



Spectral correlates of carrying power in speech and western lyrical singing according to acoustic and phonetic factors

Claire Pillot-Loiseau, Jacqueline Vaissière

► To cite this version:

Claire Pillot-Loiseau, Jacqueline Vaissière. Spectral correlates of carrying power in speech and western lyrical singing according to acoustic and phonetic factors. 15th International Congress of Acoustics (ICA), Sep 2007, Madrid, Spain. pp.1-6. hal-00530207

HAL Id: hal-00530207

<https://hal.science/hal-00530207>

Submitted on 27 Oct 2010

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.



SPECTRAL CORRELATES OF CARRYING POWER IN SPEECH AND WESTERN LYRICAL SINGING ACCORDING TO ACOUSTIC AND PHONETIC FACTORS

PACS: 43.75.Rs

Pillot, Claire¹²; Vaissière, Jacqueline²

¹Institute A. Vernes, 36 rue d'Assas, 75006 Paris, France; pillot@msh-paris.fr

²Laboratoire de Phonétique et Phonologie, UMR 7018, 19, rue des Bernardins, 75005, Paris, France; jacqueline.vaissiere@univ-paris3.fr

ABSTRACT

In order to define the variability of carrying power (sometimes called “vocal effectiveness”) indexes in speech and singing, an acoustic analysis of vowels, sentences, singing exercises, and lyrical piece spoken and sung by 23 singers, was conducted. Two parameters were measured: (i) the difference in amplitude between the highest harmonic between 2 and 4 kHz and the one between 0 and 2 kHz (“Singing Power Ratio”, Omori *et al.*, 1996); and (ii) the difference between the singing formant (near 3000Hz) and a spectral minimum that follows it. Results showed that the bringing together of the third and fourth formants between spoken and sung vowels corresponded to the singing-formant, followed by the spectral minimum more significant in singing. The singing-formant was correlated with high values of Singing Power Ratio (-16dB) and high values of the difference in amplitude between the singing formant and the spectral minimum (32 dB). Only this last parameter differentiated untrained from trained subjects. High values of these two indexes were associated with great intensities, spoken [i], and singing in comparison with speech. The same values were obtained whatever the language. These two parameters varied depending on the task (sustained vowel versus singing exercises) and on vocal category.

INTRODUCTION

Acoustics can quantify the characteristics of singing as compared to speech. Vocal effectiveness, mainly defined in singing by the carrying power, is one of the main objectives to achieve without too much vocal effort, in an unamplified vocal technique such as Western classic singing. This study shows how the carrying power in speech and in Western lyrical singing can be quantified in terms of phonetic and acoustic factors. The parameters of carrying power and vocal effectiveness may be useful to quantify the progress of a learning singer, or to evaluate the degree of vocal rehabilitation of a patient singer.

METHOD

Recordings

Among the 23 recorded subjects (12 male and 11 female subjects; average age: 38,3 years; standard deviation: 8 years), 17 were professional and were trained for Western lyrical singing. 3 were semi-professionals and 3 amateurs. They were 3 basses, 2 baritones, 7 tenors, 2 *contralti*, 4 *mezzi-soprani* and 5 *soprani*. 12 of them were a part of Choirs of the Opera of Paris. Most of them were French (17) and there were also 2 Italians, 3 Americans and 1 Greek.

The protocol was elaborated to allow for the study of the variation of two carrying power parameters. The corpus was composed of: (C1) spoken vowels [i], [a] and [u] embedded in a sentence to be spoken like - “Je répète [a] comme dans “âme” – I repeat [a] as in “âme”; (C2) *mesa di voce* sounds (same fundamental frequency by vowel, and increase then progressive decrease of intensity) on the sustained vowels [i], [a] and [u], sung in low, medium and high

pitches; (C3) vowel [a] in the singing exercise of Rossini, known for singers; (C4) one sentence by subject in French, freely chosen by the singer, read it in a normal, declaimed, and sung voice with its original melody; (C5) a lyrical piece of language chosen by the subjects

C1 was set to compare [i], [a] and [u] in a focused position (isolated one). C2 allowed to study the influence of the vowel and the intensity (at a constant fundamental frequency) on the parameters of vocal carrying power. The whole vocal range of every subject was studied through three fundamental frequencies. C3, C4 and C5 allowed for the study of the variation of carrying power parameters in a musical context. C3 allowed us to examine the influence of the fundamental frequency and the intensity. C4 allowed for the comparison of the two carrying power parameters in three contexts (spoken, declaimed, sung), in continuous speech. C5 reveal values in a more musical context, comparable with those of the other items of sometimes different language.

The recordings, of one hour by subject, were made in the sound-treated booth of the Laboratory of Phonetics and Phonology (LPP) in Paris, using a condenser cardioid unidirectional microphone (audio Technica ATM 33a; range of frequencies: 30-20000 Hz) placed 40 cm from the singer, a preamplifier, and an amplifier of measure (Sony), a digital tape recorder (Aiwa HD-S1); frequency of sampling: 44,1 kHz or 48 kHz on 16 bits.

Spectral indexes of carrying power

The recorded sounds were analysed by means of the software Computerized Speech Lab (CSL), Kay Elemetrics, the spectra averaged in the long term (LTAS) and FFT were calculated. The singing formant was considered to be represented by the highest harmonic between 2000 and 4000 Hz. This value was manually determined from the spectra (Blackman window, 2048 points (46ms), absence of pre-emphasis). In all analyzed sung signals, the singing formant was followed by a spectral minimum. Two parameters of carrying power and vocal effectiveness, SPR (a conventional measure) and MIN (a new measure proposed by ourselves), were determined: 1) SPR is the difference between the highest harmonic between 2 and 4 kHz and the one between 0 and 2kHz ("Singing Power Ratio", Omori *et al.*, 1996, [2]), figure 1; 2) MIN is the difference between the amplitude of the singing formant, as defined above, and that of the spectral minimum which follows it, figure 1.

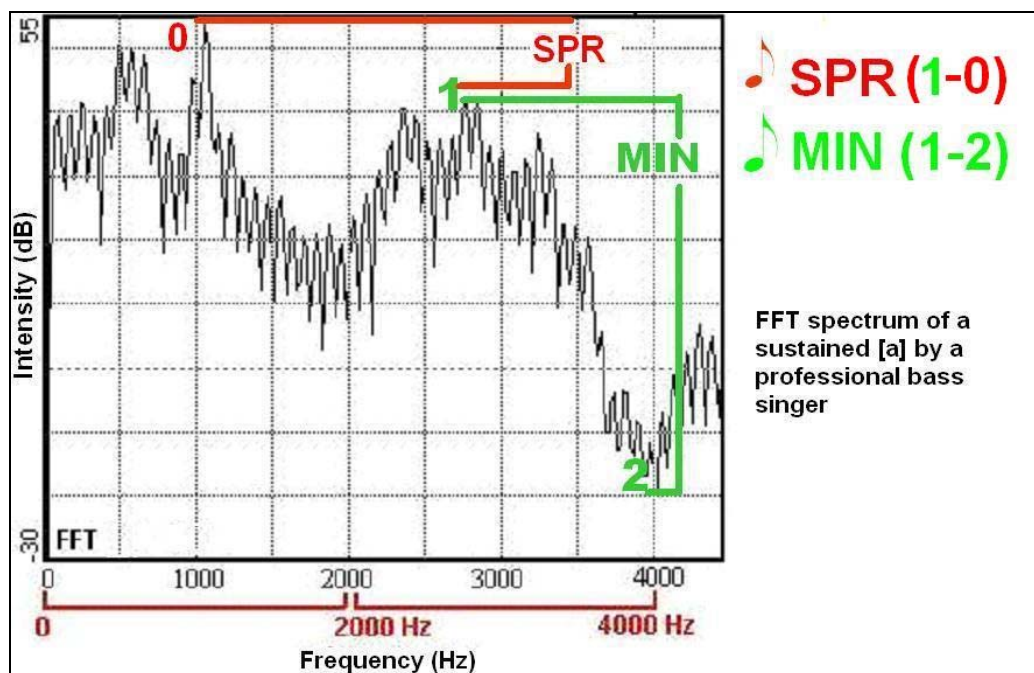


Figure 1.- Determination of the "Singing Power Ratio" (SPR) and the difference between the amplitude of the singing formant and that of the spectral minimum which follows it (MIN)

FACTORS OF VARIATION OF CARRYING POWER

In the sung material by our trained subjects, the singing formant frequency has been found independent of the vowel, the type of production, and the vocal category. It was located between 2400 Hz and 3500 Hz, and its frequency increased slightly with higher pitch. The peak corresponding to the singing formant was followed, as stated before, by an important spectral minimum. Its presence corresponded to the biggest values of the SPR (Omori *et al.*, 1996) towards 16dB, and with the biggest ones of MIN towards 32 dB on average for the professionals. The variability in that measure was due to the various items and to the vocal category of the subjects.

Acoustic parameters

Fundamental frequency and intensity

The fundamental frequency influenced the parameters of carrying power differently according to the item and the subject: we noted sometimes an absence of influence of this fundamental frequency (singing exercise of Rossini for alti: $r_{26}=0.2$, ns), sometimes a positive correlation with the carrying power parameters as for the nine mezzo-soprani and soprani ($r_{117}=0.78$, $p<0.05$).

The intensity had a significant influence on the parameters of the carrying power (positive correlations), especially for the tenors and the soprani. This means that the more the intensity increases, the more the carrying power increases (Table I).

Table I.- Correlations between the intensity and SPR and MIN for sounds *messa di voce*.

*: $p<0.05$; **: $p<0.0001$. N: number of measured and calculated points.

Correlation between intensity and ...	5 Bass and Baritones (N=65)	7 Tenors (N=91)	2 Alti (N=26)	9 Mezzo-soprani and Soprani (N=117)	All (N=299)
SPR	0.33 *	0.47 *	0.35 *	0.4 *	0.3 **
MIN	0.7 **	0.52 **	0.56 **	0.68 **	0.54 **

Singing with regard to speech

Figure 2 shows the spectra corresponding to a spoken and sung [a] with a similar fundamental frequency by a bass professional singer: the clustering of third and fourth formants contributed to the creation of the singing formant between 2000 Hz and 3000 Hz, followed by a more important spectral minimum in the sung vowel (*) as compared to the spoken one: the values of SPR and MIN were significantly great for all the vowels sung with regard to their spoken equivalents. For the professional singers, the same modifications occurred for the sentences from speech to singing, the declaimed context showing intermediate values of the parameters of carrying power.

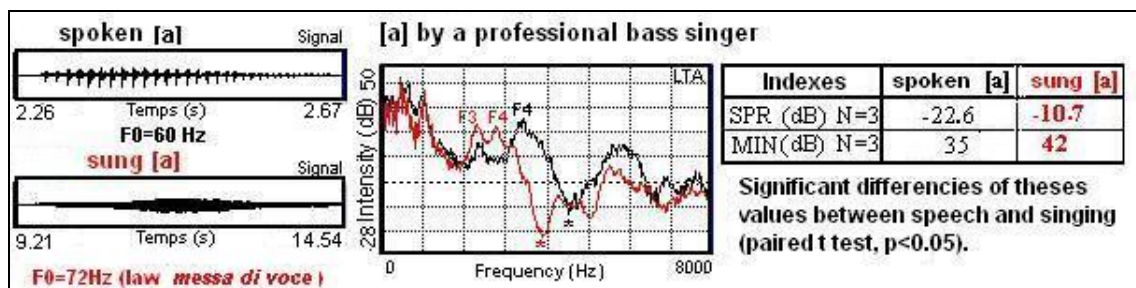


Figure 2.- Values of carrying power parameters, signals and spectra (LTAS) of one spoken (black) and sung (red) [a] by a professional bass singer. Note the clustering of the third (F3) and the fourth (F4) formants from speech to singing, and the lowest value of the spectral minimum (*) in sung vowel.

Phonetics parameters

Vowel

There was a certain variability in the obtained results. The values of the SPR (Omori *et al.*, 1996) followed the order [i]>[a]>[u] when spoken, and the MIN generally followed in the order [a]>[i]>[u]; only the difference between [a] and [u] was significant for the SPR. These data mean that the carrying power was generally the least important for [u] spoken (figure 3). The differences between the sung vowels became more and more blurred with higher fundamental frequency.

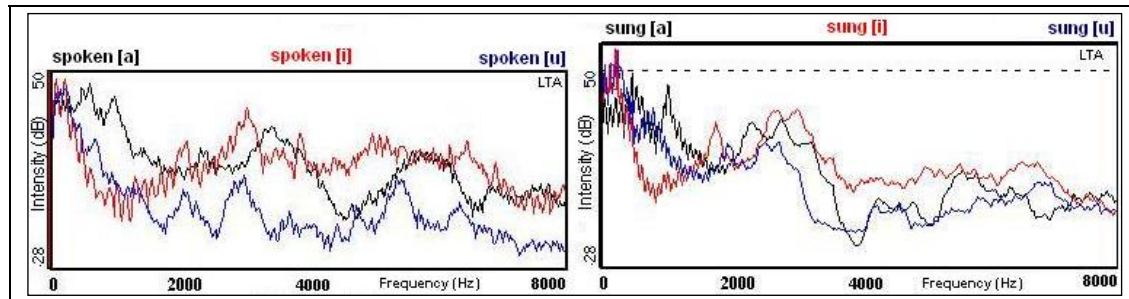


Figure 3.- Spectra averaged in the long term (LTAS) of [a], [i] and [u] spoken (to the left) and sung (to the right) by a professional bass singer: [i] has an intrinsically higher carrying power, because of the dominance of high frequencies for the [i] sound, followed by [a], then by [u], which is a grave vowel. These differences between vowels decrease in singing.

Sentence in French and lyrical piece in another language

There were no significant differences of the values of SPR and MIN in the sung sentences and the fragments of a lyrical piece (t tests between the sung sentence and the piece for our 23 subjects: $t_{23}=0.4$; $p=0.7$: figure 4).

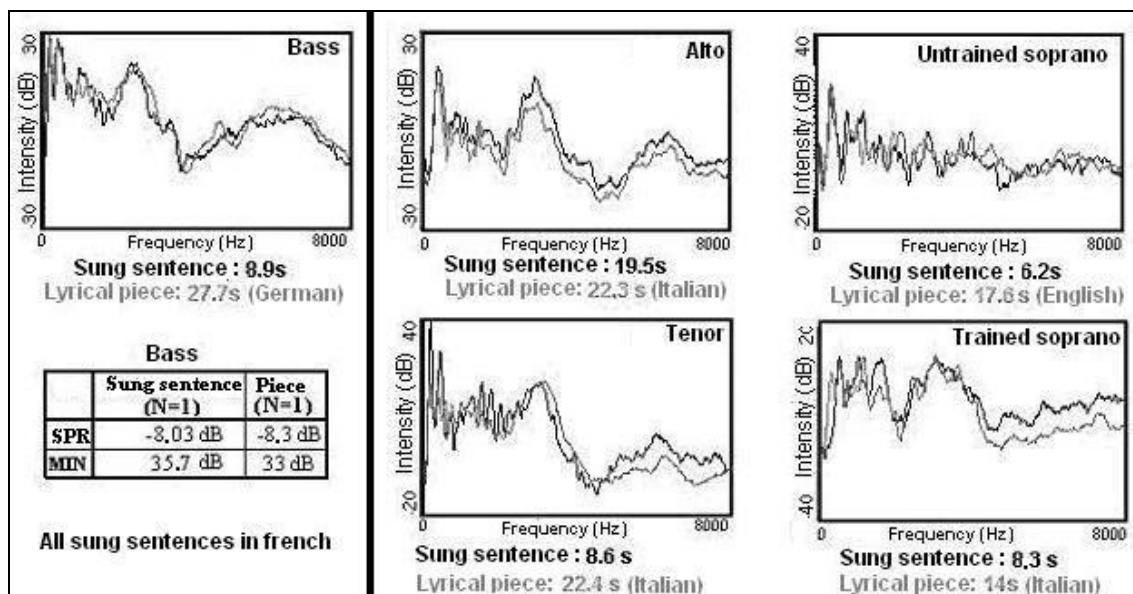


Figure 4- Spectra averaged in the long term in the sentence and in the piece sung by 5 subjects of 5 different vocal categories. The values of carrying power are almost identical in spite of differences in duration and in languages.

Variability inter-subjects (vocal category and amateurs / professionals)

The carrying power parameters seemed higher for the 12 male singers. The influence of the fundamental frequency, the intensity, the vowel and the change from the speech to the singing context on SPR and MIN varied according to the vocal categories. Only the MIN was significantly different between a professional soprano and one amateur (figure 5). The amateur increased the intensity of all the harmonics when the sound level increased, whereas the professional soprano increased selectively the zone corresponding to the singing formant. Furthermore, the professionals had the highest values of the SPR, and of MIN (Table II).

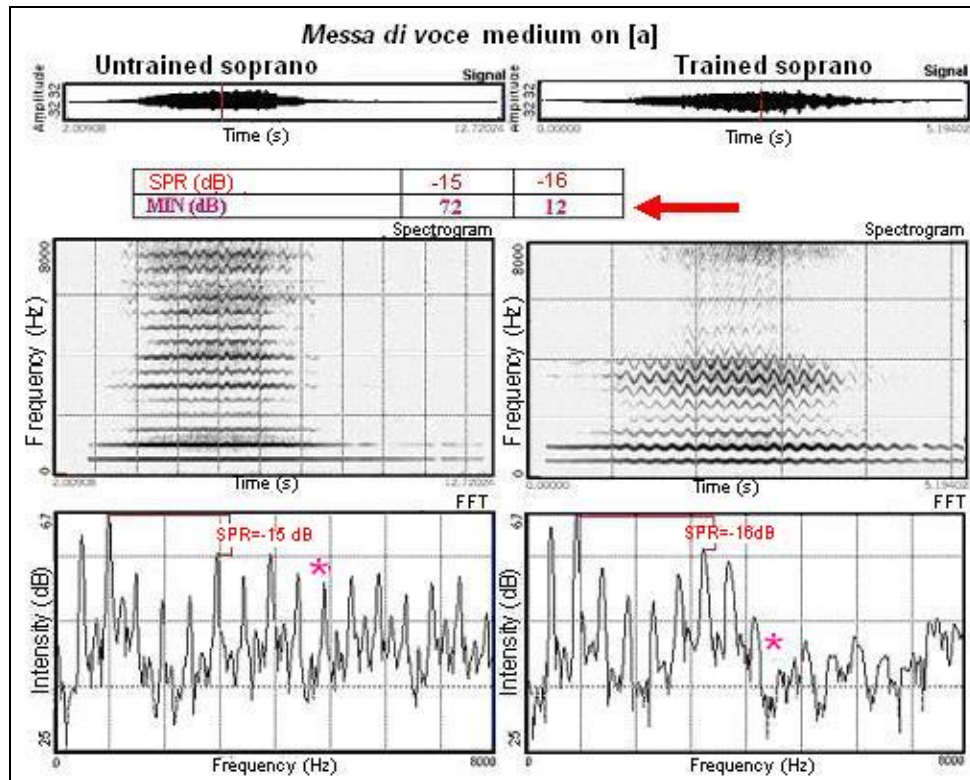


Figure 5.- Signals, spectrograms and FFT spectra of the strongest intensity of sounds *mesa di voce* sung on [a] with a fundamental frequency medium by a professional soprano, (right) and a soprano amateur (left). Only the MIN (*) distinguish these two productions.

Table II.- Examples of values of the MIN and the SPR according to the subjects' vocal level.

(dB)	Professionals			Semi-professionals			Amateurs		
	Bass	Mezzo-soprano	Soprano	Bass	Mezzo-soprano	Soprano	Soprano	Mezzo-soprano	Alto
MIN	35.7	30	32	22	15	25	13	20	16
SPR	-8	-2	-5	-10	-9	-22	-13	-15	-20

DISCUSSION

Omori *et al.* [2] studied the SPR from 37 singers and 20 non-singers from spoken and sung [a] to comfortable intensity and fundamental frequency. They obtained, professional singers and semi-professionals merged, mean values of 19 dB for the male singers, 23 dB for the female singers in spoken [a], 12 dB for the male singers, 16 dB for the female singers for sung [a], and 22 dB for spoken and sung [a] by the non singers (merged men and women).

Our values of SPR agree on average with those of Omori *et al.* [2]; they are however more variable because of the wider variety of conditions and the larger number of vocal categories investigated. Most of the authors restricted their analysis to a sustained vowel (mostly [a]). None of the preceding studies quantified, as we propose here, the difference in intensity between the singing formant and the spectral minimum which follows it (MIN). The influence of the fundamental frequency on the values of carrying power parameters, and thus on the emergence of the singing formant, was debated in the literature: Schultz-Coulon *et al.* [3] asserted that the singing formant prominence decreased as the fundamental frequency increased, while Omori *et al.* [2] evoked an absence of link between fundamental frequency and the SPR.

The authors agreed on the other hand to assert the positive correlations between the intensity and our indexes of carrying power: for Schultz-Coulon *et al.* [3], the singing formant increased with the intensity, less clearly for the female than the male singers. Bloothoof *et al.* [4] investigated the relationship between the singing formant and the overall intensity. They founded that the singing formant intensity increased between 16 and 19 dB for an increase of 10 dB in the global intensity, and the increase was dependent on the mode of phonation, on the singer and on the vowel. Sundberg [5] confirmed these relationships.

Bloothoof *et al.* [4] and Sundberg [5] studied the influence of the vowel identity on the emergence of the singing formant (calculated through the difference between the amplitude of the first formant and that of the third): [i] sung, at low frequency, evidenced the best emergence of the singing formant, closely followed by sung [a], and [u]. According to Sundberg [5], for high pitches, the intensity level of this formant of sung [a] could be equivalent to that of [i]. The order [i]>[a]>[u] was also valid for the spoken vowels studied by this author. Our data thus agree with these studies.

The results of the authors which compared the spoken and sung productions on one hand, those of trained and untrained singers on the other hand, agree with our data: Omori *et al.* [2] reported that the SPR was significantly higher for the trained subjects, and for the singing with regard to the same spoken vowel where values similar to ours were obtained (from 18 to 25dB). According to these authors, there are no significant differences between the professional singers and the amateur ones. Besides, the notable variability of all obtained values was more important for the female singers (Schultz-Coulon, *et al.*, [3]) because of higher fundamental frequency and the large distance between adjacent harmonics (Sundberg, [5]). Finally, as we have already evoked that, Sundberg [6] showed, in the singing context, the moving closer of third, fourth and fifth formants, which were separated in the speech context (except that F3 and F4 may be clustered for the vowel [i], around 3000 Hz for male speakers, in speech).

CONCLUSIONS

An objective and complete estimation of the spectral parameters of the carrying power is possible only if we suitably take into account their great variability (fundamental frequency, intensity, vowel identity, language, vocal category and vocal level of the subjects). Besides, it is advisable to take into account the presence of the spectral minimum that follows the singing formant, in order to better estimate the vocal carrying power from an acoustic point of view, as shown in the present paper.

References

- [1] C. Pillot: Sur l'efficacité vocale dans le chant lyrique : aspects physiologique, cognitif, acoustique et perceptif. Dissertation in Phonetics, Paris III Sorbonne Nouvelle University (2004) 418 pages.
- [2] K. Omori, A. Kacker, L. M. Carroll, W. D. Riley, S. M. Blaugrund: Singing Power Ratio : quantitative evaluation of singing voice quality. *Journal of Voice* **10**, No.3 (1996) 226-235
- [3] Schultz-Coulon, R. D. Battmer, H. Riechers : Der 3KHz Formant-ein Mass für die Tragfähigkeit der Stimme? [The 3kHz Formant, a measurement for the vocal effectiveness ?]. *Folia phoniatrica* **31**, No.4 (1979) 291-313.
- [4] G. Bloothoof, R. Plomp: The sound level of the singer's formant in professional singing. *Journal of the Acoustical Society of America* **79**, No.6 (1986) 2028-2033.
- [5] J. Sundberg: Level and center frequency of the singer's formant. *Journal of Voice* **15**, No.2 (2001) 176-186.
- [6] J. Sunberg: The science of the singing voice. First Edition, Northern Illinois University Press, Dekalb, Illinois, (1987) 216 pages.