singing teachers' discourse be related to oppositions in the acoustic dimensions ?
- Would there be common acoustic attributes to voice qualities which have been found to be semantically very closed in the experts' discourse ?

For this purpose, a new database was recorded. Whereas in the previous section the singing teachers' point of view was explored, we focus on the singers' one in this section. The singers who participated to the experiment were altogether producers and listeners of voice quality. Therefore, we will also examine whether their production strategies reveal some differences in the way they conceive the voice qualities they had to perform.

Material and method

Singers: three singers (different from the first corpus) were recorded: two professional bass-baritones (S1 and S2) and a non professional tenor (S3) (Sotiropoulos 2004). The two singers S1 and S2 had an extensive experience in vocal pedagogy (more than ten

years), and singer S2 had already participated as an expert to the perceptual test described previously.
Protocole: The singers were asked to perform the same musical phrase (cf. Figure 2) several times, first in their usual production, secondly at soft and loud intensity, and finally in producing two gradual levels (“a little” and “very”) of a given voice quality: “clair” (*light*), “nasale”, “métallique” (*metallic*), “soufflée” (*breathy*), “brillante” (*bright*), “timbrée”, “bâillée” (*yawned*), “sourde” (*dull*), “détimbrée”. Singers S1 and S2 also performed the musical sentence in an anterior (“antérieure”) and posterior (“postérieure”) way, and singer B2 in an open (“ouverte”) or covered (“couverte”) way. In addition, the singers were asked to listen to their production after each recording of a given voice quality, and the recording was repeated until the singer was satisfied with the target voice quality he was asked to produce.

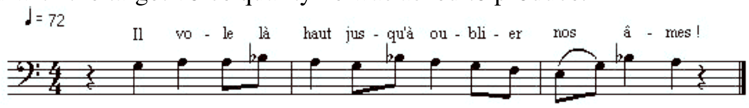

Figure 2. Melody designed for the acoustic corpus on voice quality. (Sotiropoulos 2004).

Material: The audio signal was recorded in a soundproof booth with a 1/2" pressure microphone (Brüel & Kjaer 4165), amplified via a preamplifier (Brüel & Kjaer 2669) and a measurement amplifier (Brüel & Kjaer NEXUS 2690) which delivers an internal calibration signal and thus allows a precise measurement of the sound pressure level. The microphone was placed 30cm away from the singer’s lips. The signal was sampled at 44.1 kHz with a 16 bits data acquisition card (Kay CSL 4400).
Method for acoustic analysis: Each sentence and each vowel were manually indexed using Praat software (Boersma et al. 2004). Fundamental frequency (F0) was estimated by an autocorrelation method. Frequency (in Hz) and amplitude (in tones) of the vibrato were calculated for each vowel using Matlab scripts. The frequencies of the first two formants were estimated for each vowel from the audio signal by using a conventional autocorrelation-based LPC method (Markel and Gray 1976). The number of poles has been chosen as a function of the sampling frequency. We limited our analysis to the first two formants. Higher formants may play an important role in the singing formant (Sunberg 1972) or in personal voice timber (Lavner et al. 2000), but we rather aimed here at examining the link between variations of voice quality and transforms of the vocalic system. The spectral energy in dB was measured in the frequency bands [2-4 kHz] and [4-6 kHz], and was compared to the spectral energy of the [0-2kHz] frequency band. This relative parameter allows to get rid of the variations in vocal intensity and rather focus on the energy repartition over the frequencies.
Statistical analyses were conducted for each singer to compare the variation of acoustical parameters between their “normal” production and every modification of voice quality (ANOVA test with repeated measurements on the 13 vowels of the musical phrase. One factor: voice quality and two levels: “normal” voice and “modified” voice). Statistical results will be reported as follows: NS, for “non significant” ($p>0.05$), *, ** and *** for little to high significance ($p<0.05$, $p<0.01$, $p<0.001$). These statistical tests are based on strong hypotheses: no influence of vowel type (except for formant estimation), independence of the acoustic measurements for

the different vowels, independence of "normal" voice measurement and "modified" voice ones for each singer.

Voice quality and vowel articulation

Singing teachers of our first corpus often make use of the descriptors "en avant" (*forward*), "en arrière" (*backward*), "antérieur" (*anterior*) and "postérieur" (*posterior*) to describe vocal placement. They relate these voice qualities to vibratory feelings and to the impression that sound source is located forward or backward from the face. However, the terms "antérieur" (*anterior*) and "postérieur" (*posterior*) correspond in phonetics to the forward or backward movements of the tongue, acoustically related to a variation of the second formant (Lindblom and Sundberg 1972). The second formant frequency (F2) is raised in the case of anterior vowels (e.g. [i]) and lowered in the case of posterior vowels (e.g. [u]).

Likewise, the "couverture" (*covering*) dimension corresponds for singing teachers to a raise of the velum and to an enlargement of the pharyngeal cavity. This voice quality is semantically close to the terms "postérieur" (*posterior*) and "baillé" (*yawned*) in the experts' discourse.

When the singing teachers we have interviewed describe a sound as "open", it is not clear whether they refer to the phonetic feature of mouth aperture/closure – which corresponds acoustically to a variation of the first formant frequency (F1), higher in the case of open vowels (e.g. [a]) and lower in the case of closed vowels (e.g. [i] and [u]) – or whether the opening does not deal with the jaw but rather with a dilation of the upper pharyngeal cavity. Previous studies have already pointed out the polysemy of the "open throat" notion (Mitchell 2005).

As we already mentioned it, the descriptor "clair" (*light*) seems to have several meanings in the lyrical field. Either it is related to anteriority and opening, and so potentially to vowel articulation, or it is related to a spectral reinforcement in the high frequency part of the voice spectrum. In the same way, the descriptor "sombre" (*dark*) is related to two main singing techniques by singing teachers: « couverture » (covering) on the one hand, and "tubage" (*forming a tube*) on the other hand, which consists in increasing labial protrusion and lengthening the vocal tract.

Finally, we have already mentioned the polysemy of the "nasal" descriptor, sometimes related to forward sound placement, sometimes to lateral opening of the mouth or to nose constriction, sometimes to high-harmonics spectral reinforcement, and sometimes to a coupling between oral and nasal cavities by means of the velum.

Since all these different voice qualities may be partly related to vowel articulation or to vocal-tract adjustments, it seems reasonable to start the search for potential acoustic cues related to their perception at the level of the first two vocalic formants (F1 and F2), then to explore to which extent these different voice qualities share common articulatory attributes. For each singer, the vocalic triangle has been plotted in the traditional F1/F2 phonetic chart, presenting the extrema [a], [i] and [u] of the French vocalic system for different voice qualities.

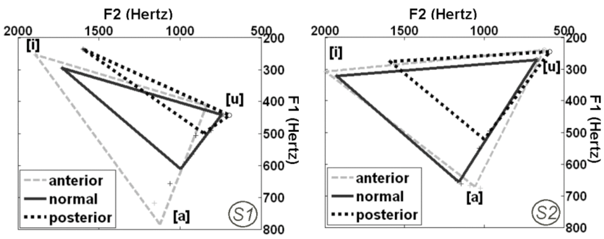

Figure 3. Transform of vocalic system for singers S1 and S2 along the F1/F2 axes between their normal, "antérieure" (*anterior*) and "postérieure" (*posterior*) productions.

Figure 3 presents the transform of the vocalic system for singers S1 and S2 between their normal productions, "antérieure" (*anterior*) and "postérieure" (*posterior*) (as judged by themselves). Vowel [u] does not seem to be articulated in a different manner for these different qualities. On the other hand, the second formant frequency for vowels [a] and [i] and the first formant for vowel [a] decrease for posterior voice quality. For singer S1 in the case of vowel [u], the qualities "antérieure" (*anterior*) and "postérieure" (*posterior*) seem to be contrasted along the dimension of reduction or expansion of the vocalic system toward vowel [u]. This dimension is more complex than the simple "anterior/posterior" phonetic dimension related to the second formant. It also seems different from the "backward/forward" singing voice quality for which Vurma et al. 2002 reported a global lowering (for "backward" singing) and raising (for "forward" singing) of F2 and F3 for the whole vocalic system. On the other hand, the articulation of the anterior voice produced by singer S2 is very close to that of his normal voice. Two explanations could be proposed for this result. It may be that singer S2 has naturally a more forward-placed voice, and, as a consequence, he could not further accentuate this dimension. It may also be that, for this singer, the posterior quality does not contrast with the anterior one, and that his perception may be influenced by other acoustic cues than the first two vocalic formants.

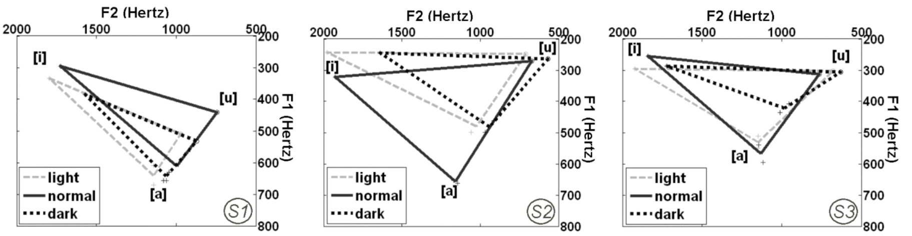

Figure 4. Transform of vocalic system for singers S1, S2 and S3 along the F1/F2 axes between their normal, "claire" (*light*) and "sombre" (*dark*) productions.

Figure 4 presents the transform of the vocalic system for each singer between his normal, "claire" (*light*) and "sombre" (*dark*) productions. As can be seen in this figure, the vowel articulation is not much modified between these different qualities for singer S1, apart from a slight modification from [u] to [o] in both light and dark productions. Concerning S2 and S3, the dark quality corresponds to a reduction of the vocalic system towards vowel [u], similarly to the posterior quality. On the other hand, the light voice is very similar to the normal one for singer S3 on the articulatory point of view. For singers S1 and S2, the light quality does not seem to contrast with the dark one along an articulatory dimension. These results seem to corroborate that the French term "clair" does not correspond to the same voice quality than the English notion of "clarity", related to focus and to the raising of the first formants of the [i] and [a] vowels for Ekholm et al. 1998.

Figure 5 presents the transform of the vocalic system for singer S2 between his normal, "ouverte" (*open*) and "couverte" (*closed*) productions. For this singer, these two qualities do not contrast along an articulatory dimension. The open quality corresponds more to a central reduction of the vocalic system related to an increased similarity between vowel articulations. The covered quality is related to a reduction of the vocalic system towards vowel [u], similarly to the previous observations for "sombre" (*dark*) and "postérieure" (*posterior*) voice qualities. This result is in line with those of Blootooft et al. 1986, Chuberre 2000, Sundberg et al. 1993 and Stone 2003 who observed a lowering of F1 and F2 in the "covered" and operatic voice, considered as more "covered" than speech or other singing styles. For singer S2, the opening quality seems neither clearly related to jaw opening nor to vowel hyperarticulation. The global reduction of the vocalic system observed for his "open" voice evocates the "nasal triangle" of Feng and Castelli 1989 and thus, may come from an increased nasalisation (in the phonetic meaning) of the vowels. This hypothesis would deserve further exploration in analysing the voice spectrum and searching for possible antiformants in the spectrum.

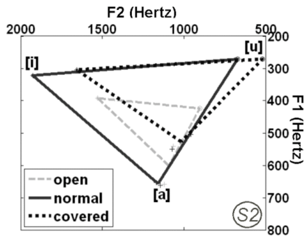

Figure 5. Transform of vocalic system for singer S2 along the F1/F2 axes between his normal, « ouverte » (open) and « couverte » (covered) productions.

Figure 6 presents the transform of vocalic system for each singer between their normal, "nasale" (*nasal*) and "très nasale" (*very nasal*) productions. The property of nasality seems related to an articulatory reduction along the jaw opening/closing dimension (reflected on F1 values) for singer S3. This result is similar with the lowering of F1 observed by Sundberg et al. 2007 for an increase of the velopharynx opening, and is all the more interesting as it only concerns the tenor singer. The coupling with the nasal cavity may indeed be used a lot by tenors or female belters to achieve high pitches (Cross 2007). On the contrary, F1 and F2 increase with nasality for singer S1, which is in line with the observations for anterior quality. Similar observations can be made for vowel [u] and [i] in the case of singer S2, but not for vowel [a] which does not vary much with nasality These results are compable to those of Titze et al. 2001 and Steinhauer et al. 1992, who observed an increase of F1 in "twang". Through the acoustical analysis, we observe again that the term "nasal" seems to present different meaning associated with different vocal techniques, and maybe with different voice classifications (tenors, barytons, etc).

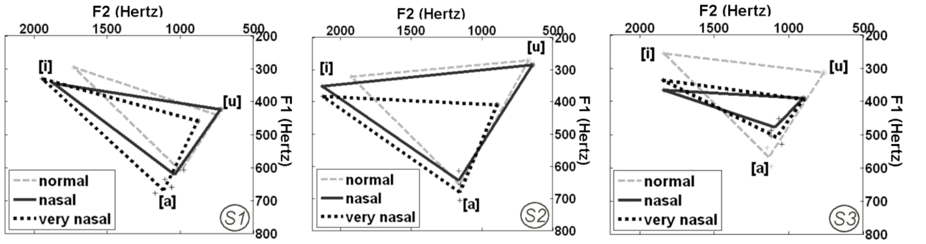

Figure 6. Transform of vocalic system for singers S1, S2 and S3 along the F1/F2 axes between their normal, "nasale" (*nasal*) and "très nasale" (*very nasal*) productions.

Finally, Figure 7 presents the transform of vocalic system for each singer between their normal, "baillée" (*yawned*) and "très baillée" (*very yawned*) productions. Once again, this quality corresponds to a decrease of both first two formants, similarly to the posterior, dark and covered qualities, and to previous observations (Titze et al. 2003).

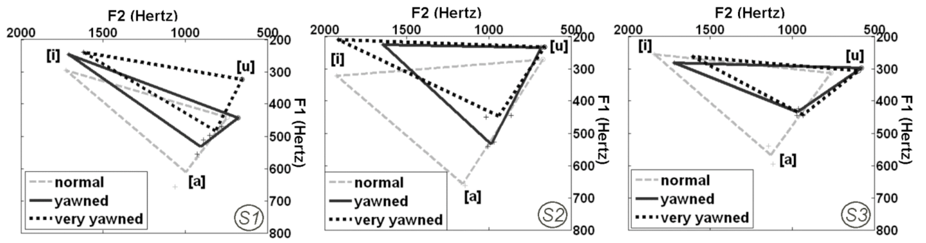

Figure 7. Transform of vocalic system for singers S1, S2 and S3 along the F1/F2 axes between their normal, "baillée" (*yawned*) and "très baillée" (*very yawned*) productions.

Voice quality and spectral balance

The notions of "richesse spectrale" (*spectral richness*), of "équilibre" (*balance*), of "balance spectrale" (*spectral balance*) between low and high harmonics, of "efficacité" (*efficiency*), of "portée" (*reach*) or of "formant du chanteur" (*singer's formant*) take an important place in the verbal description of voice quality by singing teachers. Therefore, we paid attention to the spectral difference between the frequency bands [2-4kHz] or [4-6kHz] and the frequency band of the voice first harmonics [0-2kHz]. If we come back to the audio examples used for the first perceptual test, a great difference can be noticed on the singer's formant emergence in comparison to the energy in the first-two-formants frequency band between the voice most frequently judged as dull by the experts and the one most frequently judged as bright (see Figure 8).

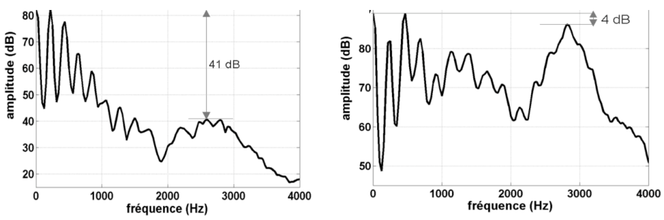

Figure 8. Comparison of the mean spectra calculated over the whole musical phrase "Ave Maria" in the case of the voices judged as the dullest and the brightest in the first corpus.

For S1 and S3, the qualities "baillée" (*yawned*), "sourde" (*dull*) and "détimbrée" show a noticeable decrease of spectral energy between [2-4kHz] and [4-6kHz] as compared to the frequency band [0-2kHz] (cf. Figure 9). Similar observations can be made for singer S2 with regard to the qualities "baillée" (*yawned*), "sourde" (*dull*), "claire" (*light*) and "sombre" (*dark*). Singers S1 and S3 have already much spectral energy in the range [2-4kHz], and may not be able to add more of that to produce a voice "timbrée", "brillante" (*bright*), "métallique" (*metallic*), "sombre" (*dark*) or "nasale" (*nasal*). These results are not in line with those of Ekholm 1998, Sundberg 2007, Vurma et al 2002, and Titze 2001, who observed a significant increase of energy in the singing formant region for "ringing/resonant", "forward", "twang" voices or for an increase of the velopharyngeal opening. A similar trend is observed in the acoustic energy between [2-4kHz] and [4-6kHz]. So the comparison between these two frequency bands does not bring any further information on the selected voice qualities. The qualities "brillante" (*bright*) and "sourde" (*dull*), as well as "timbrée" and "détimbrée", are contrasted along the spectral richness dimension for singers S1 and S3, but not for singer S2.

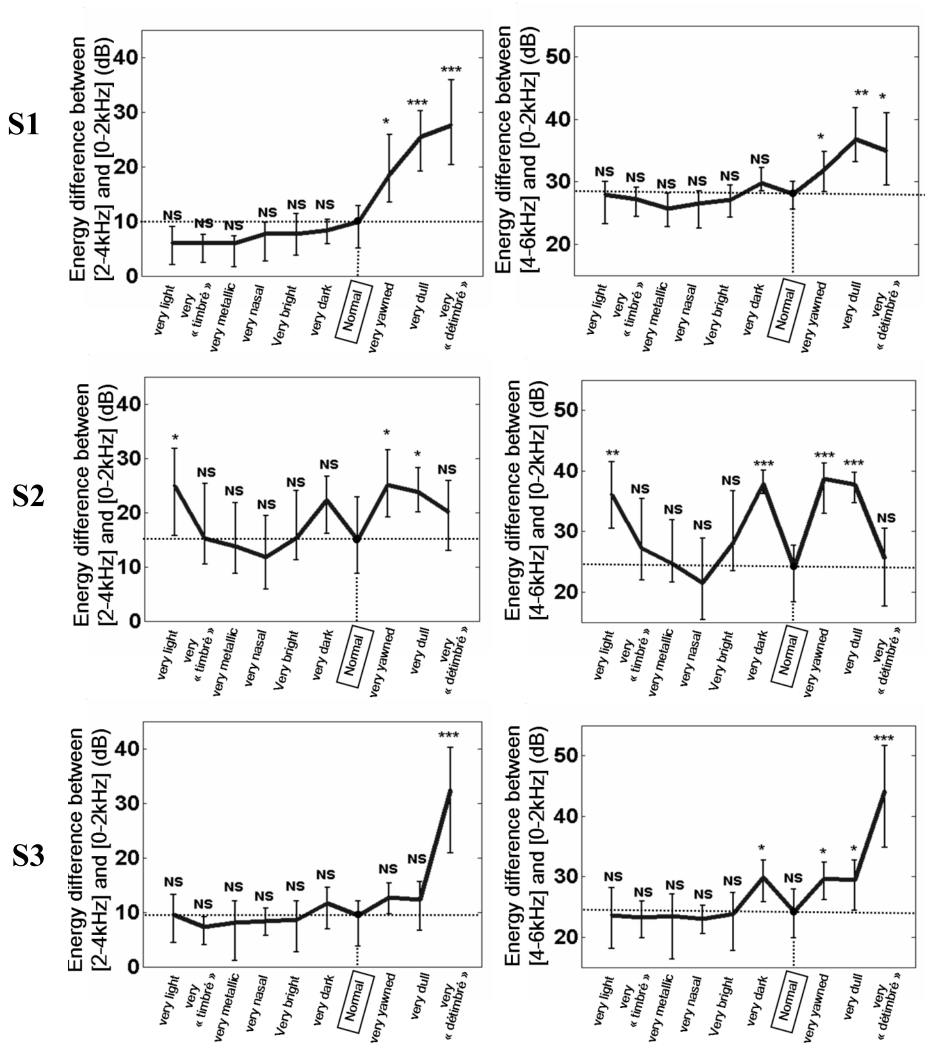

Figure 9. Spectral energy difference between the frequency bands [2-4kHz] or [4-6kHz] and [0-2kHz] for several voice qualities produced by singers S1, S2 and S3. For each quality, the two parameters have been measured on 13 vowels extracted from the musical sentence. Error bars present the values distribution around the mean value in each case. The stars or "NS" label above each error bar indicate the level of statistical significance of the difference between each quality and the normal one. The statistical analyses are detailed in Table 1.

Table 1. Statistical results of the Anova tests which compare the spectral energy difference between the frequency bands [2-4kHz] or [4-6kHz] and [0-2kHz], between the « normal » voice of each singer (S1 to S3) and several modifications of their voice quality. The label « + » represents a slight modification as compared to the normal voice and « +++ », an important modification.

		En[2-4kHz] - En[0-2kHz]			En[4-6kHz] - En[0-2kHz]		
		S1	**S2**	**S3**	**S1**	**S2**	**S3**
light	+++	3.4 dB, F(1) = 8.1, **	10.0 dB, F(1) = 5.3, *	0.2 dB, F(1) = 0.01, NS (p>0.9)	0.1 dB, F(1) = 0.01, NS (p>0.9)	12.1 dB, F(1) = 17.1, **	0.6 dB, F(1) = 0.1, NS (p>0.8)
	+	0.2 dB, F(1) = 0.02, NS (p>0.8)	8.6 dB, F(1) = 4.9, NS (p>0.05)	0.9 dB, F(1) = 0.1, NS (p>0.7)	1.2 dB, F(1) = 1.1, NS (p>0.3)	10.0 dB, F(1) = 13.4, **	0.2 dB, F(1) = 0.01, NS (p>0.9)
"timbré"	+++	3.0 dB, F(1) = 7.1, *	0.2 dB, F(1) = 0.002, NS (p>0.9)	2.0 dB, F(1) = 1.0, NS (p>0.3)	0.9 dB, F(1) = 1.0, NS (p>0.3)	3.2 dB, F(1) = 0.9, NS (p>0.3)	1.0 dB, F(1) = 0.2, NS (p>0.6)
	+	1.8 dB, F(1) = 2.1, NS (p>0.1)	0.8 dB, F(1) = 0.03, NS (p>0.8)	0.5 dB, F(1) = 0.04, NS (p>0.8)	0.2, F(1) = 0.04, NS (p>0.8)	1.6 dB, F(1) = 0.3, NS (p>0.5)	0.5 dB, F(1) = 0.04, NS (p>0.8)
metallic	+++	2.3 dB, F(1) = 2.9, NS (p>0.1)	1.1 dB, F(1) = 0.1, NS (p>0.7)	1.3 dB, F(1) = 0.2, NS (p>0.6)	2.0 dB, F(1) = 3.6, NS (p>0.06)	0.7 dB, F(1) = 0.1, NS (p>0.8)	0.8 dB, F(1) = 0.1, NS (p>0.7)
	+	3.5 dB, F(1) = 12.0, **	1.6 dB, F(1) = 0.2, NS (p>0.6)	1.0 dB, F(1) = 0.1, NS (p>0.7)	0.4, F(1) = 0.4, NS (p>0.5)	0.2, F(1) = 0.004, (p>NS 0.9)	0.6 dB, F(1) = 0.05, NS (p>0.8)
nasal	+++	2.3 dB, F(1) = 3.0, NS (p>0.09)	3.2 dB, F(1) = 0.6, NS (p>0.4)	1.1 dB, F(1) = 0.3, NS (p>0.5)	2.1 dB, F(1) = 4.4, *	2.4 dB, F(1) = 0.5, NS (p>0.6)	1.2 dB, F(1) = 0.4, NS (p>0.5)
	+	4.1 dB, F(1) = 7.5, *	3.6 dB, F(1) = 0.9, NS (p>0.3)	1.2 dB, F(1) = 0.4, NS (p>0.5)	0.7 dB, F(1) = 0.2, NS (p>0.6)	4.1 dB, F(1) = 2.4, NS (p>0.1)	1.6 dB, F(1) = 0.8, NS (p>0.3)
bright	+++	1.6 dB, F(1) = 1.5, NS (p>0.2)	0.4 dB, F(1) = 0.01, NS (p>0.9)	0.8 dB, F(1) = 0.18, NS (p>0.7)	1.0 dB, F(1) = 1.5, NS (p>0.2)	4.3 dB, F(1) = 1.7, NS (p>0.2)	0.4 dB, F(1) = 0.03, NS (p>0.8)
	+	0.9 dB, F(1) = 0.3, NS (p>0.5)	0.1 dB, F(1) = 0.0002, NS (p>0.9)	1.2 dB, F(1) = 0.2, NS (p>0.6)	0.3 dB, F(1) = 0.1, NS (p>0.7)	0.9 dB, F(1) = 0.1, NS (p>0.7)	0.1 dB, F(1) = 0.001, NS (p>0.9)
dark	+++	1.9 dB, F(1) = 2.7, NS (p>0.1)	7.4 dB, F(1) = 4.2, NS (p> 0.06)	2.3 dB, F(1) = 1.0, NS (p>0.3)	0.5 dB, F(1) = 0.3, NS (p>0.5)	14.0 dB, F(1) = 46.0, ***	5.7 dB, F(1) = 6.8, *
	+	0.3 dB, F(1) = 0.1, NS (p>0.7)	4.5 dB, F(1) = 1.5, NS (p>0.2)	1.2 dB, F(1) = 0.2, NS (p>0.6)	1.3 dB, F(1) = 2.1, NS (p>0.1)	8.2 dB, F(1) = 12.6, **	3.7 dB, F(1) = 2.3, NS (p>0.1)
yawned	+++	5.8 dB, F(1) = 10.3, **	10.1 dB, F(1) = 7.0, *	3.3 dB, F(1) = 2.7, NS (p>0.1)	1.3 dB, F(1) = 1.1, NS (p>0.2	14.8 dB, F(1) = 33.5, ***	5.4 dB, F(1) = 6.7, *
	+	4.9 dB, F(1) = 12.5, **	4.5 dB, F(1) = 12, NS (p>0.3)	2.2 dB, F(1) = 0.8, NS (p>0.3)	2.1, F(1) = 3.2, NS (p>0.08)	8.0 dB, F(1) = 9.9, *	5.1 dB, F(1) = 4.6, NS (p>0.05)

Table 1. Statistical results of the Anova tests-Continued

		En[2-4kHz] - En[0-2kHz]			En[4-6kHz] - En[0-2kHz]		
		S1	S2	S3	S1	S2	S3
dull	+++	14.0 dB, F(1) = 64.5, ***	8.8 dB, F(1) = 7.0, *	3.0 dB, F(1) = 1.4, NS (p>0.2)	7.3 dB, F(1) = 32.6, ***	13.9 dB, F(1) = 41.1, ***	5.3 dB, F(1) = 5.0, *
	+	12.4 dB, F(1) = 46.9, ***	9.5 dB, F(1) = 7.0, *	4.9 dB, F(1) = 3.8, NS (p>0.07)	4.5 dB, F(1) = 18.4, ***	13.9 dB, F(1) = 27.7, ***	7.1 dB, F(1) = 8.2, *
"détimbré"	+++	15.8 dB, F(1) = 50.9, ***	5.1 dB, F(1) = 1.7, NS (p>0.2)	23.0 dB, F(1) = 28.8, ***	5.6 dB, F(1) = 15.7, ***	1.5 dB, F(1) = 0.2, NS (p>0.6)	20.0 dB, F(1) = 27.4, ***
	+	10.2 dB, F(1) = 50.3, ***	7.2 dB, F(1) = 5.3, *	4.4 dB, F(1) = 2.4, NS (p>0.1)	5.5 dB, F(1) = 11.7, **	4.8 dB, F(1) = 4.3, NS (p>0.06)	6.0 dB, F(1) = 4.0, NS (p>0.07)

Moreover, singers S1 and S3 seem to conceive the "clair" (*light*) quality more as an excess of brightness related to a great reinforcement of high harmonics, while singer S2 seems to relate it to an untrained voice, "détimbrée". In the same way, singers S2 and S3 seem to consider a dark voice as an equivalent to a dull one (similarly to Bloothooft 1986), while the dark voice produced by S1 is very similar to his normal voice from a spectral point of view, and with a harmonic richness closer to the qualities "brillante" (*bright*) and "métallique" (*metallic*) than to the "sourde" (*dull*) and "detimbrée" one. These observations confirm the different meanings of the terms "clair" (*light*) and "sombre" (*dark*) evidenced during the psycholinguistic analysis of the singing teachers' discourse. These two qualities are contrasted along a spectral dimension for singer S3, but not for singer S1 and S2.

These varations of spectral balance could simply be related to variations of vocal intensity. We indeed observe that the three singers of our corpus accompany their modifications of voice quality by variations of vocal intensity that can reach 8 dB. It has already been shown that an increase of vocal effort induces an enhancement of the spectrum for the medium and high frequencies (Sundberg and Nordenberg 2006). Therefore the LTAS measure α of spectral balance, proposed by Frokjaer-Jensen and Prytz (1976), and defined as the ratio between the sound energy above and below 1kHz, has been computed on the different vowels of the musical phrase and for every modification of voice quality. We have first noticed that for singer S1 and S3, this α coefficient increases as a linear function of vocal intensity for the soft, moderate and loud productions of their "normal" voice quality. The linear approximation of the data gave $\alpha=0.69*Idb-60.3$ ($R=0.998$) for singer S1 and $\alpha=0.75*Idb-63.9$ ($R=0.946$) for singer S3. For singer S2, the α coefficient increases as a quadratic function of vocal intensity ($\alpha=-0.0277*Idb^2+4.39*Idb-178$). These observations are in line with those of Sundberg and Nordenberg 2006, even if these authors have rather examined the relationship between the coefficient α and the equivalent sound level (Leq).

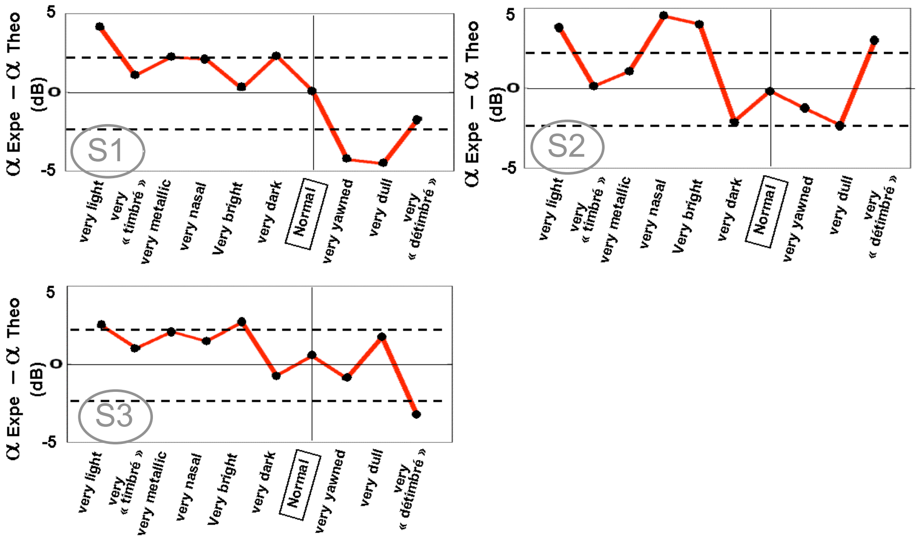

Figure 10. Difference between the experimental α coefficient of spectral balance, computed from the voice spectrum, and the "theoretical" α coefficient, computed from the measured vocal intensity and from the linear or quadratic relationship between α and vocal intensity established for each singer, from their "normal quality" productions at soft, moderate and loud intensity.

Based on these linear and quadratic approximations, as well as on the vocal intensity measurements of every modification of voice quality, we have then computed the "theoretical" value of the α coefficient that these different productions should give if the spectral balance would only be determined by vocal intensity. Figure 10 represents the difference between the measured and the "theoretical" values of the α coefficient for different modifications of voice quality. We observe that a few productions of each singer show an "error" of α estimation greater than 2.4 dB, which is the maximum error of prediction reported by Sundberg and Nordenberg 2006 for male subjects. For the corresponding singing voice qualities, the linear relationship between the α coefficient and vocal intensity seems not to be valid anymore. In these cases, and contrary to the observations of Bloothooft and Plomp 1986, the observed spectral balance may not only result from a variation in vocal intensity, but also from other laryngeal adjustments or resonantial strategies.

Correlations and sharing of acoustic properties between the different voice-quality criteria

A strong correlation between high-harmonic "richesse spectrale" (spectral richness) and "vibrato" has emerged from the analysis of the singing teachers' discourse, similarly to Ekholm 1998. Mitchell 2005 also showed that the vibrato extend was significantly reduced when singers do not use "open throat" tehnique. We also observed a strong correlation between the "manque de timbre" (lack of timbre) and the "air sur la voix" (breathy voice).

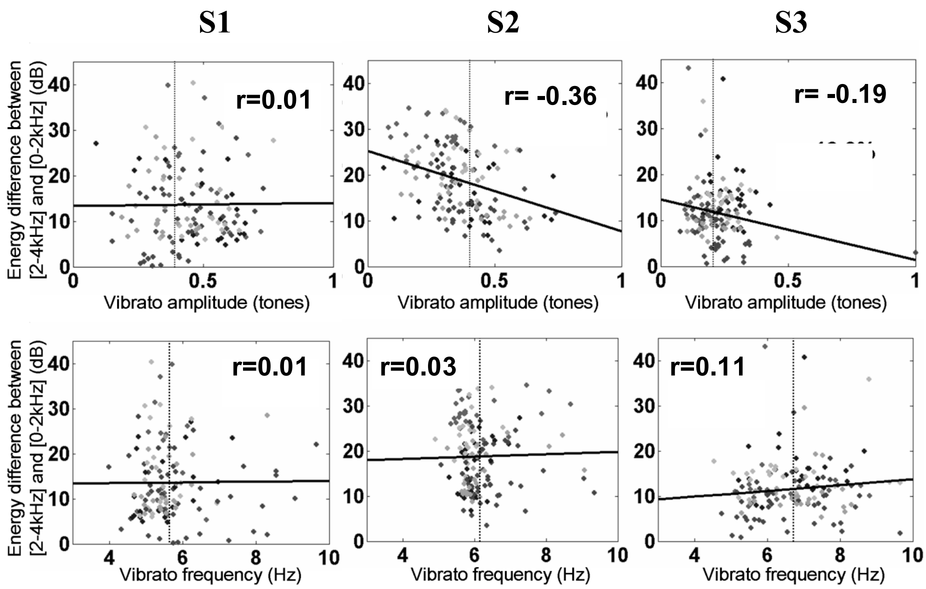

Figure 11. Spectral energy difference between the two frequency bands [2-4kHz] and [0-2kHz] as a function of vibrato frequency (bottom panels) and amplitude (top panels) for each singer and all the produced voice qualities. The dotted vertical line presents the mean vibrato amplitude and frequency.

From an acoustic point of view, no correlation is observed between the voice spectral richness in the frequency band [2-4kHz] and the vibrato amplitude or frequency for each singer and over all the different voice qualities (see Figure 11). Moreover, the vibrato characteristics are not noticeably modified between the normal production and the modified ones for each singer. For singers S1 and S2, the vibrato frequency shows a weak variability, comparable to the measurements of Rothman et al. 1990 and Prame 1994. It seems that vibrato is a natural aspect of voice, whose frequency is determined by physiological constraints, and few controlled by these singers, who seem rather to use the vibrato amplitude dimension to change their voice quality. On the contrary, the less experimented singer S3 exhibits a more variable vibrato frequency, but a relatively constant vibrato amplitude.

5 Conclusion

We have presented here the results of an ongoing research which aims at better understanding the notion of voice quality and the way it can be described.

We have not applied here a commonly-used methodology that would consist in examining how the variation of several acoustic parameters - which we would assume to be relevant - would influence the perception of different voice qualities - whose "label" would have been selected from the common sense. On the contrary, we have tried in this study to avoid basing our exploration of voice quality on assumptions or common sense, but to treat this notion in a more scientific way. This has first required a necessary come back to the expert listeners' point of view, so as to determine what are the relevant criteria and lexical forms for singing teachers to perceive and describe the voice quality of the specific category of lyrical voices. In that goal, we have developed a semantic and cognitive approach, which has allowed us to confirm the results of previous study about sound quality perception (identification of sound source preliminary to other verbal description, use of onomatopoeia and imitation, value judgments mainly negative), to validate scientifically the relevance of some criteria and verbal descriptors for studying voice quality, and to identify new semiotic properties related to the experts' specific cognitive processing which were not taken into account by the physical description alone (expert's listening aim for which analytical listening of sound provides only cues, articulatory representation preliminary to acoustic one for some given voice qualities, in particular those related to vocal placement, technical and esthetical expectations, field-specific terminology). In other words, whereas some physical parameters enable to describe voice quality adequately, others are only cues pointing to listeners' representations in memory of what voice quality is for them. These results imply that voice quality, considered as a cognitive object, has to be scientifically explored and described along the cooperation of different research fields and not exclusively restricted to physics. In particular, we wish to convince the reader that the analysis of cognitive representations given in discourse as shared representations may be a relevant and productive issue. Indeed, if one considers that the lexical forms are not labels which would take their meaning from their reference to one thing, the analysis of their meaning can be interpreted as a negotiated consensus in reference to cognitive representations which have been built on acoustic stimulations. Such representations diversely establish and configure what is referred to, ranging from an individual effect of the world to a collective shared representation of an object of the world. The verbal criteria should be carefully used, not because of their linguistic or "subjective" status, but because of the precision of their semantic contours and the variations in consensus on their meaning(s). This study has shed further light on this issue by many methodological considerations on the set-up of experimental protocols, notably the precautions to be taken when deciding semantic scales for a perceptual test.

The results presented in this paper are specific to this category of expert listeners. Would other listeners use the same lexicon to describe voice quality? Would the knowledge on vocal production matter in the voice-quality representation of non-expert listeners? To answer these questions, it would be helpful to interview other kinds of experts as well as non-expert listeners, so as to assess whether the obtained results on the notion of voice quality and the productivity of the analysis method could or not be extended. It would also be useful to extend this experiment to other types of voice, especially female voices, but also other musical styles, to allow more precise conclusions about the description of voice quality.

This first step enabled us to start looking for acoustic parameters which may account for the perception of different relevant criteria to the perception and the description of voice quality. In particular, we have shown that some variations of voice quality may be related to modifications of the spectral balance, which are not only produced by variations in vocal intensity but which might also be influenced by other laryngeal and articulatory factors. As regards the influence of the glottal source on voice quality, our first analyses of the parameter H1-H2 (difference between the first two harmonics of the derived glottal flow) as well as the Harmonic to noise Ratio have not shown any significant relation with the different modifications of voice quality considered in this study. However, we also dispose of the electroglottographic signals of this same database so that we plan to explore them in a future work. As regards the influence of articulation on voice quality, we have demonstrated that several modifications of voice quality are related to transforms of the vocalic system. We are planning to conduct this experiment again with a more precise measurement tool of vocal tract resonances instead of estimating the formants frequency by inverse filtering. This should also allow us to examine how the higher resonances of the vocal tract influence voice quality by modifying the spectral balance, especially in the singing formant area and in the frequency band around 5 kHz. These analyses should be extended to more singers and to many more data, so as to test the reproducibility of our observations, and to analyse in more detail the interindividual variations of production strategy for a given voice quality. It would also be necessary to extend this experiment to the female and child voice, as well as to explore a broader range of the singer's tessitura.

Lastly, the perceptual relevance of the acoustic parameters correlated to several voice qualities should be tested. This could be done by using a psychophysical approach with synthetic voice examples or by modifying the most relevant acoustic parameters in real voice examples, so as to assess whether the variation of selected acoustic parameters would effectively be related to the perception of given voice qualities.

Acknowledgements

The authors are very grateful to the subjects who volunteered to participate to this project: the singers who were recorded, and the singing teachers who agreed to share their experience without hesitation. The authors also thank Pascale Cheminée for her precious help on the linguistic analyses, Jacques Poitevineau for helping on statistical analyses.

References

Achard, P. 1987. On the Methodology of Discourse analysis and Content Analysis. *Bulletin de Méthodologie Sociologique* 13 : 24-32.

Bhuta, T., Patrick, L., Garnett, J. 2004. Perceptual evaluation of voice quality and its correlation with acoustic measurements. *Journal of Voice*, 18 (3) : 299-304.

Bjorkner, E. 2006. *Why so different ?* Doctoral dissertation. KTH, Stockholm.

Bloothooft, G. and Plomp, R. 1986. Spectral analysis of sung vowels. III. Characteristics of singers and modes of singing. *The Journal of the Acoustical Society of America* 79 (3) : 852-864.

______. 1988. The timbre of sung vowels. *The Journal of the Acoustical Society of America* 84 (3) : 847-860.

Boersma, P. and Weenink, D. *Praat: doing phonetics by computer* (Version 4.2.28) [Computer program]: retrieved from http://www.praat.org/ in 2004.

Busson, S. 2002. *Etude des attributs perceptifs d'un ensemble de pianos*. Master ATIAM. Université Paris 6.

Castellanos, A., Benedíb, J.M. and Casacuberta, F. 1996. An analysis of general acoustic-phonetic features for Spanish speech produced with the Lombard effect. *Speech Communication* 20 : 23-35.

Castellengo, M. 1986. *Les sources acoustiques*. Le livre des techniques du son. tome 1. Editions Fréquences. Paris: 45-80.

Castellengo, M., Chuberre, B., Henrich, N. 2004. Is Voix Mixte, the Vocal Technique Used to Smoothe the Transition across the two Main Laryngeal mechanisms, an independent mechanism ?. *Proceedings of ISMA*, Nara, Japan.

Castellengo, M. and Dubois, D. 2007. Timbre ou timbres ? Propriété du signal, de l'instrument, ou construction(s) cognitive(s). Cahiers de la *Société Québécoise de Recherches Musicales*, Montréal.

Cheminée, P. (2005). Vous avez dit clair? éléments d'une sémantique du timbre des pianos. Cahiers du *LCPE 7*.

Childers, D. G. and Lee, C. K. 1991. Vocal quality factors: analysis, synthesis, and perception. *Journal of the Acoustic Society of America* 90(5): 2394-2410.

Chuberre, B. 2000. *Les registres et passages dans la voix chantee*. Ph.D thesis, University of Nantes, France.

Cross, R. 2007. Aigu classique ou moderne ?. *Journal de l'AFPC*.

David, S., Dubois, D., Rouby, C., Schaal, B. 1997. L'expression des odeurs en français :
analyse lexicale et représentation cognitive. *Intellectica* 24 : 51-83 .

De Krom, G. 1995. Some Spectral Correlates of Pathological Breathy and Rough Voice Quality for Different Types of Vowel Fragments. *Journal of Speech and Hearing Research* 38: 794-811.

Dubois, D., Ed. 1991. *Sémantique et cognition*. Paris, Éditions du CNRS.

______. 1997. *Catégorisation et cognition*. Paris, Kimé.

Dubois, D. 2000. Categories as acts of meaning:The case in olfaction and audition. *Cognitive Science Quaterly* 1: 35-68.

______. 2006. Speaking of colors and odors. An interdisciplinary approach to cognitive and linguistic categorization of color vision and olfaction. In *From psychophysics to semiophysics: Categories as Acts of Meaning : a case study from olfaction and audition, back to colors*. M. Plumacher & P. Holz (Eds.) , Amsterdam , Benjamins: 45-119.

Dubois, D, Giboreau, A. 2006. Descriptors: attributes? labels? terms? names? A contribution of psycholinguistics to sensory evaluation. *Proceddings of 6th Pangborn, Food Quality and preferences*, 17 : 669-672.

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?* Ph.D thesis, EHESS, Paris.

Feng, G. and Castelli, E. 1996. Some acoustic features of nasal and nasalized vowels: A target for vowel nasalization. *The Journal of the Acoustical Society of America* 99 (6) : 3694-3706.

Furui, S. 1986. Research on individuality features in speech waves and automatic speaker recognition techniques. *Speech Communication* 5 (2) : 183 – 197.

Garnier, M. 2003. *Approche de la qualité vocale dans le chant lyrique: perception, verbalisation et corrélats acoustiques*. Master thesis, Université Paris 6.

Gerratt, B. R., Kreiman, J., Antonanzas-Barroso, N. and Berke, G. S. 1993. Comparing Internal and External Standards in Voice Quality Judgments. *Journal of Speech and Hearing Research*, 36: 14-20.

Gobl, C. and Ní Chasaide, A. 1992. Acoustic characteristics of voice quality. *Speech Communication* 11: 4-5.

______. 2003. The role of voice quality in communicating emotion, mood and attitude. *Speech Communication archive* 40(1-2): 189 - 212.

Guastavino, C. 2007. Categorization of environmental sounds. *Canadian Journal of Experimental Psychology*, 61 (1) : 54-63

Hanson, H. M., Chuang, E. S. 1999. Glottal characteristics of male speakers: Acoustic correlates and comparison with female data. *The Journal of the Acoustical Society of America* 106 (2) : 1064-1077.

Harnad, S., Ed. 1987. *Categorical perception*. Cambridge, Cambridge University Press.

Henrich, N. 2001. *Etude de la source glottique en voix parlée et chantée. Modélisation et estimation, mesures acoustiques et électroglottographiques, perception*. Ph.D thesis, Université Paris 6.

Henrich, N., d'Alessandro, C., Doval, B., Castellengo, M. 2005. Glottal open quotient in singing: Measurements and correlation with laryngeal mechanisms, vocal intensity, and fundamental frequency. *The Journal of the Acoustical Society of America* 117 (3) : 1417-1430.

Henrich, N., Bezard, P., Garnier, M., Guerin, C., Pillot, C., Quatrocchi, S., Roubeau, B., Terk, B. 2007. Perception and verbalisation of voice quality in western lyrical singing : contribution of a multidisciplinary research group. *Proceedings of the Third Conference on Interdisciplinary Musicology*.

Hertegard, S., Gauffin, J., Sundberg, J. 1990. Open and covered singing as studied by means of fiberoptics, inverse filtering and spectral analysis. *Journal of Voice* 4 (3) : 220-230.

Husson, R. 1962. *Physiologie de la phonation*. Masson, Paris.

Klatt, D. H. and Klatt, L. C. 1990. Analysis, synthesis, and perception of voice quality variations among female and male talkers. *Journal of the Acoustic Society of America* 87(2): 820-857.

Ladd, D. R., Silverman, K. E. A., Tolkmitt, F., Bergmann, G. and Scherer, K. R. 1985. Evidence for the independent function of intonation contour type, voice quality, and F0 range in signaling speaker affect. *The Journal of the Acoustical Society of America* 78(2): 435-444.

Laver, J. 1980. *The phonetic description of voice quality*. Cambridge University Press, Cambridge.

Lavner, Y., Gath, I. and Rosenhouse, J. 2000, The effects of acoustic modications on the identication of familiar voices speaking isolated vowels. *Speech Communication* 30:9-26.

Liberman, G. and Mattingly, A. M. 1985. The motor theory of speech perception revised. *Cognition* 21: 1-36.

Liénard, J. S., Di Benedetto, M-G. 1999. Effect of vocal effort on spectral properties of vowels. *The Journal of the Acoustical Society of America* 106 (1) : 411-422.

Lindblom, B. E. and Sundberg, J. E. 1971. Acoustical consequences of lip, tongue, jaw, and larynx movement. *Journal of the Acoustic Society of America* 50(4): 1166-1179.

Maffiolo, V., Dubois, D., David, S., Castellengo, M., Polack, J.-D. 1998. Loudness and pleasantness in structuration of Urban Soundscapes. *Proceedings of Internoise*, Wellington.

Maffiolo, V., Chateau, N. 2003. The emotional quality of speech in voice services. *Ergonomics* 46 (13-14) : 1375-1385.

Markel, J.D. and Gray, A.H. 1976. *Linear prediction of speech*. Springer-Verlag, Berlin

Maxim, L., Ganier, M. and Dubois, D. 2007. The metalinguistic activity in the speech of musical experts. *Proceedings of the 2nd Conference on Interdisciplinary Musicology* (CIM), Tallinn.

Mendoza, E., Valencia, N., Muñoz, J. and Trujillo, H. 1996. Differences in voice quality between men and women: use of the long-term average spectrum (LTAS). *Journal of Voice* 10(1): 59-66.

Mitchell, H.F. 2005. *Defining vocal quality in female classical singers: pedagogical, acoustical and perceptual studies*. Doctoral dissertation. University of Sydney

Morange, S. 2005. *Approches structurale, prosodique, psycho-cognitive de quelques propriétés linguistiques subjectives de la parole: l'exemple de trois Montmartrois âgés*. Ph.D thesis, Universite Sorbonne Nouvelle, Paris.

Morange, S., Dubois, D., Fontaine, J.M. 2007. Voices of Caruso: Cognitive evaluation and acoustic analysis of reedited editions. *Proceedings of the Third Conference on Interdisciplinary Musicology* (CIM07)

Omori, K., Kacker, A., Carroll, L., Riley, W. and Blaugrund, S. 1996. Singing power ratio: Quantitative evaluation of singing voice quality. *Journal of Voice* 10 (3) : 228-235.

Pillot, C. 2004. *Sur l'efficacité vocale dans le chant lyrique : aspects physiologique, cognitif, acoustique et perceptif*. Ph.D thesis. Universite Sorbonne Nouvelle, Paris.

Pinczower, R., Oates, J. 2005. Vocal Projection in Actors: The Long-Term Average Spectral Features That Distinguish Comfortable Acting Voice From Voicing With Maximal Projection in Male Actors. *Journal of Voice* 19 (3) : 440-453.

Pittam, J., Gallois, C. and Callan, V. 1990. The long-term spectrum and perceived emotion. *Speech Communication* 9(3): 177-187.

Prame, E. 1994. Measurements of the vibrato rate of ten singers. *The Journal of the Acoustical Society of America* 96 (4,) : 1979-1984.

Rastier, F. 1991. *Sémantique et recherche cognitive*. PUF, Paris.

Rastier, F. 1997. *Meaning and Textuality*. University of Toronto Press Buffalo.

Rioux, V. 2000. Methods for an objective and subjective description of starting transients of some flue organ pipes - Integrating the view of an organ-builder. *Acta Acustica* 86 (4) : 634-641.

Rosch, E. 1973. On the internal structure of perceptual and semantic categories. in; Moore T.E. (ed.), *Cognitive development and the acquisition of language*. Academic Press, New York, London, 111-144

Rothman, H.B., Rullman, F.J., Arroyo, A.A. 1990. Inter-and Intrasubject Changes in Vibrato: Perceptual and Acoustic Aspects. *Journal of Voice* 4 (4) : 309-316

Sotiropoulos, D. 2004. *Analyse acoustique et catégorisation d'un ensemble de qualités vocales pertinentes pour la description de voix lyriques masculines*. Master thesis, Université Paris 6.

Stone, R., Cleveland, T., Sundberg, J., Prokop, J. 2003. Aerodynamic and acoustical measures of speech, operatic, and broadway vocal styles in a professional female singer. *Journal of Voice* 17 (3) : 283-297.

Sundberg, J. 1972. An articulatory interpretation of the singing formant. *STL-QPSR* 1: 45-53.

Sundberg, J., Gramming, P. and Lovetri, J. 1993. Comparisons of pharynx, source, formant, and pressure characteristics in operatic and musical theatre singing. *Journal of Voice* 7 (4) : 301-310.

Sundberg, J. and Nordenberg, M. 2006. Effects of vocal loudness variation on spectrum balance as reflected by the alpha measure of long term average spectra of speech. *Journal of the Acoustical Society of America*, 120 (1) : 453-457.

Titze, I. 2001. Acoustic Interpretation of Resonant Voice. *Journal of Voice* 15 (4) : 519-528.

Titze, I.R., Bergan, C.C, Hunter, E.J. and Story, B. 2003. Source and filter adjustments affecting the perception of the vocal qualities twang and yawn. *Logopedics Phoniatrics Vocology* 28 (4) : 47 – 155.

Vogel, C. 1991. Expression langagière de l'expertise : problèmes d'analyse. *Intellectica* 12 : 65-100.

Von Bismark, G. 1974. Timbre of steady sound: a factorial investigation of its verbal attributes. *Acustica* 30, 146-159.

Von Helmholtz, H. L. 1974. *Théorie physiologique de la musique fondée sur l'étude des sensations auditives*. Masson, M., Paris.

Vurma, A., Ross, J. 2002. Where Is a Singer's Voice if It Is Placed “Forward”?. *Journal of Voice* 16 (3) : 383-391.

Wapnick, J., Ekholm, E. 1997. Expert consensus in solo voice performance evaluation. *Journal of Voice* 11 (4) : 429-436.