Towards the conception of a vocal synthesizer: Problems and prospects

Anastasia Georgaki
Ionian University , Music Department
Old Fortress, 49100 Corfu, Greece
georgaki@ionio.gr

Abstract

In this paper we will discuss current research on the synthesis of the singing voice, its technical problems and esthetic issues, and the perspective of creating a vocal synthesizer which could be accessible to composers, musicians and musicologists, as a creative tool for composing, performing, and interpreting forgotten voice techniques of various cultures and bygone times.

1 Introduction

" Je m'occupe maintenant à trouver la manière de faire prononcer les syllabes aux tuyaux d'orgue. J'ai desja rencontré les voyelles a, e, o et u mais il me fait bien de la peine, et puis j'ay treuvé la syllabe vê et fê. Je ne sçay si je paourray prendre le losir de trouver les autres consonnes, à raison des différentes experiences qu'il faut faire sur ce sujet , lesquels estant de coust, je laisseray le reste à ceux qui voudront passer outre."

(Marin Mersenne, Musurgia Universalis)

Since the late antiquity, scientists have been interested in the conception of automata such as talking machines, centering their interest on the speaking voice and its utilities; voice has been always the privileged model of automatophons [Tubach J.P., 1989; Georgaki, 1998a].

In the last fifteen years, there has been a special research interest on the synthesis of the singing voice; the conception of synthesizers has been quite fruitful, due to the exploitation of the data extracted from the restricted analysis of the singing voice, and due to a big effort by scientists to separate the speaking voice from the singing voice by centering their interest on their acoustical and cultural differences: frequency, displacement of the formants, vibrato, attack, spectral envelope etc. [Benade, 1986; Fant, 1973; Sundberg , 1989].

In this article we would like, on the one hand, to proceed with the presentation and evaluation of the current research in various institutes concerning the synthesis of the singing voice, in order to outline the technical problems and orientation of the research in our days; on the other hand, we would like to propose our ideas about the conception of a vocal synthesizer which could be used in various domains of the music research, creation and performance.

2 EXPLORING THE CURRENT RESEARCH ON THE SYNTHESIS OF THE SINGING VOICE

Since 1980 various projects have been carried out all over the world having as centering point the synthesis of the singing voice [Mathews et al, 1961; Chowning, 1981; Sundberg, 1989; Rodet et al, 1984, 1985, 1987,1988; Gael, 1990; Rossitier et al,1994; Berndtsson 1995; Carlsson et al, 1991, Tisato et al, 1991; Cook Perry, 1989, 1993; Depalle et al 1994, Paabon 1994; Lomax 1996; Pierucci et al, 1997; Macon, 1997; Gibson 1998].

In fact, every project has its own goals and directions. Among the various projects using a multitude of techniques and rules, concerning the analysis and synthesis of the singing voice, we have selected as a vehicle of our presentation the research projects which tend to have a complete point of view about the synthesis of the singing voice (in order to speak about 'models' which fulfill the expectations of a synthesizer not only from the acoustical point of view but also from the phonetic one).

According to our research, we concluded that the most complete vocal synthesizers in the domain of the synthesis of the singing voice are: *Musse/Rulsus* (Sweden, J. Sundberg), *Chant* (France, X. Rodet) and *Spasm/singer* (United States, P.Cook) because of reasons such as the original analysis/synthesis technique, the appropriate rules for singing, the interface for musical applications and the performability/utility by the users, as well as their prospects [Georgaki, 1998a].

In order to outline the current research on synthesis of the singing voice we would like to make some remarks concerning the advantages and status of every synthesizer, along with its applicability in contemporary music:

a) one of the most complete "cognitif" synthesizer is the *Mu(s) se/Rulsus* [Sundberg, 1987;1989] as it is equipped with many rules which are describing classical singing. It gives to the user the possibility to produce his own sung phonemes, words and phrases of reasonably good quality, allowing for vocal expressivity [Gael, 1990, Carlsson 1991, Berndtsson, 1995] and it is used also for the synthesis of polyphonic choral singing. One of its cues is the capacity of controlling the fundamental frequency and the formants with more speed and precision than human singers.
b) other research projects, like *Chant* [Rodet et al, 1985, 1995] have been oriented towards more artistic applications, equipped with the proper interface and environment, in order to afford software tools to the composers or reconstruct ambivalent voices of the past [Depalle et al,1994]. More precisely, Chant [X.Rodet et al., 1984] is a software program designed for the compositional needs rather than for the scientific ones, as it stresses the continuity between sound processing and synthesis by filtering synthetic or real voices and instruments.
c) the *Spasm/Singer* model [Cook, P. 1993} is more advantageous compared with the others, because of its proper control environment: it offers to the user a subtle control on the parameters of the vocal signal which are related directly to the vocal pedagogy and physiology of speech (tongue's or lips' position, the form of the vocal tract, the vocal effort , etc.) through a user friendly interface (the form of the vocal tract on the screen) .
d) Despite the importance of the research projects on the singing voice, the majority of these projects are based on the analysis- synthesis of the classical singing technique allowing for some exceptions such as [Tisato 1991; Rodet 1985; Kamarotos et al, 1994] who have been studying extra-European voice techniques (Diphonic singing, Thibetan singing, or traditional Greek singing).

In any case all these projects, and especially the models in which we have been referring to more extensively, differ not only in their synthesis technique or the implemented rules (describing singing) but also in the control interface and the resulting sound (every model has its own particular voice signature). They differ also in the performability of the model and its applications in the computer music field (composition, psychoacoustics, vocal training and education, or a powerful tool for performance).

Some acoustic examples will give evidence for these observations.

3 THE TECHNICAL PROBLEMS AND THE SPECIFICITY OF THE VOCAL SIGNAL

In this paragraph we are going to outline the basic technical problems for the synthesis of the singing voice.

The basic points of the philosophy for designing a vocal synthesizer are the following: the synthesis technique, the rules and the interface, as well as its control system. The analysis-synthesis techniques define the simulation and the rules ensure a better articulation as the interface allows the expression.

3.1 The complexity of simulation of vowels, consonants and their adjunction in singing phrases

The first and basic technical problems are related with the complexity of the vocal signal and more specifically to:

a) the huge quantity of parameters and data, describing the complex voice singing model, related to the incapacity of the machines to elaborate these satisfactorily (for example in order to have an entire command of the French language we must sample about 2200 sound units , just for one type of voice only) [Depalle, 1994]. This is one of the major cues for differentiating the vocal signal from an ordinary instrumental signal, as the voice is closely connected to the human being, not only from the physiological point of view but also from the acoustic one. In other acoustic signals, it is not necessary, in the same detailed manner to describe the formant trajectories or the microvariations of the signal (which, in the case of voice, are related very closely with the biological function of the vocal apparatus).

b) the specifity of the voice concerning biological functions of the human body (organic and psychological) which affect the timbre , the intensity and articulation of the voice. For example, any aleatoric microvariations due to stress or other factors, influences the periodicity of the vocal signal. [Fonagy 1987; Rodet 1994].

c) the fact that every language has its proper phonetic rules and phonemes renders the creation of an international phonetic database very difficult (a big vocabulary of phonemes and diphones in several languages and an interdisciplinary connection between them), and for evident reasons prevents the commercialisation of a vocal synthesizer.

3.1.1. In the search of the proper technique of analysis- synthesis

Apart from these general technical problems which relate to the conception of a 'universal'[1] synthesizer lets focus our interest on the current scientific research and its problems.

a) The researchers are still experimenting with several methods of analysis- synthesis in order to find the proper combination of techniques, which could better describe the vocal signal during singing and be adapted in musical applications. For the moment researchers haven't yet achieved the refinement of the techniques concerning the analysis -synthesis of the vocal signal, and in some cases they either try to establish new techniques addressed to some particularities of the vocal signal, or they use existing classical sound synthesis techniques. For example one of the techniques of estimating the spectral envelope, although it allows the precise extraction of formants of high frequencies of female voices, poses many difficulties in the level of analysis [X. Rodet, 1992].. From the other hand we are confronted with problems of non- linearity of the vocal signal [P.Cook, 1996]..

b) The other technical problem is the insuffiency of data extracted by analysis, concerning the nature of the vocal signal, the behaviour of the vocal cords and the resonators and more specially in the case of physical modeling, the dynamic movement of the tongue and the articulators , the air flux in the vocal tractus (interior form of the vocal tractus,

---

[1]In the words of Feruccio Busoni : L'esthètique musical, Minerva, Paris, 1989: ..."Louons les novateurs et les libérateurs, si mince soit leur pouboir. Car quelle serait la machine, que les hommes découvriraient et feraient fonctionner, qui ferait retentir des millers des voix? Où est et sera jamais la technique qui ferait jouer les mille registres de l'orgue universel?"

tongue and articulators, vocal cords) or the facial muscles which help the expression during singing. [Cook, 1996].

3.1.2. The singing vowels and consonants

After the fastidious study of some vocal models (or techniques of synthesis) and especially of the models Spasm, Chant and Musse [Georgaki, 1998a] we can assume that :

a) the synthesis of vowels is very satisfactory (especially in intermediate pitches) even if we perceive sometimes a synthetic attribute due to the trajectories of formants. The only problem that has to be solved, concerning the vowels, is the dynamic simulation of the signal and, especially, the control of the transition from one vowel to another.

b) The synthesis of consonants , static or dynamic, remains a problem to be solved by scientists, because of the complexity of the signal itself. Consonants are signals which evolve very quickly in time, and the pertinent tools for their analysis haven't been perfected yet. We cannot preserve the notion of formant trajectory for the control of a consonant ; the only solution is that during the emission of a consonant we extend the formants corresponding in every vowel concerned. The formants models (e.g. Chant) poses many problems in this case, and the researchers are trying to find the solution for the control of the trajectories of formants related to the consonants. The most difficult consonants to be synthesized are the fricatives and plosives (b, d, g). The sound consonats like (f, s, ch) are noisy and in order to synthesize them, researchers may calculate and use aleatoric or pseudo-aleatoric numbers by multiplication, testing the distribution energy [Depalle, 1996].

The result of the synthesis depends also on the conditions and quality of recording and the evolution of quick- rapid formants in time.

The next step of our discussion on the technical problems of the singing synthesis, is the adjunction between consonants and vowels in singing phrases as also the way that energy is distributed between diphones which determine the meaning of a word [Zera and al 1984; Rodet and al , 1988]

3.1.3 The adjunction between vowels and consonants in singing phrases

Now, in order to synthesize a word, or a complete singing phrase (in the case of singing the problem of intonation is already defined by melody), apart from the transition problems (consonant- vowel) or from one diphone to the other, we would like also to underline two other major problems:

a) the control of the evolution of the fundamental frequency. In order to describe the passage from one note to another (despite the problems that may be posed by the consonants) we must take care not only of the formant trajectories but also of the vibrato rate which determines along with *jitter*[2] the naturalness of the vowels. If we resolve the problem of trajectories between the phonemens we will be able to interpolate very easily between different registers.

b) We cannot ignore of course the affinities of the formants' vartiations which affect the quality of the synthesized singing phrase because of:
   - the unique identity of formants for every individual
   - the constant evolution of the articulation during singing speech
   - the dependence of the formant frequency to the fundamental frequency and the intensity [Fant, 1973; Sundberg, 1987].
   - the dependence of the transfer function to the dimension of the vocal tract [Fant, 1973; Sundberg, 1987]

Regarding musical applications, the need imposed by composers for the construction of quasi- vocal timbers, must urge researchers to invest more on the utilisation and control of formant bands for the voiced sounds, and of the noisy sounds for the consonants.

3.2 The control

Researchers are still preoccupied with problems concerning the control of a synthesized singing phrase;

---

[2]The *jitter* is a non- periodic modulation of the phonation frequency which is refered to the aleatoric variations of the waveform. The vocal jitter is minimal in a proffessioanl singer's voice but it affects the quality of the perceived voice.

a) on a first level, the most important element that they try to assume by manipulating parameters in the frequency domain (frequencies, bandwidths, amplitudes of formants) is that the control must be coherent with the direct perception.
b) on a second level researchers are preoccupied with rules of evolution of the formants, rules of interactions between the parameters or questions concerning the reception of control signals from physical peripherals, in order to introduce them into the synthesizer.
c) on a third level they try to ameliorate the control of the oscillation of the glottis, of respiration muscles and other acoustic mechanisms. Researchers must also construct rules related to the vocal cords' tension during singing, as also rules which describe the interaction between vocal effort (muscles, cords etc..) and vocal result (articulators).

In brief, the problem of the control can be treated not only by the physiological point of view but also from the biological one, which in collaboration with the researchers of the cognitive sciences will give the proper informations about the influence of psycological situations on singing interperation and the importance of the body anatomy on singing (resonators, tongue, vocal cords etc..).

3.3. Perception related problems

Lastly one of the problems, to which we are going to refer concerning the synthesis of the singing voice, is the problem of perception. The voice, more than any other instrument, is related very closely to the ear and the ear has the capacity to recognise whether a sound is artificial or not. These perceptive problems come forward because of:

a) the lack of naturalness in the dynamic evolution of the vocal signal due to the lack of rules needed to describe the complexity of the human activities during singing . Among the most perfect sound examples that we have in our archive[3], and concerning of course voices which are conceived in abstracto, we can conclude that: the synthetic singing voices lack naturalness. (apart from some exceptions, cf. Rodet 1981, which don't make use of consonants and are sung in high

---

[3]Kindly offered to us by the creators of the models

frequencies, where normally the intelligibility of voice is altered).

b) the sensibility of the ear which can discriminate the most subtle differences between the spectre of a real and synthetic voice (critical bands, feedback system between voice and ear.)

4 Towards the instrumentalisation of the voice: conception of a vocal synthesizer

Having already mentioned some aspects of the problems of the singing synthesis, we would like, in this last paragraph of this paper, to discuss the conception of a vocal - synthesizer, in instrument form [Cadoz,1994; Dufourt 1996] by formalising the vocal "gesture", which could be used not only by composers but also by the musicians [4].

The multitude of models and techniques doesn't provide the musical world with a solution. What is needed is a complete model that could give the user the possibility to reproduce, by the means of a written musical phrase, voices of a good quality which can be extended in several registers and treated by different techniques; this has not yet been materialized.

Let us imagine, for example, the way in which we could synthesize a singing phrase of an aria (for example 'lasciate mi morire', Cl. Monteverdi , *Ariadne in Naxos*) in different registers , and also in a different voice technique other than this of classical singing.

On the other hand our expectations of this vocal synthesizer are not restricted only in the domain of research (use the computer as a tool for the simulation and extrapolation, hybridisation of singing phrases or resurrection of forgotten voice techniques) but also as a powerful tool for composers implicated in vocal compositions and in the domain of performance. More specifically, after all these research projects done on the synthesis of the singing voice we need to start conceiving 'instruments' designed for the vocal synthesis during musical performance.

In line with the research of the "lost instrument" [Dufourt, 1996] our ambition here is to discuss the possible form of a manipulable vocal synthesizer which can reproduce a big palette of vocal sounds, control prerecorded voices, imitate a cultural vocal model,

[4] Who ‘dream’ of a vocal electronic instrument for performance.

and could give to the composer the possibility of investigation and experimentation (sound chimeras) and to the performer the liberty of expression.

From this point of view, we want to discuss the eventual conception of a musical tool at the borders of an instrument (with the proper flexible interface and control) in order to make more evident the relation between computer music research and creation/performance (leaving aside programming languages etc..).

Since it's not our objective today to talk about the exact form of the final synthesizer, we will limit the discussion to tour anticipations. We would like to witness the creation of a synthesizer that could provide us with the most concrete model for the voice in synthesis in the field of music.

Our first idea, which has been discussed with some researchers in the field, is the conception of a vocal synthesizer which could:

- - reproduce not only a wide palette of timbres (by the means of diphones), but also cover a wide range of registers whilst preserving timbre homogeneity between them.
- - be equipped with the proper rules of different vocal techniques and the appropriate modal musical systems (ancient Greek modes, ecclesiastic modes, Indian modes, etc.)
- - have the possibility to combine the singing technique and languages (for example combine the vocal technique of Byzantine singing with Portuguese language)
- - be used not only like a studio instrument but also as a performance one (like the analog and digital synthesizers)

In order to enrich this 'virtual vocal singer' with elements concerning not only timbre, but also the technique and the language, we must extract the data by analysis studying different vocal models (thibetan singing, byzantine singing, algerian singing, Mongolian singing etc..) and implement all these techniques on the computer

In order to justify our ideas, about the eventual conception of a vocal synthesizer, we are going to discuss the construction this instrument according to the classical instrumental model of the acoustician E. Leipp [Leipp, 1989] which depends on the:

a) perceptive imperatives,
b) fabrication imperatives,
c) anatomo- physiological imperatives,
d) commercial imperatives and
e) the liberty domains concerning pitch, intensity, timbre, form
f) the possibility to implement new vocal techniques in a cognitive form.

According to the *perceptive imperatives*, music is destined to be heard and we must avoid surpassing certain limits which could eventually deform the vocal character. We must not neglect our auditory system and the Fletcher- Munson's curves. According to Morozov's experiments [Sundberg, 1981] has been found that in sung female and male voices the identification of vowels becomes easier in the low tessiture. If male voices overpass G4 and female voices B5 the singing vowels and phrases are non- intelligible. The more we descend in low registers, the weaker the frequency difference is but when we are ascending towards higher frequencies, the difference between the frequencies may reach 500Hz.

According to the *fabrication imperatives* we must take care not only of the innumerable technical imperatives as for the construction of an ordinary synthesizer, but also of the control buttons which must be conceived in a different way than in an ordinary 'instrumental synthesizer'. In our case we must have the pertinent control environment in order to control not only the timbre of vowels (fundamental frequency, amplitude, formants and trajectories, frequency modulation, etc.) but also the combination of consonants and vowels into phrases as well as the adjunction of preset vocal diphones.

Following this logic, the ordinary keyboard which controls the pitches of synthetic instrumental sounds is insufficient because one of the most important factors for the simulation of a singing voice is 'artificial respiration'. Thus, we are proposing that the form of this synthesizer should have been designed more towards the form of an accordion (equipped with a soufflet.) because of two major reasons: first it is an instrument which like the church organ is functioning by air (free reeds) and after is closely carried by the body which can add more expressivity and liberty of mouvement to the interpretator.

The *anatomo- pysiological imperatives* can be overlooked in all synthesizers, as a simple movement can produce all the audible scales. The only advantage that is given by this new instrument- concerning the human anatomo- physiological imperatives- is the possibility to overcome problems that a human being cannot. For example in order to construct vocal sounds of 100Hz, with the proper intensity level which makes the sound perceptible (curves Fletcher Munson), the human lungs are insufficient. In this way a synthesizer gives us the possibility to construct voices extremely forte or piano by the pertinent equalization of the registers or by a simple gesture of control on the parameters (F0, formants positions, amplitude and bandwidth).

According to the *commercial imperatives* we could remark that unfortunately the commerce of electronic instruments often have nothing to do with the musical research; the compromise between researchers and big instrument manufacturers (Yamaha, Korg, Roland, etc.) must , if not avoided, at least be based on new terms (like the phase- vocoder of the '70s). Additionally, the fact that every language has its proper phonetic rules and phonemes, prevents the commercialization of a vocal synthesizer, for evident reasons.

Now, if a vocal synthesizer is being commercialized, first of all the problems of control (language, timbres, techniques) must be resolved in a convenient way and second the designer must preserve the sophistication of the instrument by addressing the people who have the knowledge to handle it in an artistic or scientific way (people who are working in the field of electroacoustic and computer music, institutions, composers and performers, phoneticians, singers etc.).

Another parameters which could affect the performability of the synthesizer is the liberty domains concerning the pitch, intensity, timber, form. First of all, according to the liberty domain concerning pitch we are expecting from the synthesizer to give the possibility to the user for pitch modulation, change of registers and easy transition from the low to the high register. We must also examine the case of polyphony (how many voices can be produced simultaneously in the same or different timbres?).

According to the liberty domain concerning the *intensity and the dynamics*, we are expecting a better control than the ordinary synthesizer in order to enhance to expressive quality of singing.

We have been confronted many times till now with disadvantages of expression in the most electronic instruments of the commerce. In addition, the control by sensitive touch or pedals is insufficient in the case of voice as we have already describe above.

According to the *liberty of choosing the timbres* by pertinent controllers or by selecting a pre- constructed timbre- diphone from the database, we can create instrumental sounds which approach the vocal sound by extrapolation, hybridation or interpolation.

Finally, according to the *liberty of forms,* the interpreter-performer is given the possibility of controlling with virtuosity three multi- directional parameters (timbre, intensity, pitches) in a convenient way (we must not forget the excellent example pf Clara Rockmore playing the Moog/ Theremin)[5] . An ergonomics study should help the instrument designers who are interested in the design of a 'cognitive' instrument. Finally, an interesting instrument for music is not necessarily a complicated instrument, which gives the possibility to interpret many notes (like the piano) but also an instrument which allows a number of effects and different musical forms.

Our last expectation of this vocal synthesizer concerns the field of ethnomusicilogical research : to which degree this synthesizer could be a pertinent tool for integrating several voice techniques, which could be studied and modeled with the aid of ethnomusicologists, phoneticians, and system engineers or study forgotten voice techniques of other cultures and times (e.g. Greek ancient singing)?

All tools are extensions of human intention. The musical instrument that are in essence an extension of human voice and touch, are conceived in order to combine the emotion of the speaking voice and of the singing voice, with the possibility of dexterity than can be achieved with the fingers and hands. Yet, the immediate musical expression is still the voice, the most difficult to master instrument. In order to overcome the difficulty of mastering the voice, we are dreaming of such an instrument, that we will allow everyone of us (without the gift of a resonant voice) to model with the hands our own vocal environment!

---

[5]We should not forget the excellent example of Clara Rockmore playing the Moog/Theremin.

5 CONCLUSION

Ameliorating the technical problems concerning the synthesis of the singing voice, researchers can in the future orientate their research in the construction of a vocal synthesizer which can be more accessible to composers, musicians and musicologists.

6 BIBLIOGRAPHIC REFERENCES

Berndtsson, Gunilla (1995) *Systems for synthesising singing and for enhancing the acoustics of music rooms*. Dissertation, KTH, Department of Speech communication and Music Acoustics, Royal Institut of Technology, Stockholm.

Cadoz Claude, Florens J-L.. (1989) Physical Models for music and animated image; The use of Cordis - Anima in *Esquisses*. A music film by ACROE. In *ICMC 1994* proceedings, Tokyo, p. 11.

Carlson ,G. and Ternström,S. and Sundberg, J. (1991) A new digital system for singing synthesis allowing expressive control.In *ICMC'91 proceedings,* Montréal,

Chowning John 1981) Computer Synthesis of Singing Voice. In *ICMC '81 proceedings*, La Trobe University, Melbourne.

Cook Perry (1993) Spasm, a real - time Vocal Tract Physical Model Controller; and Singer the companion Software Synthesis System. *Computer Music journal*, 17(1),MIT, Boston

Cook. Perry (1996) Singing voice synthesis: History, current work, and future directions. In C*omputer music journal*, vol. 20 no 3,. MIT Press. p.27

Depalle Philippe(1995) Conribution orale, Ircam .Paris.

Depalle Philippe, G. Garcia, X. Rodet (1994) A virtual castrato. I*n ICMC 1994 proceedings*, Aarhus, Denmark.

Dufourt Hugues(1996) L'instrument philosophe. Entretien avec p. Szendy, *Cahiers de l'Ircam, no7,* Paris.

Fant Gunnar (1973) Speech sounds and features. MIT press.Cambridge.

Gael Richard (1990) *Rules for fundamental frequency transition in singing synthesis*. Dept of Speech Communication and acoustics, Royal Institute of Technology, Stockholm.

Georgaki Anastasia (1998a) *Problémes techniques et enjeux esthétiques de lavoix de sythèse dans la recherche et création musicales.* Thèse de doctorat, EHESS/IRCAM, Paris.

Georgaki Anastasia (1998b)Synthesis of the singing voice: links betwen research and creation.In the proceedings of the *First symposium on music and computers*, Ionian University, Corfu.
Gibson, I.S., Howard, D.M., Tyrell, A.M.(1998) Real- time singing synthesis using a parallel processing system. *Proceedings of the IEE colloquium on Audio and music technology;the creative challenge of DSP*, IEEDigest 98/470, 8/1- 8/6.
Howell P. and N. Harvey (1975) Voice techniques. In *musical structure and Cognition,* .ed. P. Howell, I Cross and R. West. New York Academic Press.
Kamarotos D. , Diamantopoulos T., Philippis G. (1993) Igdis: A modern Greek text to specch/singing program for hte spasm / singer instrument. *In ICMC'93proceedings,* Tokyo..
Leipp Emile(1989) *Acoustique et Musique.* Masson , Paris.
Lomax Ken (1996) The development of a singing synthesizer. In *JIM'96 proceedings.*
Macon Michael (1997) A singing voice synthesis system based on sinusoidal modelling. *Proceedings of the International Conference on Acoustics, Speech and signal processing.*
Paabon Peter(1994) A real- time singing voice analysis/ synthesis system. *ICMC 1994* Proceedings. Aarhus , Denmark.
Pierucci P, Paladin A.(1997) Singing voice analysis ans synthesis system through glottal excited formant resonators in *ICMC '97 Proceedings,* Thessaloniki.
Puckette Miller **(**1991) Music and speech synthesis using nonlinear distortion and amplitude modulation", *Journal of audio society*.
Rodet X., Depalle Ph. (1992) *Spectral enveloppes and inverse FFT synthesis*. CNMAT, Paris.
Rossitier D., Howard D. (1994) Voice source and acoustic output qualities for singing synthesis. In *ICMC'94 Proceedings*. Aarhus , Denmark.
Risset J.C.(1991) Timbre analysis by synthesis: representations , imitations, and variants for musical composition. In *Musical signals and representations*, De poli et al.
Rodet X. et Al. (1984) The Chant project : From Synthesis of the singing voice to sythesis in general. *Computer Music Journal 8 (3)* (pp. 15- 31), MIT Press. _
Rodet X., Depalle Ph., Poirot G.(1988) Diphone Sound Synthesis", *Int. Computer Music Conference*, Koeln, RFA.
Sundberg Joahn (1981) To perceive one's own voice and another person's voice .In *research aspects of singing* (80- 84). Royal Swedish Academy of Music, Stockholm.

Sundberg Johan (1987) *The science of singing voice*, Northern Illinois University Press.

Sundberg John (1989) Synthesis of singing by rule. In *Current directions of computer music research*, ed. Max Mathews et John Pierce, MIT Press..

Tisato Graziano, Maccarini Andrea Ricci (1991) Analysis and synthesis of diphonic singing", *Bulletin d'Audiophonologie,* vol. VII no 5 et 6 - 619 - 648, Ann. Sc. Aniv. Franche - Comté.

Tubach J.P (1989) *La parole et son traitement automatique*. Calliope, Masson.Paris.

Zera Jan, Gauffin Jan, Sundberg Johan (1984) Synthesis for the selected vcv syllabes in singing. *In ICMC'84 proceedings*, Ircam , Paris .