.

Telepresence Learning Environments for Opera Singing, a Case Study

María Teresa Rojas Rajs

Thesis supervisor: Dr Francisco Alpiste Penalba

Dissertation Submitted in Partial Fulfillment of the
Requirements for the Degree of
Doctor in Multimedia Engineering

Doctoral Programme “Multimedia Engineering”
Department of Graphic Expression in Engineering

Universitat Politècnica de Catalunya UPC - Barcelona Tech

Barcelona, May 2012

A mi amado esposo Raúl y mi adorado hijo Antonio. Sois la luz de mi vida.

Acknowledgments

I would like to thank in the first place my thesis supervisor Dr Francesc Alpiste for his gentle but firm guidance through the research process. Thank you for your support and it has been a privilege to share with you such mixture of musical and technological matters.

I would also like to thank Dr Joaquin Fernandez and Dr Josep María Monguet, who have participated in the project and mentored some parts of the research process.

Opera eLearning has been possible thanks to the collaboration agreement between the City of Music of Sabadell City Council, the i2cat Foundation and the Multimedia Applications Laboratory. I would also like to thank Joachim Rockel from Dial-HD for lending us the Acoustic Echo Canceller device for testing purposes; Marc Milian for Figure 17 and the pictures in Appendix B; Jordi Torner for editing some of the statistical figures in this thesis; and my teacher and mezzo soprano Grissel Ruiz for her support in the translation of some opera singing technique terms. I also thank Paula Repetto for reviewing the German words that had been spelled from recordings. Special thanks to the opera singing professor Fernando Bañó for authorizing the reproduction of some of the figures in his book *La Antitécnica*.

The author of the thesis has completed her PhD studies at the Universitat Politècnica de Catalunya. BarcelonaTech (UPC) supported by grant 208288 from the CONACYT, the National Council for Science and Technology of Mexico.

I would like to thank my parents, Gustavo and Danuta, my sister *Caracol*, my cousin Lucía and my husband and child. Thank you for your support and patience, God bless you all.

Abstract

The present study analyzes the data obtained in the execution of the Opera eLearning project, a multidisciplinary effort to develop a solution for Opera singing distance lessons at the graduate level, using high bandwidth to deliver quality audio and video experience that has been evaluated by singing teachers, chorus and orchestra directors, singers and other professional musicians.

The research work includes the phases of design, execution and evaluation of pilot tests, followed by further development and execution of several experimental exercises with the system, all of them carried out between July 2008 and April 2009. This is an empirical research, an exploratory case study that has provided enough data to arrive to a sustainable model for a telepresence learning environment. Different usability methods have been implemented in order to assure users of the quality of the product. The main objective is to prove whether the system or artifact proposed can be used to deliver a complete remote singing class at a higher education level; for that purpose, we have defined several research categories that describe the usability of the system in multiple dimensions. We have used "design as research" approaches to promote innovation in the technological area.

The theoretical framework is based on a wide variety of fields; from acoustics, physics, music, professional singing to telecommunications and multimedia technology. However, the common thread and central issue under analysis is distance education, through the construction of a remote learning system. We have also included the corresponding justification of the scientific methodology employed.

This document is divided in the following chapters:

- Chapter 1: Introduction, objectives and methods.
- Chapter 2: Theoretical and technological framework.
- Chapter 3: Methodology and research design.
- Chapter 4: Empirical work.
- Chapter 5: Results.
- Chapter 6: Conclusions.

References are cited using APA 6th edition style (American Psychological Association, 2012) and UPC - Barcelona Tech format recommendations have been used on a general basis.

Keywords: telepresence; singing learning environments; high bandwidth videoconference; immersive sound experience; music higher education; Internet2; innovation; design as research; multidisciplinary; user-centered design.

Resumen

El presente estudio analiza los datos obtenidos en la ejecución del proyecto Opera eLearning, un esfuerzo multidisciplinario para desarrollar una solución que permita dar clases a distancia de canto lírico a nivel de educación superior, utilizando conexiones de banda ancha con el fin de proveer una experiencia de vídeo y audio de calidad, la que ha sido evaluada por profesores de canto, directores de coros y orquesta, cantantes y otros músicos profesionales.

El trabajo de investigación incluye las fases de diseño, ejecución y evaluación de las pruebas piloto, seguido del posterior desarrollo y ejecución de varios ejercicios experimentales con el sistema, todos ellos efectuados entre Julio de 2008 y Abril de 2009. Esta es una investigación empírica, un caso de estudio exploratorio que ha obtenido datos suficientes como para definir un modelo sostenible de entorno de enseñanza por telepresencia. Diversos métodos de usabilidad fueron implementados con el fin de asegurar a los usuarios la calidad del producto. El objetivo principal es probar si el sistema o artefacto propuesto puede ser usado para realizar de modo remoto una clase completa de canto lírico a nivel de educación superior; con tal propósito, hemos definido varias categorías de investigación que describen la usabilidad del sistema en múltiples dimensiones. Hemos utilizado el enfoque de "diseño como investigación" para promover la innovación en el área tecnológica.

El marco teórico se basa en una amplia variedad de campos; desde la acústica, la física, la música, el canto profesional hasta las telecomunicaciones y tecnología multimedia. Sin embargo, el hilo común y tema central bajo análisis es la educación a distancia, ya que se trata de la construcción de un sistema de aprendizaje remoto.

También se he incluido la justificación correspondiente a la metodología científica empleada.

Este documento está dividido en los siguientes capítulos:

- Capítulo 1: Introducción, objetivos y métodos.
- Capítulo 2: Marco teórico y tecnológico.
- Capítulo 3: Metodología y diseño de la investigación.
- Capítulo 4: Trabajo empírico.
- Capítulo 5: Resultados.
- Capítulo 6: Conclusiones.

Las referencias se han citado con el estilo APA en su sexta edición (American Psychological Association, 2012) y las recomendaciones de formato de la Universitat Politècnica de Catalunya han sido aplicadas de manera general.

Palabras clave: telepresencia; entornos de aprendizaje para canto; videoconferencia por banda ancha; experiencia de sonido inmersivo; educación superior musical; Internet2; innovación; diseño como investigación; multidisciplina; diseño centrado en el usuario.

Content

List of figures

List of tables

1. Introduction: goals and methods

Almost forty years have passed since the first instructive video applications saw the light in the early 1970s (James, 1970), (Lange, 1971). Soon the viability of spreading knowledge by this means was tried out by educational institutions and became a matter of investigation, surveying the use of this technology in higher education (Ellis & Curless, 1986), evaluating effectiveness and acceptance by learning style (Larsen, 1992), or discovering opportunities in specific fields like healthcare or continuous education (Chen, Eckhardt, Sinkowitz-Cochran, & Jarvis, 1999), among others. Later, technology provided the possibility of establishing a videoconference instead of producing an asynchronous, static recorded material.

Today's high bandwidth availability allows the transmission of large amounts of data, for instance with Opera and Orchestral Concerts transmission. (Gran Teatre del Liceu, 2010) (Manhattan School of Music, 2009b) The Opera eLearning project has been inspired by the possibility to deliver remote singing lessons at the graduate level that would accomplish the image and sound requirements of singing teachers and students in order to work remotely at a very specialized level.

Why should classes be delivered remotely? Some teachers are active musicians, who have opera seasons that leave their students alone for a part of the year. This is the first reason mentioned freely by students in open text fields of the questionnaires. On the other hand, teachers seek to put their students in touch with great masters who perhaps cannot travel to the country where they are, apart from facilitating the opportunity for students who cannot afford to travel themselves.

Opera eLearning has been promoted by the City of the Music (Sabadell City Council, 2011) of Sabadell City Council[1], which is expecting to deliver an Opera Singing Master Degree with this telepresence environment as part of their future City of the Music installations and activities. In the technology side, the communications area has been provided by the i2cat Foundation[2] and the video and audio by the Multimedia Applications Laboratory[3] of the Barcelona Tech. These three institutions has allocated time and resources for collaborate at several prototypes developed in specific dates; a large amount of applied technological research has been required between different prototypes, and continuous feedback from final users has been the core of the changes and adjustments for the artifact.

A fine matter in such cases is always to discern between the project itself and the scientific case study activities; such difficult work has been the job of the author of these lines, who has designed tests, registered the results at each step, coordinated that all the questionnaires and other testing materials gets to the right people, etc. And previously, has studied the field to propose a theoretical framework, construct the best of breed on technology possible, as well as understand the new generation in advanced videoconference with high band width: the telepresence.

Locating other experience of distance higher education on music has been difficult, but we have plenty of reference for graduate studies. It is interesting to learn that this project then has been pioneer from the education and technological side as well.

[1] City of the Music of Sabadell City Concil may be reached at http://www.sabadell.es/Ciudaddelamusica/p/ciutatmusica_esp.asp
[2] I2cat Foundation may be reached at www.i2cat.net/en
[3] LAM may be reached at http://lamupc.com/

1.1. Objectives

Opera eLearning project goal was designing, constructing and testing a remote learning environment for opera singing at the graduate level.

Opera eLearning study goal was evaluating the tested artifact against a set of research categories and hypothesis in a user oriented model according to usability standards.

1.2. Hypothesis

The hypotheses were user oriented: the model of remote singing learning room:

(1) Was enough for the teacher to fully evaluate the student

(2) Allowed the student to understand observations and emulate the teacher's sound

(3) Observers could use the information received as well

(4) A high quality, immersive and space distributed sound could improve teacher's listening.

The main question to be answered, at last, is prove if it was possible to deliver proper masterclass remotely with this artifact.

1.3. Methods

Opera eLearning case study was applied technological research, with an iterative prototype's process very characteristic on innovation exercises. Main methodology is case study, because an empiric research on technology is well constructed in a case study as a reliable assessment of the product (Kitchenham, 2007). In this case we had an exploratory case study (Yin, 2003), where previous hypothesis are well defined but we were committed to find additional information trough data recovered. It was a common work in engineering, because we were testing an

artifact. In this sense, we could also define the study as a validation or verification study, from the usability point of view (Rubin, 1994).

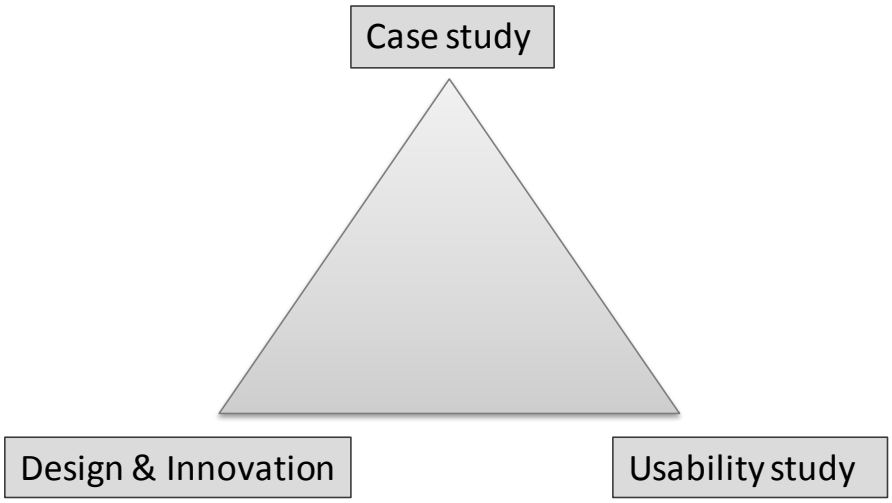

Figure 1 Methodological framework overview

As we walked through different prototypes, design activity became a crucial part of the project. Therefore we arrived to the "design as research" approach (Purao et al., 2008) , and applied as well several best practices of innovation (European Commission, 2004) as part of the methods.

Some specific facts of the methods used are listed bellow, and each of the vertices of this methodological framework is detailed in the methods chapter.

I. The study had two stages: the pilot phase and the case study activities.

II. Each stages included several prototypes, and each prototype had a user-centered design (N. Bevan, 2001) in an iterative cycle.

III. Each prototype received different testing feedback: questionnaires with open text fields as well as Likert-Scales, expert's comments and review of the recorded material to analyze the results (both video and audio).

IV. A redesign, with full documentation putted in place before the execution of the next prototype.

V. All the process, from the design of the first prototype, has been also supported by the search of bibliography in a systematic literature review (Kitchenham et al., 2009) from different fields. At the end of the pilot stage a corpus of grey and formal literature had nourished the theoretical and technological framework, properly collected and integrated, and had keep growing ever since.

VI. Case study used the stable versions of the telepresence environment artifact, as well as the tested instruments for measuring user's experience.

1.4. Contributions

Our contributions to the field of music learning, specifically using this kind of telepresence environments:

1- The design of the artifact, although it should be improved in time applying technology advances. But at the time of the study, there were not many others experiences of this kind, so is a valuable result.

2- The immersive sound model. One of best results as we discuss later is have proven that the telepresence environment is highly enhanced by the use of an immersive sound model.

1.5. Limits of the study

The study does not intend to reach further developments, like intelligent agents or specific hardware developments. We used already existing devices, but organized a specific arrangement in rooms for create the telepresence environment. We may, of course, based in our experience propose possible research lines, but such lines are beyond the scope for this study.

From the theory point of view the study does not looks for proposing new paradigms, nevertheless we found very good practices and manner of doing stuff that may be useful as know-how for other projects, specially for multidisciplinary efforts.

The aim and contribution of the study is to provide a telepresence environment and check that it has been really useful for remote masterclasses in opera singing. Those are the specific limits.

2. Theoretical and technological framework

A large issue in multidisciplinary studies is to cover enough of each of the different fields involved. This study is constructed around a technological artifact, and therefore there is a background corpus strictly related to technical, video and audio as well as web and communication engineering. But we also considered music education and vocal technical for the framework, and the distance education is the main conductor since the aim of the project is to allow a teaching and learning process.

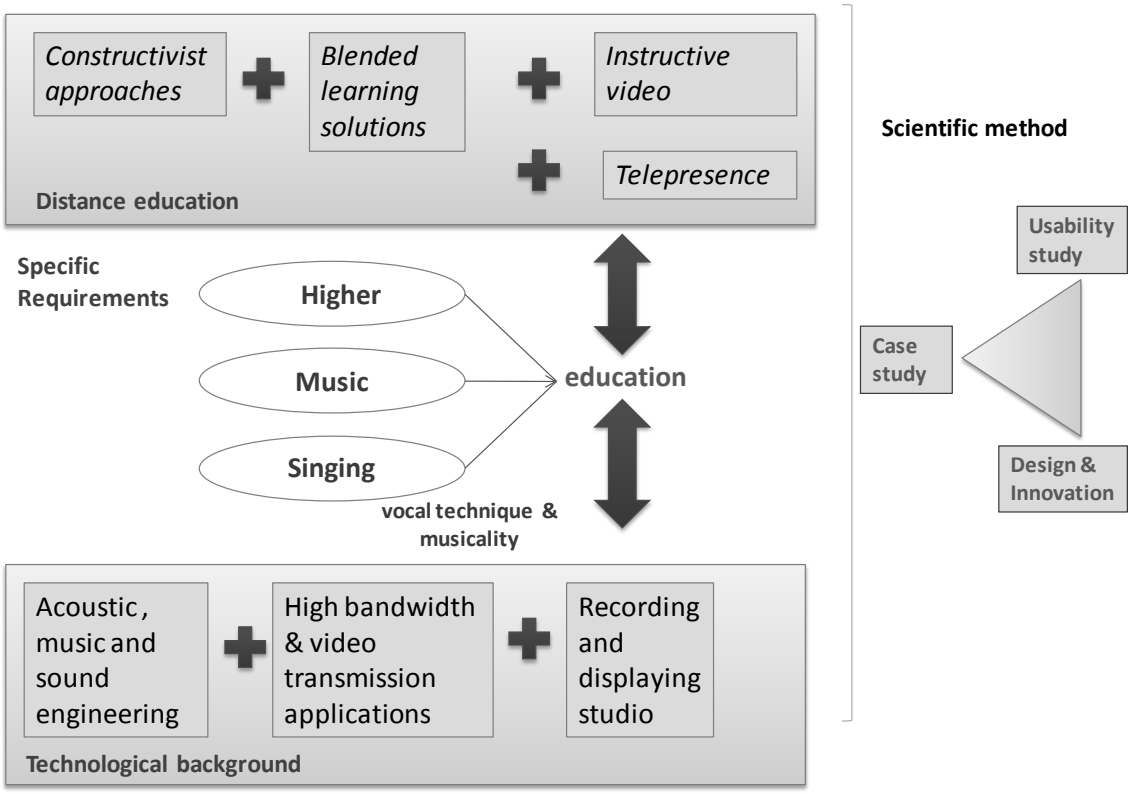

Figure 2. Theoretical & Technological Framework

This figure shows the components of the theoretical and technological framework. Matters in distance education have been split in different implementation methods and experiences, all supported by the constructivist approach to education.

Then the specific requirements of higher music singing education are treated in next place, and finally we will review the technical details.

2.1. Distance education

Opera eLearning is a remote learning environment, therefore all the theoretical argument is developed under the umbrella of distance education. It should be remarked that the theory developed for the Opera eLearning project includes all the construction of the remote, blended learning facilities, including the Web page and other services, nevertheless the innovation was focused on the telepresence environment and the case study has explored only that part of the model constructed.

In the next sections I will develop first the historic and theory for distance education, then propose some particular approaches for instructional design until comment specific practices and activities, trying to keep the thread with distance education theory all along.

2.1.1. History and definition

Distance education has a long history, starting with correspondence education in the late 1800s at the University of Chicago. Before that time, particularly in preindustrial Europe, education had been available primarily to males in higher levels of society (Gunawardena & McIsaac, 2004). At that time students used to be hosted in organized institutions to address teacher´s lessons, and change temporally their residence for that purpose. Correspondence education was intended to be an opportunity for those who

cannot afford such cost, and has been perceived as an inferior education, which is a stigma that distance education has to fight even today. Technology advances have supported distance education from the beginning, reflecting the need of democratization of the access to education; a good example is the how early radio and television were used in schools to deliver instruction at a distance, like Wisconsin's School of the Air in the 1920s (Gunawardena & McIsaac, 2004). In time video-instructive and web technologies had taken advantage of the ability to bring teacher's lesson to the student, as we may review in next sections.

But, what is distance education? Several authors had provided definitions for such question. It is a field that receives contributions from pedagogy as well as from psychology. According to the discussion developed by (Gunawardena & McIsaac, 2004), let's distinguish the following definition as the most minimal and complete, in three key concepts related to the education and learning process:

- noncontiguous communication
- two-way interactive communication
- and the use of technology to mediate the necessary two-way communication (Garrison & Shale, 1987)

Many authors include the active participation of the educational organization or group, the education as an industrialized product or the ability to be available anytime, anywhere; but for some distance education applications perhaps one of such characteristics may not be accomplished (for instance an online video conference may happen at some specific time and date, and not be able to be published afterwards as compiled video because of author rights).

2.1.1.1. A particular case: the e-learning

E-learning is the particular case of distance learning where technology used for mediate the communication is web, teleconference and all the modern technologies that may be named as "electronic learning". We may trace its history at the 1950s (Ravenscroft, 2001). An interesting change of focus may be noticed in the e-learning name itself: the active role of the student, or the "learner". But let us have a rough view on the different theories that has supported e-learning approaches in the last 50 years, so we may argue the choice of some of them for Opera eLearning framework. The decades mentioned related to the authors are not precise (most authors can be publishing during 40 years), but give an idea of the kind of thought reflected on each of those approaches.

The *e-learning* main educational theories (according to (Ravenscroft, 2001))		
Theory name	Key concepts or statements	Authors and decades
Behaviorism	Learning as a shaping behavior, small chunks of information presented with questions and immediate feedback. Try to perceive education as an industrial product.	Skinner (50´s)
Cognitive style, holistic and Cybernetic	Distinguish different learning styles in a learner-centered approach, allows the student to choose its own path through a specific topic. Conversational guidance needed to be introduced to ensure an appropriate mapping between learning styles and teaching strategy. Such conversation may be between a human and an intelligent system.	Pask (70's)
Cognitive Constructivism	Humans generate knowledge and meaning from an interaction between their experiences and their ideas, in	Piaget (60's),

	a "learning by discovery" process. Generate individual meaning in an e-learning environment expressing learner's ideas in a set of instructions to the computer, and testing the results. Lack of tutoring involved leads to student to define their own goals.	Papert (80's)
Social Constructivism	Learning is a high level mental process, which means involve language and social interaction. The cooperative interaction with a more learned one gives the information in its full social context and meaning.	Vygotsky (60's)
Activity learning	Meaning arises and evolves during interactions that are influenced by the social relations within a community of practice. Activity is the minimal meaningful context for individual actions. "You are what you do".	Vygotsky (60's)

Table 1. The *e-learning* main educational theories (Ravenscroft, 2001)

From the table above, social constructivism and activity learning are the most suitable theories for supporting Opera eLearning case study. As will be discussed in the following lines, the paper of the most learned one or tutor in the system is central, as well as the exercise of the activity of singing at a professional level and context. And certainly other authors have developed and completed such approaches, adding further terms and experiences.

2.1.2. The constructivist approach

We will not go deeper in the philosophical, detailed differences that as contemporary authors Piaget and Vygotsky had, since the pieces that we would pick up

from each of their works are perfectly coherent: accepting the cognitive approach of generating knowledge interacting in experiences for trying out and construct learner´s ideas, we will emphasize the social constructivism including in such interaction the participation of the professor that conducts the masterclasses.

"*Vygotsky draws a clear distinction between lower level mental processes, such as elementary perception and attention, and higher level mental processes that include verbal thought, logical memory, selective attention and reasoning*" (…), where "*the higher level is mediated through cultural symbols and tools; self regulated rather than bound to a stimulus context; social in origin and the result of conscious awareness rather than an automatic response*" (Ravenscroft, 2001, p.141). This higher level thinking is characterized by the language, wich is interesting for our study, because besides the spoken language there is a musical language involved in the learning process of our research.

Another crucial contribution from Vygotsky is the activity learning that gives the precedent for the next two parts of our theoretical framework: the learning objects and the learning by doing. Both inheriths the community of practice and the importance of the learner activity (later developed and extended into problem solving (D. H. Jonassen, Howland, Moore, & Marra, 2003)). "*An activity is considered the minimal meaningful context for individual actions (...) These activities contain various artifacts, such as signs, methods, machines and computers, that serve as mediational tools to facilitate the operationalisation of conceptions in was that lead to 'higher levels of thinking'* ". (Ravenscroft, 2001, p.148).

Constructivism has been explored as an approach for technology supported learning environments. An important author is Jonassen, who develops a wide job from

the theoretical as well as practical matters. From the theory we will remark a fine argumentation on the constructivism phylosophical base:

"*The revolution in learning theory and instructional design has transcended the behaviorism-cognitivism dialectic and entered a new era of theorizing. The symbolic reasoning paradigm does not accommodate the dynamic nature of learning, emergent properties of thinking, plausible rather than exact reasoning, learning situated in context, and the indeterminism that always seems to subjugate our expectations ahout learning outcomes. These issues are philosophical as well as psychological. On a philosophical level, the symbolic reasoning-situated learning dialectic is discussed in terms of objectivist and constructivist epistemologies (Jonassen 1991; Duffy and Jonassen 1992). The dominant and traditional objectivist paradigm (which provides the foundation for symbolic reasoning) assumes that the world is structured, that structure can be modeled and mapped onto the learner, and that the goal of the learner is to "mirror" reality as interpreted by the instructor. Knowledge is external to the knower and so can be transferred (communicated) from one person to another. The learner´s role is to remember and reproduce the knowledge that is transmitted by the teacher or professor." (...) "Constructivism (which provides the psychological/philosophical foundation for situated learning) begins with a different set of assumptions about learning. Constructivists believe that our personal world is constructed in our minds and that these personal constructions define our personal realities. The mind is the instrument of thinking which interprets events, objects, and perspectives rather than seeking to remember and comprehend an objective knowledge. The mind filters input from the world in the process of making those interpretations.*

The important epistemological assumption of constructivism is that knowledge is a function of how the individual creates meaning from his or her experiences; it is not

a function of what someone else says is true. *Each of us conceives of external reality somewhat differently, based upon our unique set of experiences with the world and our beliefs about them."*(D. Jonassen, Davidson, Collins, Campbell, & Bannan Haag, 1995, p.4)

Now we arrive a to a new concept: *situated learning*. This is a development of the constructivist approach theory that is in contraposition to the symbolic reasoning paradigm. Next table shows the differences in both approaches.

Symbolic Reasoning		**Situated Learning**
	Knowledge	
objective		subjective
independent		contextualized
stable		relative
applied		situated in action
fixed		fluid
	Learning	
objectivist		constructivist
product-oriented		process-oriented
abstract		authentic
symbolic		experiential
	Memory	
stored representations		connections, potentials
	Knowledge Representation	
functionally equivalent to real world		embedded in experience
replication of expert		personally constructed
symbolic, generalized		personalized
	Instruction	
top down		bottom up
deductive		inductive
application of symbols		apprenticeship
	Computational Model	
symbolic reasoning		connectionist
production rule		neural network
symbolic manipulation		probabilistic, embedded

Table 2. Contrasting Assumptions of Paradigms (David Jonassen et al., 1995, p.3)

In a more applied level, Jonassen has described how to apply the constructivist approach to computer based learning environments; there are two famous images that

resumes his capability of sythesis for this *know-how* proposal. He suggests that any technology must allow students to be engaged in meaningful learning. He refers to the qualities that must be used as guidelines when designing Constructivist Learning Environments. The following diagram identifies these qualities and how they are interconnected.

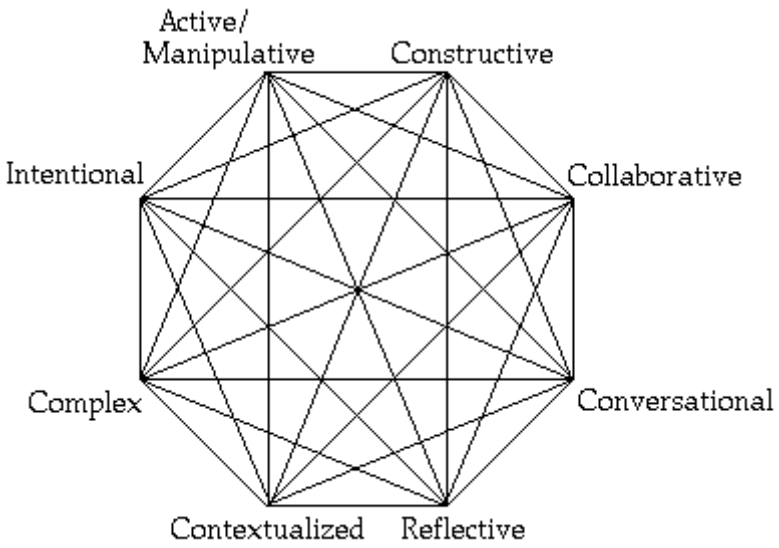

Figure 3. Qualities interconection for CLE. (D. Jonassen, 2002, p.17)

Second one propose a model for designing constructivis learning environments, where we find tools required for this learning approach in layers. In terms of the Opera eLearning project, we can find the problem or project in the execution of the masterclass, the related cases in the videos available online from other student's sessions, we do not have automatized exercises included that may reach the congnitive development, but yes we have conversational tools in the web forums, and the masterclass has a social meaning, and can be commented later for other peers.

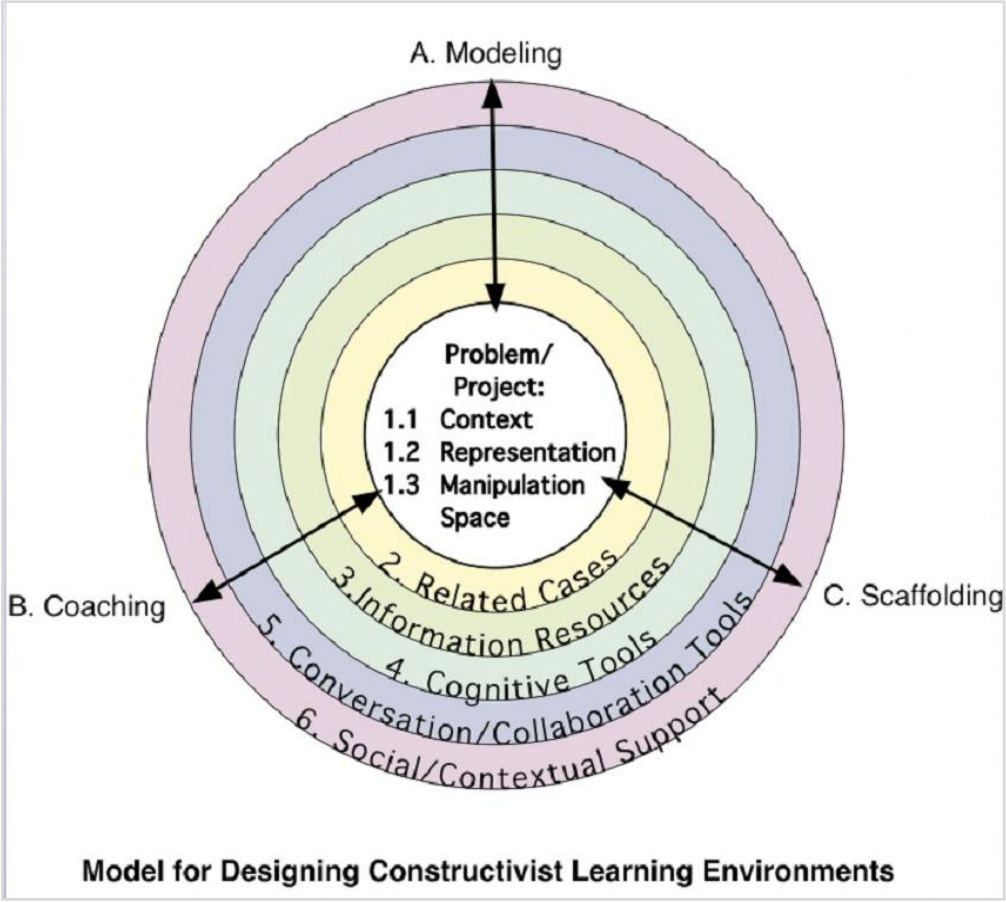

Figure 4. Model for Designing CLE. (D. H. Jonassen, 1999, p.12)

Constructivism is accepted nowadays as the paradigm for supporting distance education for most authors; there is a huge amount of scholar literature related and this study only includes a systematic literature review applied to the particular system proposed for the Opera eLearning research.

2.1.2.1. Our learning objects

Learning objects come from the idea of object oriented programming (Budd, 2002; Pierce, 2002), a software development paradigm that allows reusing pieces of code and linking them into new products at different contexts in order to minimize

efforts and use proven solutions, as well as get immediate benefits from new versions. With the formal framework from the founders of UML (Booch, Jacobson, Rumbaugh, & Safari Tech Books Online, 1999) objects have an interesting, powerful behavior by heritage, modularity and the hot deployment of changes at once in several sites by the update of a single object, which has being used with languages like Java, C++, LISP, etc. It is easily maintainable, and optimizes work.

Since most of the distance learning that is being discussed today has to do with the huge change in human communications introduced by the World Wide Web, is not rare that we have a proposal from the technological area to solve the instructional matters. As we shall discuss in this chapter, this not necessarily links with the education theory, but is an interesting deal that has a lot of sense. Basically the idea is that instructional models could break their information, activities, etc. in these reusable chunks, which are understood to be digital (no one is discussing to design a printed book as learning object, for instance).

Formally speaking, learning objects are being developed and ruled as an instructional technology standard by the Learning Technology Standards Committee (LTSC, 2010) from the Institute of Electrical and Electronics Engineers IEEE. The define the learning objects as: "*Learning Objects are defined here as any entity, digital or non-digital, which can be used, re-used or referenced during technology supported learning. Examples of technology supported learning include computer-based training systems, interactive learning environments, intelligent computer-aided instruction systems, distance learning systems, and collaborative learning environments. Examples of Learning Objects include multimedia content, instructional content, learning objectives, instructional software and software tools, and persons, organizations, or*

events referenced during technology supported learning." (Learning Technology Standards Committee, 2010)

They have been working since 1996, involving projects and publishing, even developing open standards like Learning Objects Metadata (Learning Object Metadata, 2011) , that would provide a common framework for universities, educators and entities interested in use this instructional technology for construct compatible objects.

However, several enterprises have developed solutions for educations and named part of them as learning objects without necessarily respect the LTSC orientation (Wiley, 2000), and even some of the authors that are writing about it have some differences. A more wide definition is provided by Wiley as *"any digital resource that can be reused to support learning."* (Wiley, 2000, p.7) This definition rejects any non-digital approach, and I agree with that for the learning environments that are being constructed today. Wiley also contributes with some sharp observations: "*While groups like the Learning Technology Standards Committee exist to promote international discussion around the technology standards necessary to support learning object-based instruction, and many people are talking about the financial opportunities about to come into existence, there is astonishingly little conversation around the instructional design implications of learning objects*." (Wiley, 2000, p.9) If we do not agree on instructional design, then is very unlikely that learning objects from different environments may be combined; another issue is the granularity, that is, how big a learning object could be. And, at last, "*the major issues facing would-be employers of learning objects, granularity and combination, turn out to be perhaps the two considerations known best to instructional designers: scope and sequence. There is a number of existing instructional design theories that provide explicit scope and sequencing support that, while not intended to be, are applicable to learning objects.*

Reigeluth's Elaboration Theory (Reigeluth, 1999b), van Merriënboer's Four-component Instructional Design model (van Merriënboer, 1997), and Gibbons and his colleagues' Work Model Synthesis approach (Gibbons et al., 1995) come to mind, among others." (Wiley, 2000, p.13)

Finally I will reproduce an accurate taxonomy developed by Wiley as:

- "*Fundamental - For example, a JPEG of a hand playing a chord on a piano keyboard.*
- *Combined-closed - For example, a video of a hand playing an arpeggiated chord on a piano keyboard with accompanying audio.*
- *Combined-open - For example, a web page dynamically combining the previously mentioned JPEG and QuickTime file together with textual material "on the fly."*
- *Generative-presentation - For example, a JAVA applet capable of graphically generating a set of staff, clef, and notes, and then positioning them appropriately to present a chord identification problem to a student.*
- *Generative-instructional - For example, an EXECUTE instructional transaction shell (Merrill, 1999), which both instructs and provides practice for any type of procedure, for example, the process of chord root, quality, and inversion identification*." (Wiley, 2000, p.22)

As previously mentioned, link learning objects with learning theories is crucial for the proper design of education processes and contents to be delivered. A fine resume of constructivist learning theories that may be mapped into learning objects can be found at (Bannan-Ritland, Dabbagh, & Murphy, 2000). Given that "*Constructivism is*

an educational philosophy or perspective that encompasses a wide variety of views, theories and instructional models. These views seem to converge on at least two principles according to Duffy & Cunningham (1996),

(1) That learning is an active process of constructing rather than acquiring knowledge and

(2) Instruction is a process of supporting that construction rather than communicating knowledge. " (Bannan-Ritland et al., 2000, p.12), they propose the following table as description of possibilities.

Learning theory	Situated cognition; Distributed cognition	Cognitive Flexibility Theory	Social interaction; Action learning	Social interaction; Activity theory; Distributed cognition	Generative Learning theory	Inquiry theory; Activity theory
Instructional Model	Cognitive Apprenticeship; Situated Learning	Random Access Instruction	Problem-based learning	Distributed expertise; Knowledge management	Generative teaching model	Situated Learning; Experiential learning
Instructional Strategy	Coaching, Authentic Activity, Modeling, Articulation, Exploration, Scaffolding	Cased-based Learning, thematic-based learning, self-directed learning	Collaboration, guided inquiry, authentic activity, small group instruction, self-directed learning	Collaborative learning; Learner-centered instruction; Goal-based instruction	Organization Conceptualization Integration Translation	Self-directed learning; Collaborative learning; Authentic activity; Exploration
Instructional Application	Apprenticeship; Internship; Story-based instruction; Situated narration	Cognitive Flexibility Hypertext	Problem-centered instruction	Virtual learning communities, Communities of practice	CSILEs, Anchored Instruction	Microworlds, Simulations

Table 3. Contrasting learning theories, instructional models, strategies and applications (Bannan-Ritland et al., 2000, p.24)

Depending on the kind of application to be built, some of these learning theories and related instructional model and strategy may be the most suitable. In the case of the Opera eLearning project, the social interaction, the activity theory and the problem-based learning are the most suitable choices. Their instructional applications generate the learning objects constructed for this project.

The particular opinion of this thesis's author is that the learning objects applied to telepresence virtual environments inherits the conception of the activity learning proposed by Vygotsky: in the interaction between the learner and the activity there are several artifacts, methods and computers that are mediational tools for experience and testing the ideas in order to generate knowledge. Well, those may also be called learning objects, and in this case are the communication human-to-human artifact and the videos and other media available. This idea needs further work from the epistemological point of view, and could provide a deep study in the theory field.

Finally, we would distinguish Opera eLearning objects, most of the time of the kind combined-open, as:

- The telepresence learning environment as a complex artifact, but that may be reusable for another kind of remote conference.
- The Web media elements that includes PDF files with scores, videos and audios and other reference material.
- The forums and other socials tools.

It must be said that in those learning objects are not LOM compliant.

2.1.2.2. The telepresence masterclass as a concert:" learning by doing"

So far we have established a learning theory and mapped the instructional strategy into the learning objects to construct the artifact of the telepresence environment. Now we shall review in detail the role and importance of the student's activity in such environment.

Merrill define some principles for instructional design after analyzing different instructional methods (Merril, 2002), most of them recompiled in a very famous and cited book from instructional design theories and models (Reigeluth, 1999). In this book appears a chapter titled "Learning by doing" (R. C. Schank, Berman, & Macperson, 1999), where Schank explains the instructional approach developed searching applications for industrial areas with virtual learning environments (R. Schank, 1997). Let´s review briefly the relation between those two focuses and how they describe the activity of the student at Opera eLearning case study.

The principles for instructional design from Merrill are:

"*Principle 1—Problem-centered: Learning is promoted when learners are engaged in solving real-world problems." (...)*

"Principle 2—Activation: Learning is promoted when relevant previous experience is activated." (...)

"Principle 3—Demonstration (Show me): Learning is promoted when the instruction demonstrates what is to be learned rather than merely telling information about what is to be learned." (...)

"Principle 4—Application (Let me): Learning is promoted when learners are required to use their new knowledge or skill to solve problems."

"*Principle 5—Integration: Learning is promoted when learners are encouraged to integrate (transfer) the new knowledge or skill into their everyday life*" (Merril, 2002, p.45)

Here we can identify all the steps in the one-to-one masterclass delivered through Opera eLearning telepresence environment; as told, the student should prepare and solve the masterclass session as a concert in terms of preparation, relaxation, concentration and scenic presence. It is solving a real-world problem, where previous concerts and presentations will be used to apply his or her experience to the present experience. As well, when the teacher interrupts to correct the student we have the demonstration and application principle into practice. Teachers can be singers themselves and then sing along to demonstrate how to accomplish the phrase, or they can be experienced accompanist that shows the phrase at the piano as well as explain in words that any advanced student may understand the changes to be applied. Then the student has to repeat until getting the expected result. Usually when interpreters gets a piece worked out, they integrate the changes on further interpretations.

"*Learning by doing model was clearly problem-centered with a very strong emphasis on the application phase of instruction. In this model there was limited emphasis on activation and demonstration and, while integration was certainly the goal, there was very little in the model to direct the integration process per se*" (Merril, 2002, p.56). Schank develop virtual scenarios for their students: "*GBS [goal-based scenario] is a learn-by-doing simulation in which students pursue a goal by practicing target skills and using relevant content knowledge to help them achieve their goal*" (…)

" There are seven essential components of a GBS: the learning goals, the mission, the cover story, the role, the scenario operations, the resources, and the feedback, including coaches and experts" (R. C. Schank et al., 1999). In this model Schank remarks the activity of the student as the main process for knowledge generation, and most important for this thesis, formalize the "learning by doing" instructive design.

It´s important to stress that for Merril, Reigeluth and other authors the "constructivist learning environments" from Jonassen and the "learning by doing" from Schank are different instructional design theories. We are using both of them for the design of Opera eLearning project, where the learning by doing is more applied for the telepresence masterclass itself, and the constructivist learning environment includes the Web Site and further work from the student that actually has not been measured yet with empiric work.

2.1.3. Blended learning

Blended learning is a concept that mixes the presence and remote tutoring in a learning process. It uses to have part of the contents and activities online, but includes the participation of the student and tutor in some presence sessions. "*Blended learning is the latest step in a long history of technology-based training*" (Bersin, 2004, p.2). A quick view on technology based training is well documented in the next figure:

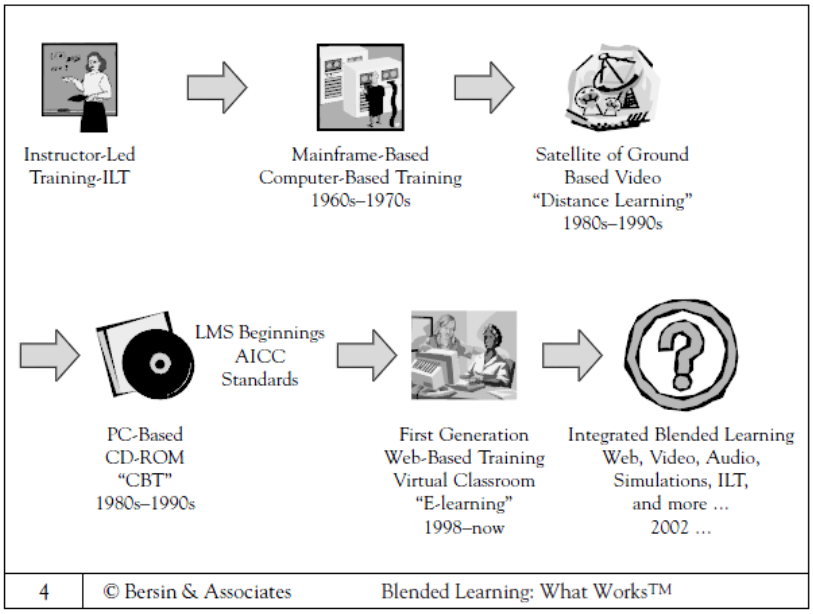

Figure 5. Evolution of Technology-Based Training (Bersin, 2004, p.2)

The benefit of integrate mainframe or computer-aided, satellite and video or CD-ROM with media content technologies to the learning process is well documented: increase the scale of possible students, provide a support for long-life learners, allows the interaction with multiple representations of the knowledge, give access in remote places and at any hour, and with proper course designs obtain feedback and evaluation. However, in time the participation of the tutor or professor in close contact with the student, as well as the contribution of the community to the process has been proved as required in several studies and fields (Bonk & Graham, 2006), (Abrahmov & Ronen, 2008). And actually these two columns, the role of the more learned one and the community practice are indeed what we are proposing from the Activity Theory and Social Constructivism of Vygotsky.

The lack of the contact with the professor and the learning community are the main issues that leaded to a limited scope some excellent ideas, like the CD-ROM

media productions or the TV transmissions. The raise of the Internet has allowed to facilitate more contact one-to-one and in community. (Bersin, 2004)

Another term to speak about blended learning environments is "hybrid model", that includes all the following characteristics that we pursuit as part of the whole Opera eLearning design (Alpiste, 2008):

- A virtual space, which includes the virtual classroom as well as the media library.
- A contents administrator that allows the integration and production of new contents as well as the knowledge management.
- The tutoring tools that in this case are not only the forums and mails, but the telepresence environment.
- The evaluation and activities of the student.

Next figure shows the whole design for the Opera eLearning project, but it is important to stress that only a part of this have been constructed, and only the telepresence environment for the remote masterclass sessions have been tested at the empiric work. Web Site has been constructed but has not been used by the learning community or tested as part of the study.

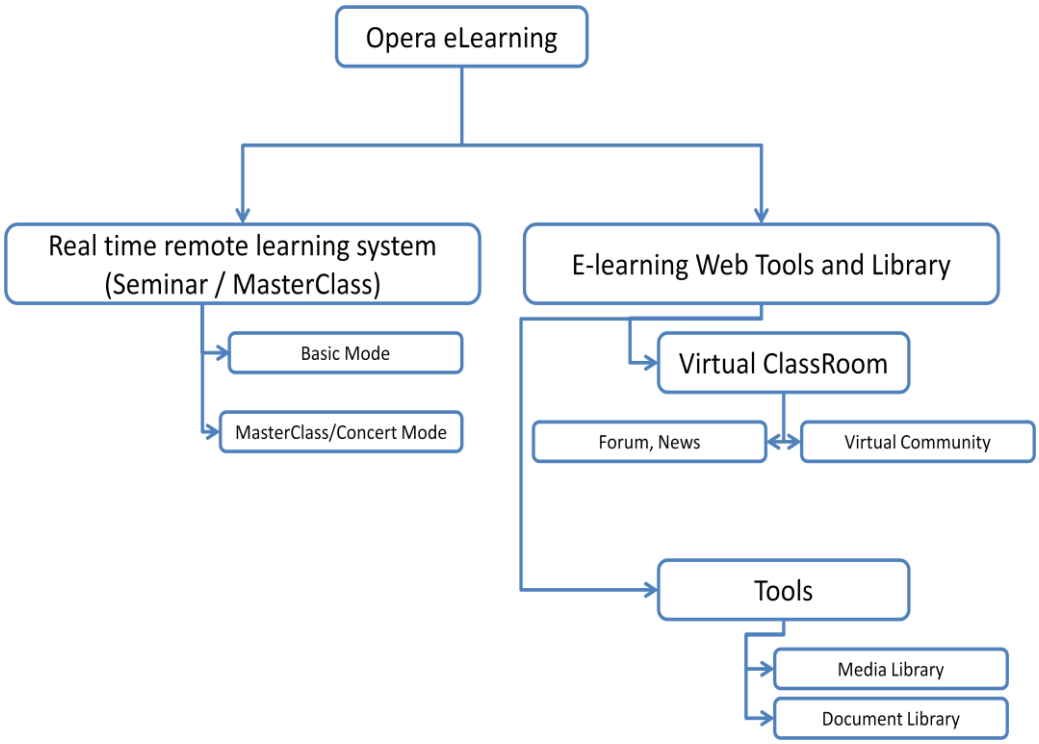

Figure 6. Whole design for the Opera eLearning project.

Those remote seminar and masterclasses session through the telepresence artifact should be anyhow complemented with presence sessions for a complete blended learning approach.

2.1.3.1. Remote resources and daily practice

Regarding music education there are interesting examples in blended learning solutions, mixing presence tutoring with the availability of contents, community support and automated review of exercises, like VEMUS (Virtual European Music School, 2009). This and other projects had advanced on supporting the execution and evaluation of the student in automated manners, like controlling gestures, measuring pitch and tone, involving sensors, visual aids and computer-assisted feedback. Such researches are crucial for the future of music education, and are maturing to the stage of even deliver

standards for further applications, like the Symbolic Music Representation (i-maestro, 2009). From the musical point of view, the instruments selected are winds, strings, percussions or piano.

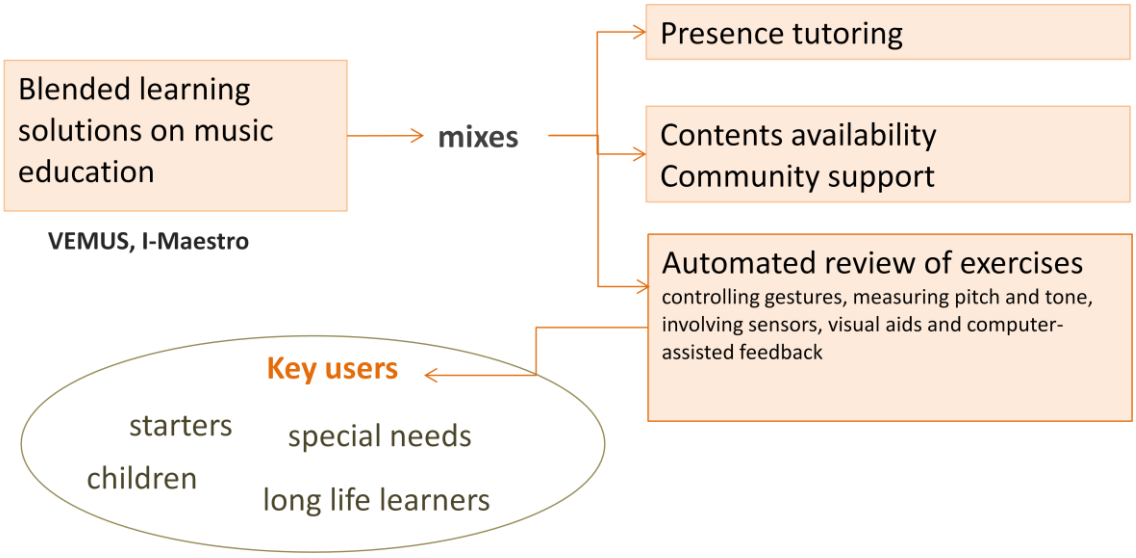

Figure 7. Blended learning solutions on music education.

Nevertheless this study has not included any remote practice tool, and although for higher education this may not be the main benefit expected, it was important to mention this possibility for the integration of blended learning at music education.

2.1.4. Instructive video

Almost forty years have passed since the first instructive video applications saw the light in the early 1970s (James, 1970), (Lange, 1971). Soon the viability of spreading knowledge by this means was tried out by educational institutions and became a matter of investigation, surveying the use of this technology in higher education (Ellis & Curless, 1986), evaluating effectiveness and acceptance by learning style (Larsen, 1992), or discovering opportunities in specific fields like healthcare or continuous education (Chen et al., 1999), among others. Later, technology provided the possibility

of establishing a videoconference instead of producing an asynchronous, static recorded material.

Today's high bandwidth availability allows the transmission of large amounts of data, for instance with Opera and Orchestral Concerts transmission (Gran Teatre del Liceu, 2010) (Manhattan School of Music, 2009).

In the Opera eLearning Web Site we have the recorded videos of the masterclasses published available for the students to take advantage and learn more from teacher's observations. We have followed in their edition, publication and format compilation some of the most important instructional strategies mentioned by (Anderson & Ellis, 2005), specifically for music education, that are listed below:

- "*The ability to easily and precisely pause, forward, and rewind video files helps students and teachers to analyze their technique and make adjustments where necessary.*
- *Split screen effects can be useful for showing: (a) two performers playing a duet; and (b) the same technique simultaneously from different angles.*
- *Nonverbal communication cues and gestures such as smiling and waving can be important and has implications for the positioning of cameras to allow occasional headshots during desktop videoconferencing or the prerecording of instructional videos for streaming via the web.*
- *Web-based instructional video files gave students a useful point of reference for their interlesson practice.*
- *Web-based instructional video files allow the teacher to readily update or refine web-based videos and supporting material and provide access to large volumes of files via searchable online databases.*

- *Avoid using extraneous material such as superfluous special effects and video titles that usually look washed out and unfocused on the web or too small to read clearly in smallvideos (eg, 160 pixels x 120 pixels or less). Instead, post supporting text on the web page or slide supporting the video.*
- *World Wide Web Consortium and government content accessibility guidelines should be followed when producing and distributing materials in the context of desktop videoassistef music teaching via the web.*
- *The ease with which digital media can be manipulated means copyright must be respected at all levels: instructional designer, teacher, and student*."(Anderson & Ellis, 2005)

2.1.5. Telepresence in education

First indexed reference specifically entitled as Telepresence in scientific literature leads to 1998 (Draper, Kaber, & Usher, 1998), where we find a fine taxonomy and review on the work to that date of the telepresence concept, that starts in the 70's related to the remote teleoperation of artifacts, for instance in space engineering. They make differences from telepresence, which is the presence of one person projected into a physical remote site, from "virtual presence", where the remote space does not exist in real world but is generated by a software; and also should be differentiated from the "ego presence" which is a projection of the presence in a remote representation, for instance in a forum, video game or desktop videoconference. (Draper et al., 1998)

As formal definitions, is common the use of three types:

- The simple definition, where "*telepresence refers to the ability to operate in a computer-mediated environment*",

- The cybernetic definition, where "*telepresence is an index of the quality of the human-machine interface*" and
- The experiential definition, where "*telepresence is a mental state in which a user feels physically present within the computer-mediated environment*." (Draper et al., 1998, p.7)

Although this experiential definition involve terms related to psychology, which are difficult to evaluate, it was proposed in the framework of space investigation (Akin, Minsky, Thiel, & Kurtzman, 1983), where yes indeed the mental and physical state of the users are controlled in detail.

In a worldlier plane, we may qualify our experiment according to this taxonomy as experiential telepresence environment, because promotes that the users feels to share space with distant interlocutors loosing awareness of the distance, and emulates the remote space, especially in the sound experience for the teacher.

What about telepresence in education? In the following sections there is a special section for related work and most important cases, nevertheless now is important to remark the wide use of telepresence technologies in education applications on last years. And again, as we took Draper's telepresence definition, we should exclude all the virtual worlds' projections and games. Main fields and kind of recent applications that we found are:

- Medicine and "telemedicine", for instance an interesting artifact for giving remote support to a local (resident) surgeon through augmented reality and a virtual environment, where the more learned one (attending) surgeon may advise the local surgeon since is getting a detailed, real-time insight of the surgery (Shenai et al., 2011).

- Again in medicine, but under the "telehealth" concept we found several telepresence applications using high quality videoconferencing (Conde et al., 2010).
- In higher education, a large amount of universities has already some classrooms with videocameras and high bandwidth for remote transmission, like the University of Porto (Martins & Martins, 2010) or the Manhattan School of Music (Manhattan School of Music, 2009b). Although there are several universities delivering higher grades between several countries (Multimedia's Lab from Barcelona Tech, 2012), some of them had also integrated telepresence tools for their day to day remote lessons (Giraldo, Jimenez, Trefftz, Restrepo, & Esteban, 2010).
- Engineering and science courses can also be found in telepresences kind of efforts (Lustigova & Lustig, 2009).

  In business and economy the trend is to have virtual worlds where simulate the decision over a usually virtual environment, but with little importance for the physical awareness of the student, or where the special movements and sensory feedback are not the main issue. There are many courses of music, few of theater and so far we have found none of dance, which make sense since the requirements of such remote environment may overload the available technology yet.

2.2. Specific requirements: higher music singing education

Music education has its own language, rules and is a very particular kind of learning process. Needs the student to get compromised into a daily, continuous practice and in

time require a minimum of talent at the professional level. Music education has a huge brand devoted to childhood.

"*Music education is a field of study associated with the teaching and learning of music. It touches on all domains of learning, including the psychomotor domain (the development of skills), the cognitive domain (the acquisition of knowledge), and, in particular and significant ways, the affective domain, including music appreciation and sensitivity. The incorporation of music training from preschool to postsecondary education is common in most nations because involvement in music is considered a fundamental component of human culture and behavior. Music, like language, is an accomplishment that distinguishes us as humans*" (Wikipedia, 2011).

Regarding theory, nowadays there is a discussion between an aesthetic and a "*praxial*" approach to music education; second one privileges the return to an applied, social meaningful music appreciation on a daily use, versus the classical conception of music appreciation as contemplation (Regelski, 2005). This affects all the levels of music education, since the possibility to include the use of popular music for teaching at child education level, for instance hip-hop (Irby & Hall, 2011), until the conceptions of aesthetic theory at higher education. "*The idea of* praxis *is first featured in the ethics of Aristotle who distinguished three types of knowing.* Theoria *was knowledge created and contemplated for its own sake*", (…) "Techne *was associated with poiesis, the making of things*" (…) and "Praxis*, however, involved not the making of things but 'doings' of various kinds that concern*". (Regelski, 2005, p.16) In the context of the present work, it is impossible to lose the opportunity to relate this modern, applied approach to music education with constructivist education theories, including the involvement of the learner with the activity and establishing relations from this activity with previous experiences. There is a lot of activity on music education research, since involve several

perspectives that can be attacked by musicians, psychologists, sociologists, etc. A renamed compilation is the "*The New Handbook of Research on Music Teaching and Learning*", which has been reedited in 2002 adding "new" to the title, but it´s original version have been developed 20 years before (Colwell, Richardson, & Music Educators National Conference (United States of America), 2002).

There are several recognized instructional methodologies for music education, which start at beginning stage exercises but usually allows development until the professional performance for the student. Most of them are available, perhaps having two or more, at the higher education institutions. Based on authors that perhaps have died half a century ago, usually there are some national associations that certify professor's knowledge in some of these methods, like the Dalcroze Society of UK (Dalcroze Society of UK, 2011) or the Dalcroze Society of America (The Dalcroze Society of America, 2011). Also there are schools that adopt one or other method, like the KMEIA (Kodály Music Education Institute of Australia, 2011). Let´s have a brief view on some main methods:

- **The Dalcroze method**: proposed by the musician and educator Émile Jaques-Dalcroze (1865-1950), it is based in specific exercises named Eurythmics.

  "*There are three branches of Eurhythmics*:

  - *Rhythmics: Understanding all the elements of the language of music by experiencing them through the body and learning about the basic principles relating to them.*
  - *Ear-training (Kodály in the UK): Developing skills and knowledge about pitch, intervals, scales, harmony and tonality leading to full comprehension of the building blocks of music and an ability to use these for improvisation.*

- *Improvisation (vocal, instrumental and in movement): Developing the ability to express musical concepts, feelings and form through spontaneous creation; being able to initiate movement through sound*."

  (Dalcroze Society of UK, 2011)

- **Kodály method:** Zoltán Kodály (1882–1967) was a prominent Hungarian music educator and composer who stressed the benefits of physical instruction and response to music. "*Some of Kodály's trademark teaching methods include the use of solfege hand signs, musical shorthand notation (stick notation), and rhythm solmization (verbalization)*" (Wikipedia, 2011)
- **Orff Schulwerk:** this technique is very child-oriented. "*Orff Schulwerk is a way to teach and learn music. It is based on things children like to do: sing, chant rhymes, clap, dance, and keep a beat on anything near at hand. These instincts are directed into learning music by hearing and making music first, then reading and writing it later. This is the same way we all learned our language.*

  Orff Schulwerk happens in a non-competitive atmosphere where one of the rewards is the pleasure of making good music with others. When the children want to write down what they have composed, reading and writing find their moment. Orff Schulwerk uses poems, rhymes, games, songs, and dances as examples and basic materials. These may be traditional or original. Spoken or sung, they may be accompanied by clapping and stamping or by drums, sticks, and bells. The special Orff melody instruments include wooden xylophones and metal glockenspiels that offer good sound immediately. Played together as in a small orchestra, their use helps children become sensitive listeners and considerate participants." (American Orff-Schulwerk Association, 2011)

As the name shows, such instructional methodology was developed by Carl Orff (July 10, 1895 – March 29, 1982), a German well-known composer.

- **The Suzuki Method:** "*The Suzuki approach was developed in Japan by Dr. Shinichi Suzuki (October 17, 1898 – January 26, 1998) in the mid 20th Century. He believed that every child, if properly taught, was capable of a high level of musical achievement. His philosophy was that if every child has the inborn ability to learn their native language, then they also have the ability to learn and become proficient on a musical instrument. Begin children early and excite their own inborn joy of music, or the* ***mother tongue approach***."(Suzuki Books, 2011). Dr. Suzuki has written what could be named traditional methods in scores: a series of exercises from the easiest to the more complex, for different instruments (viola, guitar, is quite named for the violin, etc.) Certainly he didn't touch the voice matters.
- **Music Learning Theory by Gordon:** this is the approach of a nowadays working professor, that was published a work with such name (Gordon, 1997). "*Music Learning Theory is an explanation of how we learn when we learn music. Based on an extensive body of research and practical field testing by Edwin E. Gordon and others, Music Learning Theory provides the music teacher a comprehensive method for teaching musicianship through audiation, Gordon's term for hearing music in the mind with understanding. Teaching methods help music teachers establish sequential curricular objectives in accord with their own teaching styles and beliefs*." (The Gordon Institute, 2011).
- **Willem method:** this is a method more oriented to improvisation at the graduated and higher education method. "*The method takes into account the difficulties students encounter when beginning to improvise. Some basic elements are the*

training of expressive abilities and the study of musical rhetoric's".(Bergstroem-Nielsen, 2011)

In terms of our experiment all those educational trends must be mentioned as part of a full and well documented theoretical framework, but they make no difference at the time of the remote communication between teacher and student. What should be stressed again is that most of the activities that happen at the early stages of music education may use some grade of automation and computer-aided support, but as we review in the next section, higher music education is a personal or person-to-person matter, and therefore technology in this case is the mere manner to communicate this remote participants of the learning process. Later we will review the specific requirements for the voice as an instrument, closing the music education sections.

2.2.1. Higher music education: a delicate and personal matter

One-to-one tutoring is at the core of learning process at the conservatory level; in a recent study (Gaunt, 2008), teachers have characterized such relationship like parental, friendly, collaboratively curious or like a doctor and patient. Teachers seek transmitting the skills acquired in their own professional life and experience, and are strongly committed to achieve students' own development, literally "*developing a personal artistic voice*" (Gaunt, 2008, p.222). Even more, when it comes to music at advanced higher education or graduated level, the council and support given by an experienced teacher is not oriented to teach the instrument: this knowledge is supposed to be mastered already to a high degree of proficiency. Therefore the teacher is devoted to review beyond the technique the interpretation of musical repertoire, conducting the expression (Karlsson & Juslin, 2008), the emotions and artistic quality of the performance, often using images and metaphors that make easier for the student to understand the feeling required (Woody, 2006). This is very well documented in our

study, we have video recording of sessions where teacher ask a student to "open the voice like a rose" or to the pianist to "think about this beautiful women coming inside the room". But it means a very important hint in terms of technology: this kind of emotional approach, this personal appreciation of feelings and other´s people feelings will be very difficult to be supplied by automata. Again, at the higher education level of music and it´s very likely that for all scenic arts, all that we can do is to communicate people with fine artifacts. This complex, emotional, artistic knowledge still resides in humans.

For the present study is important to denote that a very common practice, the masterclass, that involves one teacher reviewing the work of one student with other students attending as observers, it is also a one-to-one tutoring (Gaunt, 2008), since it does not involve a proper peer review, that is, the other students are usually not allowed to participate in the evaluation process. In a later work leaded by her students, Gaunt found that "*The students considered masterclasses to offer valuable performance opportunities, fresh ideas for approaching musical interpretation and technique and access to a professional community of practice*" (Creech, Gaunt, Hallam, & Robertson, 2009, p.326).

In terms of Opera eLearning, several sceneries justify the need for a remote class. Some opera singing teachers are active musicians, so they have seasons of concerts and travel part of the year; the need to continue the tutoring with their students is a good chance for remote learning sessions. Offering a masterclass remotely could mean that students that may not be able to travel can reach an important master. Such virtual classroom involves requirements on audio and video quality: fast speed synchronous communication, excellent view, and a high quality sound approach. And

yes indeed a Web Site platform may enhance a lot the community of practice of the students.

2.2.2. Vocal technique and singer's education

Inside music, voice is a particularly difficult instrument to command. It is inside performer's body, which cannot see exactly what is happening, and requires skills of drama and scenic presence beyond that any other instrument. Singers use to require learn foreign languages as part of their training, practice some sport as part of the throat-elongation-and-relaxation needs, and in general is a complex, difficult instrument to master at the professional level. For this study we need to understand not the whole but some of the keys of the vocal technique and singer's education that Opera eLearning project is supporting. We shall review the physiological matters, then have a brief sight on opera technique and later describe a typical opera singing class at the advanced graduate level.

2.2.2.1. Physiological considerations

What are the mechanisms that are used to construct this particular instrument? We could conceptualize human voice as a wind instrument, where the cranium's resonant cavities as well as the chest are used to amplify the sound that starts in the larynx by a glottis impulse to the vocal chords. (Baño, 2003)

Next slide shows how the flow of air should be distributed between the nose and the mouth, assuming that the tone is well placed by the singer (according to opera technique). Air flow should be supported by the diaphragmatic muscle. The palate should be in a high, curved position, allowing the free flow of the air and contributing to the resonance.

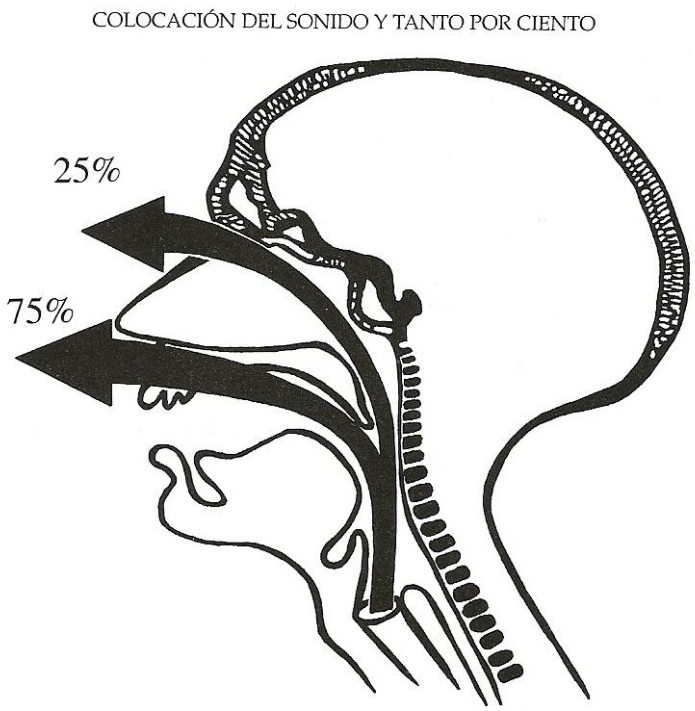

Figure 8. Placing the sound with air flow percentages extracted from (Baño, 2003, p.62).

Acoustic resonance in physics is the tendency of an acoustic system to absorb more energy when it is forced or driven at a frequency that matches one of its own natural frequencies of vibration (its resonance frequency). (Beranek, 1962) The effect of acoustic resonance is the amplification of a sound. The human body has cavities that may be used as resonators, mainly two: the chest and the cranium bones in the nose and forehead. Both are used, perhaps exploiting more one or another according to singer´s needs, obtaining the three kinds of voices from any singer: the chest voice, the head voice and the passagio or vocal transition points. (Johnson & Skinner, 2009) That means in terms of our artifact that such differences should be appreciated and qualified at remote communication.

The chest voice communicates the air flow from the diaphragm to the mouth, teeth, nose and all the cranium resonators. In classical singing, the placing of the sound

into these cranium resonators is called “use the mask” (Baño, 2003), while the right position of the mouth and throat in order to really allow the controlled flow of the air is called the “open throat” technique. This second one has been studied nowadays processing the signals, finding to be the responsible of the vibrato (H. F. Mitchell & Kenny, 2004), and certainly both are documented as part of the classical opera singing methods (Vaccaj, 1990). The next figure shows the “fish mouth” that a relaxed singer uses when accomplish this open throat and mask placing of the sound together, for instance for arising the voice to the maximum of the volume or obtain a vibrato.

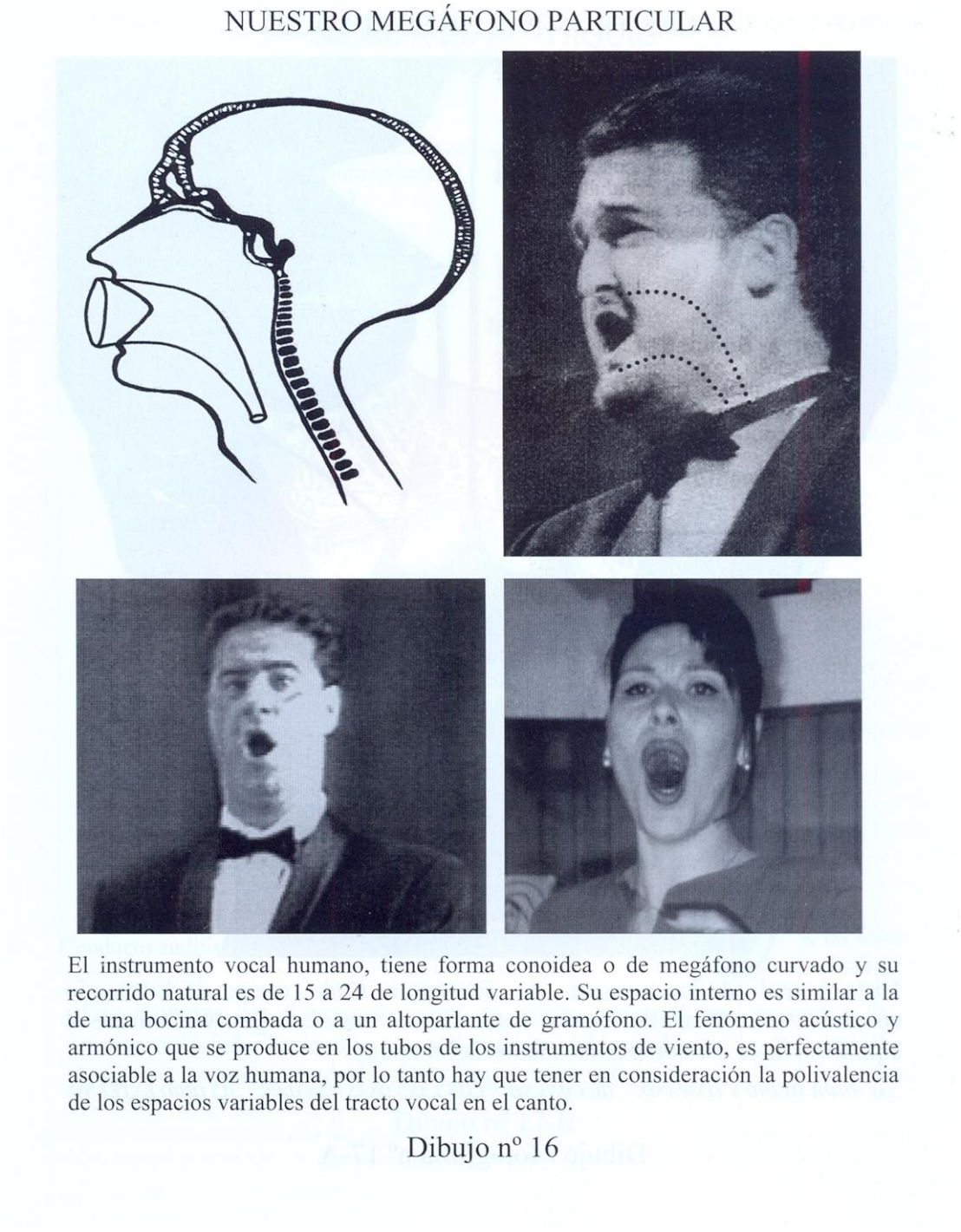

Figure 9. The "fish mouth" extracted from (Baño, 2003, p.65)

So far we have not discussed yet other characteristics of the voice like pitch range and timbre. In classical singing we have the voice classification, range or *tessitura*[4] (soprano, mezzo-soprano, contralto, counter-tenor, tenor, baritone and bass) but there is plenty more in the specific classification of such voices: coloratura, vocal weight, timbre, *portamento*, etc. The consideration of all those matters lead to the sub classification of the voice: for instance exists a soprano *spinto* or a dramatic soprano, and they are quite different, even they have specific roles in opera's productions (For instance Floria from the Puccini's opera *Tosca* should be a dramatic soprano, therefore it has some lower tones in the score that a lyric soprano may not reach comfortably, and in general is needing a more robust and darker timbre (Baño, 2003)). All such characteristics have been suspected to depend on the harmonics generated by the physiological differences between singers. A recent study has been able to predict a singer's voice classification by taking X-rays of their vocal fold length (Roers, Murbe, & Sundberg, 2009) The next slide shows this X-rays example, including a metal calibrator that is used to execute the measurement of such length.

[4] Italian words used for classical singing terms are spelled on italics characters.

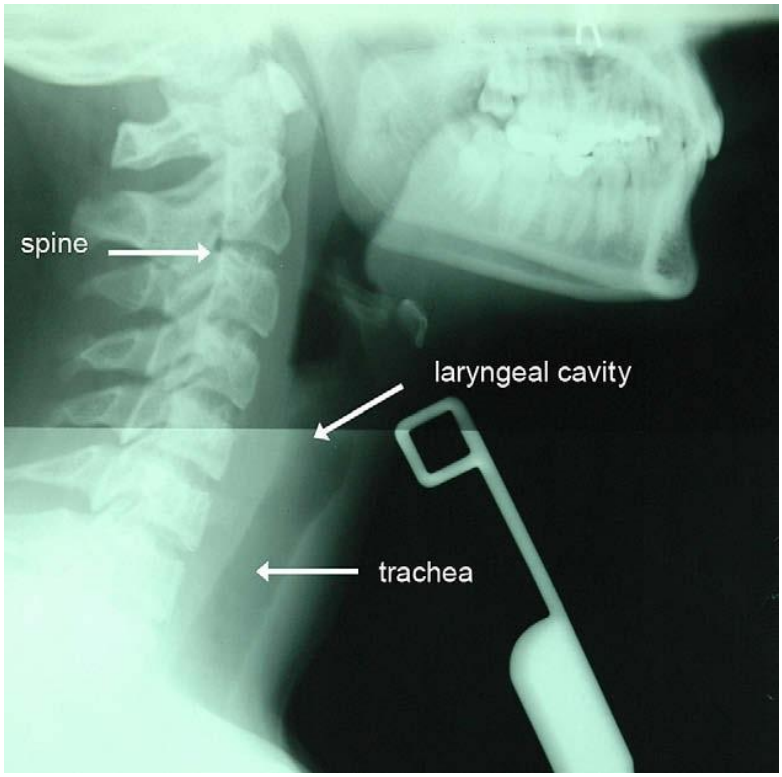

Figure 10. Sagittal X-ray image of the upper airways including pharynx, larynx, proximal trachea, cervical vertebrae, and parts of the skull. Further, the image includes a metal-distance measurement calibrator. Extracted from (Roers et al., 2009, p.409)

That is, the singer has to cultivate a technique but also learn to master the own body to exploit the most of its possibilities. Some of a singer's characteristic are nothing but physically inherited.

2.2.2.2. History and main components of opera singing technique

Until renaissance, singing and playing an instrument was a fine art that most of educated people, usually from nobility or clergy, could manage; with the baroque, a professionalization of the forms starts to cultivate genres like camera music or the Italian opera, which requires virtuosity. (Torres Mulas, Gallego, & Álvarez, 1976) We can fetch authors like Monteverdi, devoted to opera during the 1600's. In time, opera knew different stages, including the serious or comic kinds, until romanticism when many new theaters are constructed for this spectacle as part of the burgess taste.

Another genre that developed beautiful, complex songs for voice have been the German Lied (Torres Mulas et al., 1976).

In terms of technique, the so-called *bel canto* is named after the first monasteries that cultivated the religious singing at the IX and X century having an explicit knowledge, against the popular singing that have always been happening. At the Gregorian singing, that separates the voices, the principal line for a "tenor" (from the Latin words "*tenere vocem*", in the sense of maintain the sound) appears. Later the term have been used or not for different authors in the opera technique, but is certainly the root of the classical singing technique in occident. In the 1700's, Gluck propose a reform to the opera, simplifying the excess of baroque ornaments and allowing a more sentimental and dramatic experience, but incorporating most of the effects of Italian opera virtuosity. (Torres Mulas et al., 1976) At the XIX century authors like Verdi starts writing the kind of opera that we know today, for instance the famous opera "*La Traviatta*".

We may identify a set of requirements for the singers; from the beginning, we were talking about people singing without any amplification. In time, and for the new huge opera theaters, that included a full orchestra at the 1800's, so singer needed to accomplish a really loud volume. But of course, this is not a scream, but a beauty and controlled sound, and perhaps for some passages of the story told it should be a quiet, intimate tone that still is being heard at all the theater, mixed with a scenic representation of the role.

Given the previous framework, let us distinguish the main parts that compose the opera singing technique (Baño, 2003), (Vaccaj, 1990):

- Breathing: air flow like a controlled column from the diaphragm until the mouth and nose.
- Placing: the tone or the sound in the "mask" (or cranium resonators).
- Covering: the chest voice management using the palate to cover the sound.
- Timbre, coloratura, and harmonics: manage different timbres and colors, causing the changes in placing the sound to obtain different *tessituras* (from the Italian "textures").
- Phonation and pronunciation: the right pronunciation of the words, as well as use the consonants and mainly the vocals to produce the sound. Sound should use the consonants to start and open the timbre and volume at the vocals, which is actually where singing happens.

As this well defined technique exist, one may think that all the results are physically measurable, that is, that a good singer could be identified by measuring the acoustic wave produced by him or her in terms of sound. There is a lot of research on it, and it has been proven that the emotional changes in the interpretation of a piece can be measured in terms of the Fourier components of the sound wave resultant. (Rapoport, 1996) But does it means that we could measure a good or beautiful voice? Well, a very interesting result from a doctoral thesis was published in several journals and conference articles, studying the sound wave's decomposition emitted by a sample of singers and analyzed by the LTAS technique (decomposition of the wave in their long term average spectra). Comparing the results with perceptual rankings from experts gave the following results:

- Yes, identify whether or not the open throat technique was being used (H. F. Mitchell & Kenny, 2004)

- No, it didn't was relation with the quality of the voice (D. T. Kenny & Mitchell, 2006),
- And when it came to the concept of beauty, it was absolutely no relation with the perceptions and the physical measures (D.T. Kenny & Mitchell, 2004)

The conclusion is that further physical and acoustical measures may need to be taken in account for define how to qualify a voice as good or beauty according to expert´s and singing teachers (H. Mitchell, 2005). That means in terms of our experiment that at this stage of the science and technology development we could not propose the automated qualification of the singer's performance. Only the judgment of an experienced teacher is the way.

2.2.2.3. A typical opera singing class at the graduated level

A typical opera singing regular class starts with a first stage of warming the voice, by some tone placing exercises guided by the teacher, usually accompanied by a piano or keyboard. After that, some more advanced exercises attack specific troubles or abilities, for instance trills for light sopranos or sustained notes for a baritone. Finally there is the revision of some repertoire pieces, repeating some phrase that needs work. According to the teacher, class may include some relaxation or imagination exercises.

In a masterclass students uses to come already warmed, and starts directly with repertoire. However, teacher may ask the student to execute some exercise to check up the tone or sound placing in order to facilitate the interpretation of some part of the song.

Opera eLearning project have studied in prototypes schemas, the one to one class and the masterclass configuration, but this last one was the used for the case study.

2.3. Technological background and related work

Next sections will describe the technology used, as well as the scientific consideration required in order to provide the stable solution for the artifact constructed.

2.3.1. Acoustic, sound and music

Although our research does not has a specific orientation to physics, (for instance we decided not to include the wave decomposition of the recorded sound as part of the test in order to delimit the efforts), it has been necessary to assure the constructed environment's quality, and therefore some basic considerations on the mechanisms of emission, reception, perception and manipulation of the music have been done.

A wave sound is an oscillation of pressure alternatively positive or negative in the atmospheric pressure, or in any other liquid or gas material. Wave sound are longitudinal and/or transversal, and all of them propagates at a fix velocity (the sound velocity, 343 m/s at 20ºC), no matter they wavelength, frequency or direction (Rigden, 1977). It is measured in *decibels*.

Sounds from musical instruments usually support sound's tone emission in the resonance, and generates complex sound waves that have several frequencies interweaved. A musical or speech sound wave typically contents a fundamental frequency or first harmonic, and also other higher frequencies that are multiples from this fundamental frequency, and that are known as second harmonic, third harmonic, etc. Our ears interpret the frequencies mix in terms of the tone, which depends on the fundamental, while the rest contribute to the timbre of the instrument. (Moravcsik, Dorati, Rosenbluth, & Moravcsik, 2002)

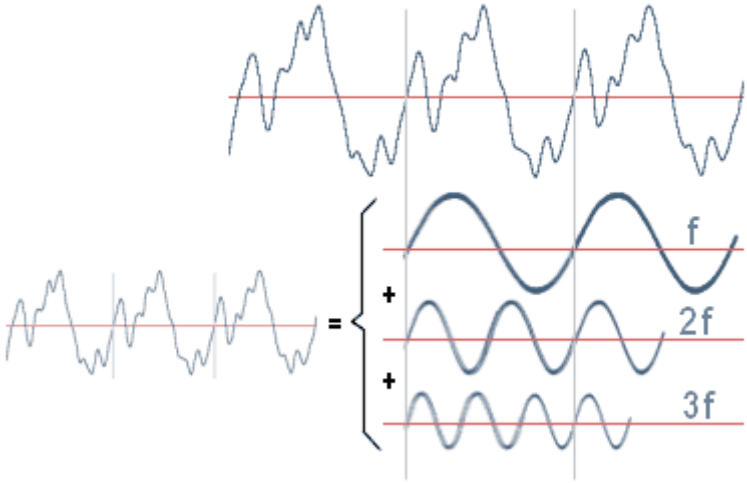

Figure 11. Wave sound decomposed in Fourier fundamental harmonics functions. From (Beranek, 1962, p.35)

Propagation of sound has a spherical form from it source; in a room, the wave sound crash into the walls. Depending on how much energy is absorbed by the material of the walls, we may have one of the following behaviors: reflex (like a light on a mirror), diffusion (like a light on a matte white paper) or absorption (like a light on a black paper). (Barron, 1993)

Those applied parameters are part of the variables considered by architectonic acoustic for auditorium's construction; whether for theater, opera or multipurpose forums, combining the described effects in the room morphology may obtain different values. Specifically to the acoustic with musical objectives the following parameters are defined: reverberation, resonance, warmth, volume, reverberation volume, brightness, diffusion, balance, mix, response or attack, texture, silence, dynamic range, uniformity and is better to obtain the lack of echo and distortion (Beranek, 1962). In terms of our experiment, we had basic recommendation for the reverberation time and levels of noise for the rooms.

About the audition, the process begins when sound wave enter by the external auditory canal, which condenses the waves and take them to the tympanic membrane, that vibrates, slowly for high frequency tones and quickly for high frequency tones. The tympanic cavity execute a secondary mission, only passing the wave sound to the middle ear, where it hits three mobile little bones: the Malleus, the Incus and Stape, that form a bridge intercommunicated. The Malleus recover the sound vibrations that are conduced to the internal ear by the Incus and Stape, that contact the middle ear stimulating the Corti organ, which transmits the reception of sound in form of cerebral stimuli. (Stevens & Warshofsky, 1981)

After the reception of the sound wave stimuli, the brain starts the interpretation that depends on knowledge and cultural baggage. Professional musicians has specialized they ear, so they have a special sensibility. Several studies support this affirmation, mentioning one recent finding, it has been demonstrated that there is a sensible difference between the ability of musicians and the rest of the people for distinguish a conversation or particular sound from the background noise. (Parbery-Clark, Skoe, & Kraus, 2009) Another matter inherent to sound interpretation, particularly with music, is the cultural and cognitive context (Morrison & Demorest, 2009).

For the Opera eLearning case study, we had the participation of experts who have all the experience for an educated ear in the satisfaction rating questionnaires. About the acoustics of the room's requirements, we could not afford to be too picky, since part of the requirements of the project is to be able to use a room in another schools or wherever the teacher is available. However, we settled the minimum sound requirements as:

- Reverberation time: is the time required for reflections of a direct sound to decay by 60 dB below the level of the direct sound. According to how big is the room, values recommended for opera singing are 1 to 1,5 seconds. Reverberation time has a lot to do with the absorption and reflection of the sound in the walls, so if we found a value too bad, we can always use a curtain or a reflecting surface to adjust this value.
- Ambient noise: it should be necessarily less than 40dB.
- Pink noise: this test has been done to have registered other acoustic phenomena that may affect the results later, like the frequency response of the amplification at the room, amplifying or neglecting some of the frequencies.

After testing those values some adjustments were done at the sound equalization of the engineer mix table or at the speaker's position at each of both sites (local and remote).

Finally in this auditing at distance exercise the influence of the acoustic configuration of each room, both transmitter and receptor, are overlapped in the speaker's resultant. That´s why each testing exercise has required a large interval of try-and-test with the position of speakers, this noises testing, equalization and previous singing exercises of until four hours before the masterclasses.

2.3.2. High bandwidth and video transmission applications

Internet 2 is a consortium of entities from the United States that today has more than 200 members, after start working at 1996 with 34 members, with high bandwidth at high network speed available thru optic fiber or satellite transmissions. (Internet2, 2009) At the beginning there were only universities, but soon received private corporations' memberships. Nowadays continues the labors of research and education, and has its

European counterpart at the Géant2 (Géant2, 2009), that has 34 countries and started to work at 2004.

In terms of the impact that this wide access to high bandwidth had done at the audio and video transmission technology, the need of compression formats could be dismissed, for instance having 10 GB per sec we could send several times high definition video signals. However, main limit so far has been to connect new nodes to these networks, since the initial investment required for the installation of required infrastructure is very expensive.

When first high bandwidth or Internet2 transmissions started a quick incorporation of many nodes at worldly level it was expected. That was not the case. In the other hand, in the last years the ability of sensible lossless compression for domestic internet makes that some applications that were supposed to require Internet2, like sharing videos at an acceptable definition, are available already for best quality domestics networks. The conclusion is that high bandwidth has allowed the transmission of Opera Concerts alive, for instance, or other huge spectacle supported by satellite or optic fiber, but in the other hand the domestic formats are growing in quality and capacity, and in time is very likely that a regular user may afford to participate of such transmissions not from another theater, but directly from home.

From the tech point of view, besides the ability of sending and receiving at high bandwidth we need a software application that manages such transmission. The application would accomplish the common matters in communication, as encoding and decoding the information, the division of the information on packages and the management of the route of such packages from the two points. We have used three of them, selected by their quality, usability and economy:

- DVTS is an application for sending and receiving DV (Digital Video) streams using the Internet. Currently DVTS works on FreeBSD and MacOSX. IEEE1394 (*Firewire*) cables are used for connecting DV equipments to a computer. However, the length of a single IEEE1394 cables cannot be longer than 4.5 meters. Using DVTS, DV data can be sent anywhere using the Internet.

  DVTS is an Internet Protocol based high quality real-time audio-visual (AV) communication tool. As DV camcorder and VCR are products available for the consumer market, they are adopted as AV CODEC. Thus, DVTS does not require especial and expensive equipments. The AV devices having IEEE1394 interface are connected to the IP host with its interface. (Digital Video Transport System, 2009) This application uses some CODEC compression and is able to send video and audio, nevertheless we have used it as the video sending signal for our most economic design, sending audio separately. The philosophy of DVTS is to interconnect two video cameras; which allows the bidirectional communication at not too expensive high bandwidth availability, and is designed for general purposes videoconference communication.

  DVTS has been developed mainly in Japan and is an open source project.

  DVTS uses a standard definition and requires 30 Mbps of band width, quite affordable for most institutions, and provides resolution (PAL 720x576) enough for big screens projection, allowing to distinguish the basic postures of singers. DVTS allows manual changes in transmission's configuration on real time according to connection's behavior, modifying the frames per second and therefore the dimension of the IP packages, allowing a clear communication in hostiles or saturated networks; can manage a stereo audio signal, besides the video. The delay

obtained is about 300 milliseconds from the application, besides the delay from the transmission medium, for instance 5μs/km at fiber optic networks.

- Ultra-Videoconferencing is a project from the Ultra-Videoconferencing Research Group located at McGill University in Montreal, Quebec, Canada. Most of their projects are based in the Centre for Interdisciplinary Research in Music Media and Technology (CIRMMT) of the Faculty of Music. Its main focus is continuing development of the McGill Ultra-Videoconferencing System for high-bandwidth, high-quality, and low-latency videoconference interaction, such as that required for high fidelity instruction and distributed musical performance. The overall objective is to create the feeling of "being there" wherever "there" may be. The Group has made their software available free for broad public use while restricting resale and incorporation in commercial products. (McGill Ultra-videoconferencing, 2009). Ultra-videoconferencing supports high-resolution multichannel audio and video from frame grabbers using v4l or v4l2 interfaces (e.g., Bt878 chipset), digital video (DV) cameras, DCAM (raw, uncompressed video) 1394 devices, and standard or high-definition Serial Digital Interface (SDI). Nowadays they are working on minimizing latency to facilitate the interpretation of music online between remote interpreters, which is one of their main interests. This is a fine application, that consumes a lot of resources but obtain an excellent audio signal, specialized for music transmission. The application allows managing the parameters for the A/D conversion, like the sampling rate and quantifying bits, also allows synchronizing the lips via software (there is always a delay and the sound gets first than the video). But is limited to two channels, so it has been used for testing the stereo configuration as a good audio choice. Uses 2Mbps per channel at 96 KHz/24bits. The delay introduced by this application is from 20 to 50 milliseconds.

- Ultragrid UltraGrid from Laboratory of Advanced Networking Technologies (ANTLab) is a software implementation of high-quality low-latency video and audio transmissions using commodity PC and Mac hardware. Supported resolutions range through HD (1920x1080) up to 4K (4096x2160). The high-quality is achieved either by using uncompressed streams or streams with very low compression ratio. End-to-end transmission latency (i.e., all the way from the camera to the display) is about 100ms, but it varies based on camera and capture cards being used. UltraGrid was originally a research project used to demonstrate the possibilities of 10Gbps networks and to study multi-point data distribution in such environments. Recent advances in the field of GPU-accelerated low-latency codecs extend its usability also to Gigabit networks. UltraGrid is supported on PCs with Linux operating system and Macs with MacOS X. The software is open-source distributed under BSD license, i.e., they are interested in both research/academic and commercial applications. Nowadays, main application areas are collaborative environments, medical, cinematography and broadcasting applications, as well as various educational activities. (SITOLA, 2009). This software has the ability to send in different formats the information, particularly the non compressed high quality video. Have been used for testing the HD transmission formats. The critical point with this software is the high bandwidth requirement, that may reach the 1,5 GB per second using HD-SDI with 10 bits per color (3 components x 10 bits = 30 bits). That means to have around 990 MB per sec; let´s say a 1 GB per second dedicated connection. With the cards used for HD-SDI capture it is possible to transmit until 6 audio bidirectional channels. This software is specially intended for cinema screens or theaters, since it manages a wide video resolution (1920x1080). If necessary because of transmission's condition we can cut some of the frames sent, but it

affects the sound. The delay of this software is about 2 frames, that is, 2x1/30 is about 66 milliseconds, plus the transmission's delay of 5μs/km.

2.3.3. The recording and displaying studio

Besides the room's acoustic and the signal transmission matters, we shall describe the audio and video recording material, and justify the selection of such devices. It must be said that we have registered all the material in high definition video and DAT recorder (professional sound), because for a masterclass perhaps the material obtained may be used for some educational DVD or other publishing purposes. However, in terms of the remote conference the signal may be of lower resolution, for instance some of the applications described in the last part make some compression of the video signal. This is very important to keep in mind for the stereo, one-to-one configuration of the artifact, since not always is necessary to deploy a full masterclass environment for a follow-up class of a particular teacher with some student.

2.3.3.1. Professional video cameras, microphones and other toys

The particular selection and use of the equipment is described in "tools and materials" sections, both for the pilot stage as well as for the case study. In this section we shall review some of the characteristics and requirements for these equipments. Sometimes only professional equipment was used only because a small but relevant characteristic.

- **The video cameras:** it was necessary to have at least a professional video digital camera that would be able to record in separate channels audio and video. Nevertheless the audio was going to be recorded in others separate channels, the original audio from the video helps to edit the video later. The 720 x 576 PAL standards were enough, but at least we needed the 25 fps and a signal of PCM 16 bits output. Of course digital outputs at high speed compatibles with IEEE at the

computer were required, and it was nice to have remote stop and pause abilities. High definition video was required only for the high definition testing prototypes, but we had most of the time a HD camera available.

- **The condenser microphones:** against the dynamic microphones, which are cheaper and mostly used for on-stage work, musicians use to have condenser microphones for recording sessions. "*A microphone is an acoustic-to-electric transducer or sensor that converts sound into an electrical signal*."(Wikipedia, 2009). A particular manner that this sensor can use is the condenser. "Condenser *means capacitor, an electronic component which stores energy in the form of an electrostatic field. The term condenser is actually obsolete but has stuck as the name for this type of microphone, which uses a capacitor to convert acoustical energy into electrical energy."(…) "A capacitor has two plates with a voltage between them. In the condenser mic, one of these plates is made of very light material and acts as the diaphragm. The diaphragm vibrates when struck by sound waves, changing the distance between the two plates and therefore changing the capacitance. The resulting audio signal is stronger signal than that from a dynamic. Condensers also tend to be more sensitive and responsive than dynamics, making them well-suited to capturing subtle nuances in a sound*." (Media College, 2010)

  Although the use of condenser microphones made more difficult the handle of possible saturation at the speakers, the transmitted sound really improved, as was confirmed by testing in some prototypes.

  Another characteristic from microphones is the direction; some of them are design so they sense the sound in a spherical region in front of them, or in any direction (*omnidirectional*). This is called "the polar pattern" of the microphone.

Most used for singers are cardioids or hypercardioids (Larson Guerra, 2010). Actually for our experiment this has nearly not relevance at all, since singers were almost in the same place most of the time, and didn´t walked on stage or changed their position at the microphone.

- **The A/D converters:** the analog to digital converters where required for the audio signals. For the multichannel Dolby 5.0 simulation prototype required five channels were required from the students to the teacher, plus two from the teacher to the students. In the market the nearest solution is getting an A/D with 8 channels. Professional extern audio cards are available for that purposes, and we used a rate sample of the sound of 16 bits and 48 kHz, which is less than the CD quality that most of the audio cards with A/D converter has available.
- **The speakers:** auto amplified acoustic boxes with a wide frequency range (60 Hz to 55 kHz) were good enough. Regarding the speakers, a fine adjustment was to test and tune the pink noise parameters at the mixer tables, and in the other hand, the case study showed that the best thing to do is to distribute them in the room space.
- **The screens:** plasma monitors at 50" minimum, HD 1920 × 1080i30 & PAL 720 x 576 were used. However, smaller monitors in HD and even with less resolution could be used with good results.
- **The mixers:** we used best professional audio mixers, but actually the requirements for the audio mixer are quite modest. We only need to control pink noise and equalize if required each channel separately, and is absolutely required that we may apply milliseconds of delay to several channels coordinately, and perhaps tune this value in time. This is because the sound signal will arrive to

the transmission before that the video, and this tiny adjustment in the audio mixer fixes that matter.

2.3.4. Related work

Music higher education institutions are already delivering remote classes and educative content through high quality videoconference and broadcasting (Manhattan School of Music, 2009b). High-bandwidth availability allows improving videoconference's quality. Such is the case of the use of Internet2 (Internet2, 2009) in Miami for orchestral higher education level (New World Symphony, 2009) . Our case study would be the first experience only devoted to opera singing, and is using the Géant2 European network platform (Géant2, 2009).

Next table shows some of the references found through this research. It is very likely that in time there will be more institutions in remote learning of music, and of course this table only have a few that have been relevant to this research because some articles of their results have been cited, their work shows some particular way to do things, or simply because the software that we used made us contact with such research group. Several commercial web sites provide media distribution of music lessons in low resolution videos, mostly at starter level and are not being considered for this list.

Related work references	
Institution, link and reference	**Key Features**
European Comission founded projects	
I-Maestro (i-maestro, 2009) http://www.i-maestro.org/	Huge Project between top line european and UK universities and some technology providers, has features like the remote control and review of posture and gestures, as well as a cooperative environment. One of the most interesting proposals is the Symbolic Music Representation standard to be included also in MPEG-4 format as metadata, already recognized by the ISO. They have curricula proposals and standards as well.

	Recording sessions and the use of sensors to correct the body gesture, as well as machine recognition of online video has produced interesting results properly documented in the academic literature.
VEMUS (Virtual European Music School, 2009) http://www.vemus.org/	Devoted to provide online resources for wind instruments, as a social manner to offer training for European popular music in a blended learning approach.
OpenDrama (Open Drama, 2012) http://mtg.upf.edu/opendrama/	"*The project aims at the definition, development and integration of a novel platform to author and to deliver rich cross-media digital objects of lyric opera and other vocal dramatic music, opening this heritage to a dimension of learning, exploring and entertainment*". It´s remarkable the use of MPEG-7 for storing metadata. Includes an Opera Tutor for recording the singing voice while reading the scores from a Java application at any desktop station.
Public or private universities	
Manhattan School of Music (Manhattan School of Music, 2009a) http://dl.msmnyc.edu/	At the beginning of this research the only delivered some media content and were exploring remote videoconferencing. Today they have putted in place a Global Conservatory to be able to deliver full remote classes, although they do not provide the degree online, but rely on a remote institution to certify the knowledge of the musician in other country. They use a very similar way to deliver masterclasses using high bandwidth in auditoriums supported by audio and video engineers, as Opera eLearning project does. They teach singing, mostly jazz. For one to one sessions they use and loan a commercial solution (Polycom, 2012) that was not used for Opera eLearning because it is a closed box and we wanted to experiment in detail with audio signal manipulation.
Carnegie Mellon School of Music (Carnegie Mellon School of Music, 2012) http://music.cmu.edu/	In a conservative way, this important School of Music has only developed webcasts (in HD from 2008) for delivering their live presentations to the public.
New World Symphony (New World Symphony, 2009) http://www.nws.edu/	This private institution located in Miami has taken advantage of high bandwidth technology and participate from the Internet2 project from the beginning. They tried to deliver the whole music degree online (accessed on 2009), but now (accessed on 2012) they remove this option, requiring that students go physically to the school for some time in a blended learning approach. This institution does not teach singing.

Berklee School of Music (Berklee School of Music, 2012) http://www.berkleemusic.com/	For the whole degree they have putted in place agreements with some music schools in other countries (Japan and Mexico), while for credits have a wider range of institutions around the world for sharing courses. The online courses allow obtaining minor certificates for music production, songwriting, etc. There are jazz and rock voice training courses, and all of them use video recorded sessions, a web site with interactive media content and some remote sessions.
Open College of the Arts (Open College of the Arts, 2012) http://www.oca-uk.com/	This institution delivers whole professional degrees on fine arts, and has some music modules on history and composition, but not devoted to some specific instrument. These modules are part of degrees like creative writing, and do not include any videoconference, but e-mail and phone feedback from the tutor.
Research groups	
Center for Computer Research in Music and Acoustics (CCRMA, 2012) https://ccrma.stanford.edu/	This center includes a lot of research on education and practice of music with technology support, besides music research itself, has events, publications and a wide range of interests. There is a lot of open source activities.
Centre for Interdisciplinary Research in Music Media and Technology (CIRMMT, 2012) http://www.cirmmt.mcgill.ca/	This research group located in Montreal is who provide the Ultra videoconferencing software used in Opera eLearning case study, and belongs to the Schulich School of Music of McGill University. They have explored several matters about perception, technology and music. One of most fascinating laboratory facilities they had built is a listening room that has a sphere of high quality speakers, where listener located at the inner space of the sphere gets to experience different surrounding sound sensations.

Table 4. Related work references.

3. Methodology and research design.

This chapter summarizes the scientific method followed, explaining the theoretical framework for the case study and describes the research design, giving support to the treatment of the data obtained.

3.1. Case study

The case study is a research strategy for empiric studies that comes from social science and that has proved to be useful for technological projects evaluation in the last years:

"The nature of industrial software engineering does not match well with the restrictions imposed by formal experiments. We cannot usually perform randomised controlled trials in industrial situations and without randomised controlled trials we cannot assess the actual impact of competing technologies, nor can we assess the context factors that influence outcomes in industrial situations." (Kitchenham, 2007, p.31)

Then, a reasonable approach is a case study, because allows the analysis of the phenomenon on the real life, for instance in social science a case study may get to observe the habits from a certain community, while in technology we may observe a system being used by a community of users.

Three points must be assured for a good case study research (Yin, 2003):

- Constructs validity, that is, that the concepts to be measured, as well as the dependent and independent variables involved are all well defined and treated with the proper tools and scales for their measurement and later analysis.

- Extern validity for knowledge domain, that is, that the study is well justified and can be defended in the academic (or industrial) media.
- Experiment reliability, which means the replication possibility as well as the control of the course of the case study, understanding that we won´t be able to manage all and every strange variable, but we have to provide a stable, reproducible environment.

The methodological check list of activities and questions to be solved in order to assure constructs validity, extern validity for knowledge domain and experiment reliability includes choosing a kind of case study depending of the conditions, and the focus is on properly define the hypothesis and collect and process the data. (Yin, 2003) When researchers are facing a phenomenon that was not available to scientific observation before, then it is justified to execute the case study with only one experience, whose nature should be exploratory, since there is no manner to have previous experience. This is the case to brand new technological proposals.

"An exploratory case study must be faced with courage. At the beginning there may be uncertainty about some important issues of the real case study – the questions to ask, the hypothesis, the data collection methods, data access o data analysis methods- and therefore such matters requires a pilot research. Once those are defined, and the pilot stage is complete, we are ready to start the case study from the beginning…" (Yin, 2003, p.82).

Such consideration has been very useful for this project that has split the activities in two stages: the pilot and the case study. There is a common error in confuse the data resulting from pilot stage and include it in the case study, that should be avoided. But as

we will discuss in the case study chapter, such pilot stage have been crucial to de success of the project and has constructed the hypothesis and tests for the case study.

In order to defend the case study in the external domain, that is its validity for the scientific community, it should be defended as it does an experiment with analytic results, and the generalization obtained is analytic, qualitative and quantitative, but not statistic (Yin, 2003). The case study is different from the survey as research methodology, because in the survey a representative sample of the universe is chosen, and therefore the generalization of the results is statistic.

Does that mean that we will not use statistics for our data? No, this is not the case. We shall and will use statistics tools for analyze the data, as part of the quantitative methods. It means that the results from a scientific point of view have another interpretation. A survey or other statistic supported method, as a formal experiment, sustain that the result can be reproduced in the controlled environment: if we take other sample of the universe, we should find a similar result in a statistical range of reliability. This is the classic manner of generalization. In the other hand, the deductions obtained from a case study do not always lead necessarily to generalization, or if they do it is only in the framework of conditions of the experiment. A case study result argues that in similar conditions the phenomena studied should present similar behaviors. This is what Yin calls "replication logic". (Yin, 2003) In a colloquial manner to speak, we should be positive about that we can replicate with success the Opera eLearning environment in another university, but is very likely that if we measure the network connection or compare the rating questionnaires results, we may find differences that cannot be predicted statistically.

Another step that can be used to reinforce the extern validity of the case study is the triangulation of data recollection, so we may check the same result from different perspectives.

Our investigation group at the Multimedia Laboratory has a wide experience on the use of case studies; a fine comment on how to design good case studies may be found at (Ferruzca, 2008, p.110): "…a lot of case studies are done, but few accomplish criteria enough to be meaningful. A good case study implies: specify the study's hypothesis, to establish de project selection and data analysis criteria, establish a good comparison basis, carefully planning and the use of the proper data analysis and results presentation." In this definition we found a new issue for the themes treated so far: comparison. A case study should be able to be compared with other studies as part of the results analysis. That´s why in this case study, that is exploratory since we have no precedent, it has been a very important task to recover and identify similar projects that may serve as comparison basis.

3.1.1. Hypothesis, variables and constructs

In the development of the text regarding the empiric work we may trace how these values have been proposed, tested and discarded or formally defined as part of the study and mapped into some construct for the data recollection. However now as part of the research design it is important to stress which are the questions that this case study is intended to solve.

The hypotheses were user oriented: the model of remote singing learning room:

(1) Was enough for the teacher to fully evaluate the student

(2) Allowed the student to understand observations and emulate the teacher's sound

(3) Observers could use the information received as well

(4) A high quality, immersive and space distributed sound could improve teacher's listening.

The configuration of each Opera eLearning environment or prototype provides the independent variable context. By modifying the conditions on register, transmission and reproduction of the sound we control the environment. We have two kinds of variables:

1. Variables affecting the register, transmission

In terms of variables, we have two kinds.

1. Variables that are defined by the register, transmission and reproduction of the sound conditions. Controlling those variables on Opera eLearning configuration measures we have a stable, reproducible environment.
2. Variables measuring user's perceptions about class reception, classroom effectiveness and communication fluency. From these variables on perception the lyric singing education hypothesis are argued, with source on:
    a. Rating questionnaires
    b. Open interviews only with teachers
    c. Spontaneous dialogs in the class that has been of interest in the further review of the recorded material. Such recorded material is available in HD definition video, as well as recorded in separate or mixed channels in professional quality. Are treated with a qualitative conversation analysis approach.
    d. Network status statistics, wave form of recorded material, and a lot of recorded material available.

Variables that have been researched all along, including the pilot stage, from the audio configuration are:

- Micropohes: types and space location
- A/D converters
- Transmission software
- Speakers: types and space location

From the user's perception of audio quality:

- High, medium and low tones
- Harmonics
- Timbre and dynamics (between *piano* and *forte*)
- Communication fluency
- Perception on differences on speakers and microphones position in the room.

Another group is the "Classroom effectiveness" construct, which has more to do with the matters of educational value or curricula importance of this artifact from the pedagogical point of view.

For the analysis, categories of research were defined, which are listed in the following table. Such categories can be reviewed against each of the hypothesis. The research categories defined and tuned after the pilot stage may be reviewed in the **Table 14** from this Thesis.

Each subcategory is mapped into a set of constructs that shows specific information about that subcategory, for instance: high, low and medium tones reception in a Likert scale from 1 to 5.

As strange variables control, previous tests on several values like noises values on rooms where we were planning to execute a masterclass, as well as the network

connection data statistics. A lot of important variables could not be controlled, like student's selection in the population available.

3.1.2. Quantitative and qualitative analysis

So we have all this data about questionnaires and video recordings. What is the most meaningful way of obtain data from them?

The rating questionnaires

On the quantitative side statistics from Likert Scales are somehow restrictive, since they are nominal, in the best of the cases ordinal type variables. But again this study from the scientific point of view is mainly argued on the analysis of one case that is one of a kind. Actually formal statistics treatment would need other cases or samples, and there is none. We cannot define samples or a universe. Anyhow variables have been analyzed with statistic frequency descriptors and deviation values, and will be exposed. A

Statistics is defined as the practice or science of collecting and analyzing numerical data in large quantities, especially for the purpose of inferring proportions in a whole from those in a representative sample. There is a lot of literature on statistics (Freedman, 1991; Hernon, 1991; Kshirsagar, 1983) as the pillar of the scientific method for experiments analysis in most of the fields. One very well known author is Tukey, who worked on the field since the 60's and had a big interest on education (Tukey, Hoaglin, & Mosteller, 1985).

The recorded video sessions

First work about the video session is to arrange a video compilation with both sides at the time, by carefully mix and coordinate both recordings. Then we can appreciate the classes comfortably.

Videos where exhaustively reviewed, integrating notes that later were the source for conversation analysis, which is a social science method that formalizes the discourse analysis with a strong theoretical basis.

Conversation analysis investigates interaction as a social institution. First author in propose this though is Goffman at the 60's, as described in (Heritage, 2001). Erving Goffman was an American social scientist, whose "*fundamental achievement was to establish that social interaction is a form of social organization in its own right. Social interaction, he argued, embodies a distinct moral and institutional order that can be treated like other social institutions, such as the family, education, religion, etc. Goffman came to term this the* interaction order" (Heritage, 2001, p.2). He founds this order reflected in the syntax of the interaction, conceptualized as a socio-logic of interaction that provides for the sequential ordering of actions. The main idea is that we can obtain identities, power or institutional representation or discourse, intentions and other information from studying the conversation. However Goffman gave a lot of examples of how to use this syntax analysis from conversation, perhaps missing a more substantial argument from the theoretical point of view (Schegloff, 1988)

This approach was soon developed by another important American scientist, Harold Garfinkel, who "*argued that all human action and human institutions rest on the primordial fact that persons are able to make* ***shared sense*** *of their circumstances and act on the shared sense they make. Garfinkel wanted to know how this is possible, and*

he hit on the notion that persons use shared methods of practical reasoning ('ethno-methods')" (Heritage, 2001, p.3). These ethno-methods are being widely used until today for analyzing conversations, pieces of recorded text, written texts, etc. in social science. Garfinkel had a very scientific oriented thought that made him execute quasi experimental procedures for develop his proposal, delivering a set of empiric working practices well documented.

"*Conversation analysis (CA), developed by Harvey Sacks in association with Emanuel Schegloff and Gail Jefferson, emerged in the late 1960s at the intersection of the perspectives developed by Goffman and Garfinkel. From Goffman, CA took the notion that talk-in-interaction is a fundamental social domain that can be studied as an institutional entity in its own right. From Garfinkel came the notion that the practices and procedures with which parties produce and recognize talk are talk's 'ethnomethods.' They form the resources which the parties unavoidably must use and rely on to produce and recognize contributions to interaction which are mutually intelligible in specific ways, and which advance the situation of interaction in an incremental, step-by-step fashion*."(Heritage, 2001, p.5)

Inside the CA we have different practices; in the case of our study we shall distinguish the sequence organization and word selection (Prevignano, Thibault, & ebrary Inc, 2003; Schegloff, 2007), that will be applied by the search of specific terms used or some sequences of meaningful interactions, like the review of dynamics on interpretation from the teacher to the student.

"*The field also began to diversify into domains of interaction - such as legal proceedings, doctor-patient interaction, calls to the emergency services, news interviews and classroom interaction - which are socially and organizationally distinct*

from ordinary conversation (…) *CA studies of institutional talk effectively examine the management of social institutions in talk (Drew and Heritage, 1992; Heritage, 1997)*" (Heritage, 2001, p.7).

This is an important hint in terms of our research: "*Conversational analysis provides a detailed methodological tool for recover the details of a discourse interaction in the framework of an institutionalized, intentional and asymmetric situation, as the medium level classrooms*" (Escudero, González, & García, 1999, p.12). And we shall be able to deeply notice these three characteristics in the analysis of the recorded sessions.

A typical manner to display a CA study is to list some of the most important pieces of conversation transcribed, and comment the details of the lines according to the research categories previously designed. Usually occurrence count of the terms is also a common activity. Both of them will be used in this study.

Other registered data

Other registered data was used at convenience and was available on request in case of troubles. That is the case of some network connection issues.

For tuning the prototypes on the pilot stage, we had a few of sessions at the Liceu Opera House with the head audio engineer, hearing the separate channels and exploring possible alternative configurations in idea storms sessions. Manipulation of audio channels recordings has not being analyzed later.

Lastly we may summarize the case study methodology applied to the Opera eLearning study as the guide to organize the activities of research: we have split the study in the pilot and the study stages; we have characterized the study as an exploratory case study; we have defined the hypothesis, variables and constructs; we have reviewed the validity of constructs and extern domain; we have organized the recovery and analysis of the data, and have comparison basis from other projects to evaluate our results. Now we have an organized case study to go on.

3.2. Usability methods

Usability has been a polemic term; begins to appear in the scientific literature in the early 80's, and is related to concepts like user friendly and user's learning curve. For a while it was not clear if the systems should have a user's oriented design, or be designed in order to optimize the business process or the computer power available. An interesting sample of such dichotomy can be found in an article that proposes "conceptual detours", that is, changes in the manner of doing stuff, so the user may be adjusted to a system that is designed *"in order to exploit characteristics and potential of the computer-based medium*" (Wright & Bason, 1983, p.1). Nowadays that seems strange, but at that time it was difficult to foresight the power of the processors and memory that we have today, or even though about a WWW with high speed.

In time usability began to be mentioned related to a key concept for this study: testing, and in particular, testing of the process of the use that the humans make of one system. In this sense is related to human-to-computer interaction (HCI). At the early 90's, we have already several authors providing proper definitions of usability, as: "*that usability can be explicitly defined and tested, in a way that can lead to improvements in design. The essential prerequisites to usability are considered—know the users, their*

characteristics, their tasks and their environment. Usability is then defined in a manner that can incorporate users, tasks and environments to set explicit performance goals relating to the effectiveness, training time and subjective opinions of sample users." (Heaton, 1992, p. 149)

The formalization of usability as one of the software´s attributes that should be evaluated became when the International Organization for Standardization, ISO, mentions usability as part of the ISO 9126 rule (International Organization for Standardization, 1991). That standard is a milestone for software quality, since there is no previous standard that is devoted to software product. That is, all the previous ISO standards where used to evaluate the software productions process. Then, you could have an ISO certificated company producing very bad software, certainly in an ordered and well documented process. That was a major issue for industry at those days.

ISO 9126 distinguish six factors for measuring a software product: functionality, reliability, usability, efficiency, maintainability and portability. (ESSI-SCOPE, 1991) Regarding usability, divides its measures in three dimensions:

- Understandability (Attributes of software that bear on the users' effort for recognizing the logical concept and its applicability)
- Learnability (Attributes of software that bear on the users' effort for learning its application)
- Operability (Attributes of software that bear on the users' effort for operation and operation control)

ISO 9126 it is used nowadays, and has its last reviews at 2004 for specifically define and review metrics recommended for the measure of usability (International Organization for Standardization, 2004).

A mature definition of usability is provided by Bevan at 1992, like: "*usability should be defined as the ease of use and acceptability of a product for a particular class of users carrying out specific tasks in a specific environment ...*", although there are "*different approaches to usability based on the product, the user, ease-of-use, actual usage and the context of use*" (Bevan, Kirakowskib, & Maissela, 1992, p.1).

Soon several methods of usability testing where formalized and published; an scholar sample is a very hands-on manual on usability testing by Rubin that have help situate and define the environment testing of this study in a user centered approach (Rubin, 1994)

In Opera eLearning the goal was to provide seamless interaction between humans developing a remote singing class at a distance. To achieve the proper design and implementation for the project, a multidisciplinary team was created. Network and audio engineers, multimedia experts, singers and professional musicians, as well as singing professors, chorus conductors and theatre directors have worked together in designing, testing and evaluating the proposed model. A pilot stage served to define and test the model through six different prototypes, and therefore the case study had two functional prototypes, all of them oriented by a user-centered design (Bevan, 2001), including all the activities defined by the ISO 13407 "Human centered design processes for interactive systems" standard for each of the generated prototypes (International Organization for Standardization, 1999). These activities are: plan the user-centered process, specify the context of use, specify user requirements, produce design solution and evaluate, in an iterative cycle (Niegel Bevan & Curson, 1998).

For the usability evaluation we resorted to experts' inspection methods, as well as user-satisfaction questionnaires and interviews (Rubin, 1994). Usability and user-

centered design were so important for the pilot and further case study, that the working hypotheses are related to the usability goals defined and tuned during the pilot stage.

3.3. Design & Innovation

Such an effort drives to innovation since it generates valuable knowledge, using design management techniques like prototyping and usability approaches (European Commission, 2004), apart from generating feedback between the university, the research community and government institutions in solving a specific problem. We found that this kind of work has also been described as "design as research" (Purao et al., 2008), and it is characterized by the multidisciplinary approach, as well as the design activity as knowledge-generator, that can be described into a model, construct, method or framework. An important contribution to this "design as research" (Purao et al., 2008) theoretical corpus is made by Carroll and Kellogg (Carroll & Kellogg, 1989), who describes artifacts as the nexus of theory-based design and hermeneutics. The process used can be also described as either an early phase of an innovative project (Marcandella, Durand, Renaud, & Boly, 2009), or as a collaborative research across discipline boundaries (Cummings & Kiesler, 2005).

3.4. Systematic literature review

Technology studies needs to be updated all the time, since advances and new opportunities are available all the time. That does not means that you should *change* your model at each new trend from the industry, but yes is very important to be aware of new products and trends. In the scholar literature, it is as well relevant to have a continuous review of advanced proposals for new theories and contributions that may be affecting the field of distance education.

In the present study a systematic literature review have been putted in place. "*SLRs are secondary studies (i.e. studies that are based on analyzing previous research) used to find, critically evaluate and aggregate all relevant research papers (referred to as primary studies) on a specific research question or research topic*" (Kitchenham et al., 2009, p.336).

We have included in such review more than papers, because previous telepresence or video experiences in music education have been also allocated and documented.

The terms that have been searched are listed in the table below, and includes as told different fields. For determine historic evolution of some terms in the formal literature, results from search engines have been filtered by year, trying to find the first occurrences of the term in a three year period, and later analyzing the state of the art of the concept today with the same three year frame time. Another search refining criteria used is to mix with another concept, for instance we may have tried with telepresence, and then refine with the word "education" or "music". While some of the concepts are the same spelled in different order (like distance education and a refining criteria as music against music education and a refining criteria as distance) the articles found where different, because the search engines selected different order priority of such combinations, although of course some works and authors were found several times.

Concepts searched in literature review for Opera eLearning study	
Concept	Refining criteria
Distance education	music, higher, internet2, video, telepresence, hybrid, blended, voice
Remote learning	music, higher, internet2, video, telepresence, hybrid, blended, voice
Blended learning	Method, case study, music, higher
Video education	Blended learning, higher, first and last years, internet2, high bandwidth
Virtual classroom	Case study, music, higher, telepresence, Internet2 , first and last years, Geánt2
Telepresence	Learning, education, classroom applications, first and last years
Internet 2	Education, remote learning, first and last years
Case study	Methodology, design, scientific method, first and last years
Innovation	Methods, design as research, knowledge management
Music education	singing, distance, higher, vocal technique, voice, opera

Table 5. Concepts searched in literature review for Opera eLearning study.

Regarding the used engines, we may distinguish between the academic or formally published material, and the grey literature. Next table lists the main search engines used.

Search engines used for systematic literature review		
Search engine	Type of results	Link
Web Of Knowledge (WOK)	Articles, papers and patents	http://www.accesowok.fecyt.es/
ACM Digital library	Articles and papers	http://portal.acm.org/
Google	Webs, grey literature, patents, commercial products, related work and projects	http://www.google.com
Bibliotècnica/UPC Catalog	Articles and books	http://bibliotecnica.upc.edu/
UPC Commons	Articles and thesis	http://upcommons.upc.edu/

Table 6. Search engines used for systematic literature review.

The selection criteria was to find similar experiences, although they would be similar in only one point, for instance to be related to the field of distance music education but not in the graduate course, or be a virtual classroom for higher education in medicine. In the methods and theoretical corpus matter certain treads have been previously selected, like blended learning or constructivist approaches and refined selecting best authors or those who matched the kind of experiment we have developed.

In time, some authors where standing out as part of the main reference for this work, so some specific searches have been asked for following up the work of such authors. Those authors are: Yin and Kitchenham for the case study methodology; Bevan for the usability methods; Purao for innovation and design as research, Heritage for the conversational analysis; Vygotsky, Merrill and Jonassen for constructivist approaches, and lastly Bañó, Vaccaj, Mitchell and Gaunt for the music and singing education matters.

Not all the articles have been collected at the beginning of the study, because as told the literature review has been continuous, therefore is difficult to provide exact numbers on time, because even they have been recorded in a database of EndNote (Web of Knowledge, 2012), sometimes such record in the database was done time after the source have been included in the study. However, using the export function of the EndNote we may found next numbers about the sources.

Next table details the taxonomy on source's fields. There are many articles that should be classified in more than one class, but for simplicity only one has been chosen and codified.

Field domain of references	**Number of references**
Telepresence, Multimedia & Streaming	37
Singing & Education	17
Music Education	54
Methodology	60
Distance Education	32
Acoustical	6
Total	***207***

Table 7. Number of references by field domain.

From these 207 sources studied and included in the database, only 122 were actually included in this PhD Thesis. Newest source was included in 2012, and the older one is the Vaccaj method of singing, which is reedited in 1990 but actually was written on 1833. 24 from these sources are books, 55 are articles, 40 are web pages and 3 are other kind of sources.

Another natural source of our literature corpus is the doctorate program at the Multimedia Laboratory, which have provided main authors for this study in two ways: by receiving some lecture materials as part of regular courses, or by reading thesis from other students that have already accomplished the grade.

4. Empiric work

The prototypes or live masterclasses delivered were divided in two large stages:

- the first pilot stage, that allowed innovation exercises for the generation of the *best-of-breed* in the environment, as well as the tunning of the case study constructs and measurement methods, and
- the case study itself, that delivered the 14 international masterclasses at professional level

In this chapter we shall review the prototypes in such order; however, some of the descriptions that are used for the pilot stage remain valid for the case study, for instance the final tools and material configuration are exactly the same, actually the last configuration of the pilot stage was a merely quality assurance test the day previous to the case study execution.

A final note for this section is to stress that if we are including the results from the pilot stage in this section is *precisely* for not mixing those results with the case study results properly speaking, that are the ones treated in the "Results" chapter. Such results are briefly discussed only in order to document the work developed for estimate and test the prototypes at the pilot stage, as case study methodology requires.

4.1. The pilot stage

The pilot stage conducted six designed, planned and full-documented prototypes as huge experiments including large amounts of people and resources. However, a lot of specific experiments have been developed among such six prototypes, for try and test different configurations or devices. For simplicity, we shall review in detail those six prototypes.

4.1.1. Goals for each prototype

The following table shows the goals, date and place, and testing purposes for each of the prototypes at the pilot stage.

Table 8. Prototypes' description for the pilot stage of Opera eLearning.

Number, goal and date	**Students' side: city, room and team**	**Teacher's side: city, room and team**	**Statistics, scenarios and "hypotheses"**
I – Preliminary test 07/24/08	Castelldefels Large classroom 1 professional singer	Barcelona Large classroom 1 professional musician	Preliminary technical test and setup Transmission and A/V conversion adjustments "The transmission can be delivered"
II – First sound evaluation 07/28/08	Cornellà Multipurpose auditorium 2 professional singers	Castelldefells Large classroom 1 singing teacher Evaluation commission	Scenarios A1, A2, B1, B2, obtained from testing different microphones and speakers 3 experts and the teacher answered a questionnaire and evaluated the results 24 questionnaire entries "Professional condenser microphones improve the quality of the transmitted sound, and it is better to have separate speakers"
III – Stereo and Dolby 5.0 simulation 12/12/08	Valencia Music auditorium 4 professional singers 1 piano accompanist Evaluation commission	Sabadell Large classroom 1 singing teacher Evaluation commission	Scenarios E1, E2, D1, D2, obtained from testing: Stereo VS. Dolby 5.0 simulation speakers configuration Headphones equalization, or not 10 experts, the teacher and the students answered a questionnaire and evaluated the results 72 questionnaire entries "Dolby 5.0 simulation sound may improve teacher's listening" and "students could use headphones for further equalization and manipulation of the sound, in order to simulate theatre acoustic"
IV – HD and Dolby 5.1 simulation 02/12/09	Castelldefels Large classroom 1 professional singer 1 guitar accompanist	Barcelona Multipurpose auditorium 1 professional musician	General testing of the model, including technical details on Dolby 5.1 simulation and High Definition Video transmission using different software. "High Definition Video and the subwoofer simulation of the Dolby 5.1 simulation may improve remote transmission quality"
V – Acoustic	Castelldefels	Castelldefels	Technical testing on an acoustic echo cancellation device.

Echo Cancellation 02/28/09	Large classroom 1 professional singer 1 guitar accompanist	Large classroom 1 professional musician	"The acoustic echo canceller may solve feedback and echo issues"
VI – Quality Assurance testing for International Masterclasses 03/11/09	Sabadell Large classroom 1 professional singer 1 piano accompanist	Amsterdam TV studio 1 singing teacher 1 professional musician	Quality assurance testing in order to try out the final model to be implemented in international masterclasses delivery.

Table 8. Prototypes' description for the pilot stage of Opera eLearning.

4.1.2. Tools and materials

All telepresence classrooms were installed in single rooms, either large classrooms in a university building, auditoriums, and even a TV studio. Nevertheless, we had no control on the acoustics of the rooms, and a minimum requirement of ambient noise < 40 dB had to be previously tested at each site. For opera singing it would be preferable to have a reverberation time from 1 to 1.5 sec. (Beranek, 1962), but this was not a requirement for the rooms. The masterclass zone included the piano or guitar, lighting, one or two plasma monitors (50" minimum, HD 1920 × 1080i30 & PAL 720 x 576), speakers, and one or two video-cameras. Teachers delivered the masterclass sitting at the piano, while singers stood next to it, so each camera had to be accommodated to take a mid-shot of the teacher at the piano and a wide-shot of the singer including the piano or guitar accompanist. Video-cameras were connected to a computer (double processor at 2.40 GHz, 2 GB RAM, Windows Operating System) through the IEEE 1394 port for the video transmission.

The network engineer and the audio engineer had a specific space, including their equipment. When experts acting as evaluators or observers attended, we had a separate area for them. Within large auditoriums this was easy to organize, but difficulties arose when trying to set up the telepresence room in large classrooms.

Regarding audio and video equipment, after Prototype III only professional condenser microphones were used. The Stereo system requires two speakers, while the Dolby 5.0 simulation requires five of them, as well as the same number of microphones recording and sending signal from the students' side. It is important to underline that we did not use any Dolby compression or codification. Instead, we sent the whole audio signal to the remote site. The simulation wanted to emulate the surrounding experience obtained in fine Dolby auditoriums using five channels of recording, transmission and

reproduction of the sound. All microphone inputs entered into a professional audio-mixer at each side. The mixer sent the audio channels to a professional computer audio-card with an AD/DA (24 bit/96kHz) converter and recorder, installed on a computer (double processor at 3.16 GHz, 2 GB RAM, *Linux Operating System*) for the selected software audio transmission and reception. Since the audio signal arrives prior to the video signal, we used the mixers' delay function to coordinate sound and image by delaying the sound.

A crucial material for these experiments is the network and communication software. We chose to use GÉANT2, "the high-bandwidth, academic Internet serving Europe's research and education community"(Géant2, 2009), equivalent to the Internet2 in the U.S. Video- and audio-signals were sent separately most of the times, testing three different kinds of communication software: Ultra-videoconferencing (McGill Ultra-videoconferencing, 2009) for audio and DVTS (Digital Video Transport System, 2009) for video, and Ultragrid (SITOLA, 2009) to send video, or both signals.

Next table shows in detail the tools and materials required for the six prototypes developed at the pilot stage.

Table 9. Tools and materials per prototype.

Number, goal and date	**Students' side: microphones, speakers, display**	**Teacher's side: microphones, speakers, display**	**Communication software, video and audio resolution, AD/DA converter**	**Required bandwidth**
I – Preliminary test 07/24/08	Cardioid dynamic microphone 20 to 16.000 Hz frequency response Regular desktop computer monitor display and speakers	Tie clip microphone 80" plasma monitor with speakers 2 Separate speakers	DVTS DV 720 x 576 PAL 25fps 48kHz, 16 bits PCM 16 bit AD/DA at 48kHz from a video mixer to the video camera	30 Mbps + 1 Mbps per audio channel = 32 Mbps in each direction

II – First sound evaluation 07/28/08	Cardioid dynamic microphone 20 to 16.000 Hz frequency response Hypercardioid condenser microphone 20 to 20.000 Hz frequency response Regular desktop computer monitor display and speakers	Tie clip microphone 80" plasma monitor with speakers 2 Separate speakers	DVTS DV 720 x 576 PAL 25fps 32kHz, 16 bits PCM 16 bit AD/DA at 48kHz from a video mixer to the video camera	30 Mbps + 1 Mbps per audio channel = 32 Mbps in each direction
III – Stereo and Dolby 5.0 exploration 12/12/08	Hypercardioid condenser microphone 20 to 20.000 Hz frequency response 2 Hypercardioid condenser microphones 20 to 20.000 Hz (piano) Regular desktop computer 80" plasma monitor and 2 Separate speakers	Tie clip microphone 2 Hypercardioid condenser microphones 20 to 20.000 Hz (piano) 80" plasma monitor with speakers 5 Separate speakers	DVTS Ultra-videoconferencing DV 720 x 576 PAL 25fps 32kHz, 16 bits PCM 16 bit AD/DA at 48kHz. 5 channels	30 Mbps + 1 Mbps per audio channel = 35 Mbps in each direction
IV – HD and Dolby 5.1 testing 02/12/09	4 Hypercardioid condenser microphones 20 to 20.000 Hz frequency response 2 Hypercardioid condenser microphones 20 to 20.000 Hz (guitar)	Tie clip microphone 2 Hypercardioid condenser microphones 20 to 20.000 Hz (piano) 80" plasma monitor with speakers 5 Separate speakers 1 subwoofer	Ultragrid HD(1920 × 1080 30i) HDSDI at 990 Mbps without compression	990 Mbps without compression for video and audio
V – Acoustic Echo Cancellation 02/28/09	Hypercardioid condenser microphone 20 to 20.000 Hz frequency response 2 Hypercardioid condenser microphones 20 to 20.000 Hz (guitar)	Tie clip microphone 2 Hypercardioid condenser microphones 20 to 20.000 Hz (piano) 80" plasma	DVTS DV 720 x 576 PAL 25fps 32kHz, 16 bits PCM 16bit AD/DA at 48kHz Echo Canceller Equalization protocol with	30 Mbps + 1 Mbps per audio channel = 35 Mbps in each direction

	Regular desktop computer 80" plasma monitor display and 2 Separate speakers	monitor with speakers 2 Separate speakers	pink noise	
VI – Quality Assurance testing for International Masterclasses 03/11/09	3 Hypercardioid condenser microphones 20 to 20.000 Hz frequency response 2 Hypercardioid condenser microphones 20 to 20.000 Hz (piano) 80" plasma monitor 2 Separate speakers	Tie clip microphone 2 Hypercardioid condenser microphones 20 to 20.000 Hz (piano) 80" plasma monitor with speakers 5 Separate speakers	DVTS Ultra-videoconferencing DV 720 x 576 PAL 25fps 32kHz, 16 bits PCM 16 bit AD/DA at 48kHz 5 channels	30 Mbps + 1 Mbps per audio channel = 35 Mbps in each direction

Table 9. Tools and materials per prototype.

4.1.3. Team, activities and ideas management: iterative methods

The pilot phase deployed six full prototypes; each of them had been carefully designed, planned and evaluated, not only in order to define and control the prototype's goals, but also due to the cost involved. Additional and partial experiments were conducted as required, with some specific part of the technology tested. Regarding the staff, we had two different teams: the complete team, which involved a commission of evaluative experts, several different musicians, professional singers as students and teachers, and was used for remote classes delivery testing, applying versions of the user's satisfaction questionnaire, and getting feedback from experts; and the reduced team, with most of the technical staff included and a reduced staff of musicians, at least one singer and one accompanist musician at the students' side, and one teacher or experienced musician at the teachers' side. The second configuration was used when some technical detail required hours of testing different configurations, and was evaluated by head managers.

Having these two teams is recommended as part of the innovative project implementation:

"*– a 'central project group' gathering a restricted number of key interlocutors, responsible for project realization and management (essential competences), and*

- *An 'enlarged network' which will include experts in emergent disciplines and new competences (...) in order to widen the project prospects and to promote innovation*". (Marcandella et al., 2009, p.11)

Experts were opera singing school directors, orchestra or chorus conductors, or audio engineers. All experts have an extensive experience (over 25 years) and work nowadays at the most renowned institutions related to opera in Catalonia, Spain.

The direction team of the project was composed of a chief director, who is a professional opera singer and singing teacher, and a chief of the technical staff, which was divided into the multimedia and the network team. The multimedia team also had the task of coordinating and organizing the work to be done.

Experts where interviewed, participated in brainstorms, or delivered feedback before (for planning and design) and after (for evaluation) the assembling of each of the prototypes. Some of the experts where present throughout the entire project, while others made a specific contribution, like the Dolby expert who only reviewed the design of the implementation of surrounding sound. The user-centered design was mainly oriented to the teacher as user, in order to accomplish all the requirements for the pedagogical experience. However, the students' feedback was taken into account, and as we learn from the results, they even rejected one of teacher's proposals: the optional use of headphones.

The context of use (Niegel Bevan & Curson, 1998) is the same for all the prototypes. Having considered the user's activity and requirements, a detailed technical planning was set up, adjusting when necessary the tools and materials used, and estimating or optimizing the costs involved. When it was required, a detailed technical test was carried out with the small team of musicians. We kept specific documentation of the planning, execution and feedback from each of the prototypes. In fact, the first three prototypes served for deeper definitions, while the other two explored certain technical possibilities, or where used as quality-assurance tests. New ideas were transformed into the testing hypotheses for the next prototype

4.1.4. Results and findings from the pilot stage: defining the learning environment for the case study

The first prototype adjusted some technical parameters for transmission, and allowed the technical staff to achieve the know-how to set up the remote learning rooms. In this test we only used DVTS for transmission, and the audio was converted by the video camera itself (PCM 16bit AD/DA at 32 kHz), and then sent encapsulated as part of the video and audio signal.

4.1.4.1. The right speakers and microphones

Even though we could assume that a condenser microphone would be better than a dynamic one, we wanted to make sure that the difference was substantial even for the transmitted sound. We had a stereo signal that could be heard in two different ways: through separate speakers or through the speakers integrated to 80" plasma monitor. This test also served to prove how to get the feedback from experts, and to learn how to organize the whole process. The scenarios where:

- A1: dynamic microphone, two separate speakers.

- A2: dynamic microphone, integrated speakers.
- B1: condenser microphone, two separate speakers.
- B2: condenser microphone, integrated speakers.

Twenty-four entries of the questionnaire were recovered, including the constructs from the sound quality objectives (tones, dynamics and harmonics). Since we did not arrive to 30 entries, we do not have enough data for descriptive statistics (SPSS, 2009). The next figure shows a histogram bar of the count results for the dynamic constructs, processed in SPSS, that reveals a preference for the A1, B2 scenarios. Experts agreed on getting better dynamics (difference between piano and forte) when perceiving the sound from different speakers. This was later improved by the surround system design.

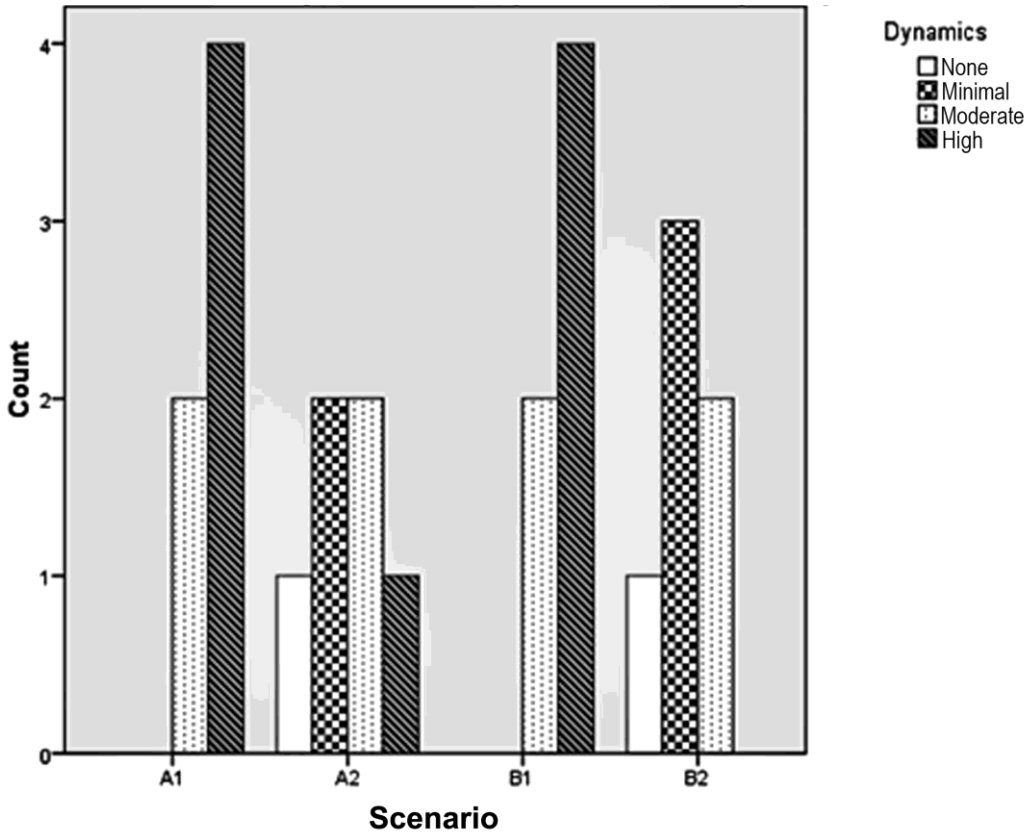

Figure 12. Dynamics count result by scenario.

When we analyze the following figure, we notice a preference for the B1 and B2 scenarios regarding the timbre construct. As expected, a better quality and warmness of the sound could be heard from the condenser microphone. Qualitative commentaries qualified condenser microphone sound as "cleaner" or with "better focus", while separate speakers delivered a sound "with more harmonics", against the "lifeless" sound of the integrated speakers.

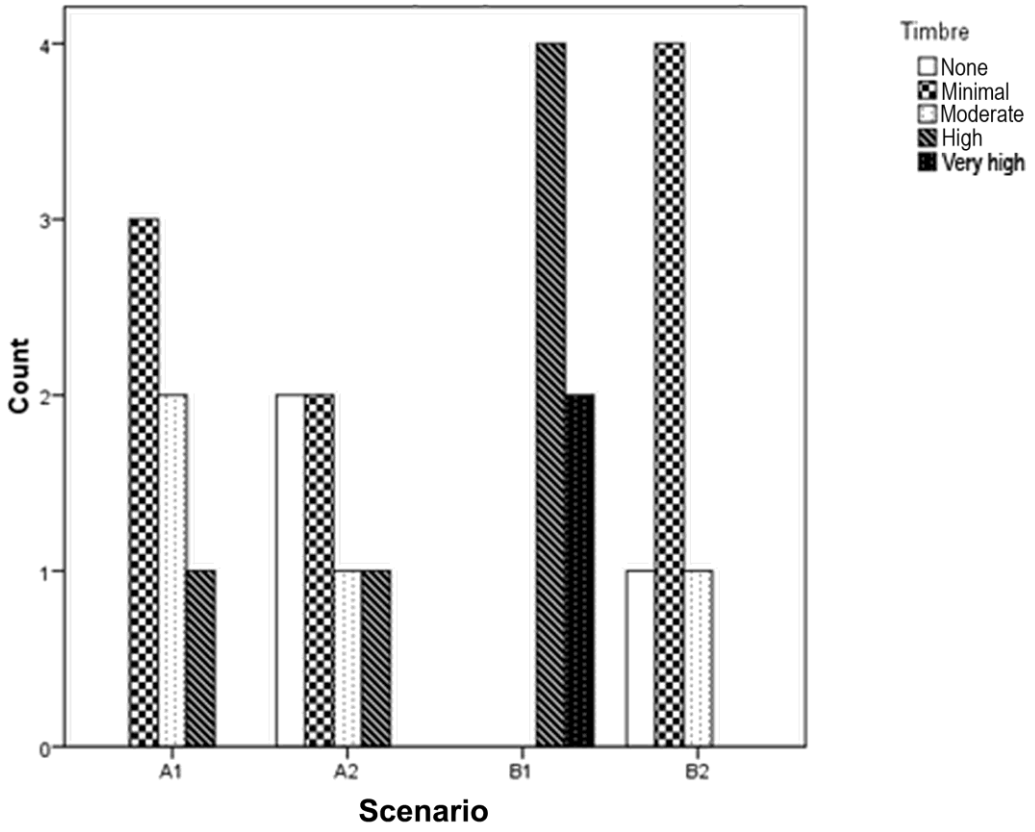

Figure 13. Timbre count result by scenario.

4.1.4.2. The Dolby 5.0 simulation in space – a substantial improvement

Before developing prototype III, we conducted reviews of the artifacts' design with two experts: the chief audio engineer of the Gran Teatre del Liceu, whose work team

transmits online some of the Opera performances at the theatre, and also produces DVD recorded material of high-definition standards for sales purposes; and the Dolby 5.0 expert, who supervises films' audio versions and cinema's installations for the Spanish market. Both of them agreed about the superiority of our new design, which included more speakers surrounding the room, and thus emulating the physical position of microphones distributed in the remote learning classroom, and both of the experts included some additional recommendation.

However, our task was to provide a minimal model that could optimize the available resources. Singing professors needed to evaluate the sound, and they were likely to be in a better position to discern little nuances in sound (Parbery-Clark et al., 2009). On the other hand, students were more likely to be devoted to their own performances, attending to observations while performing. Therefore, the decision taken was to provide the Dolby 5.0 simulation for the teacher's listening, leaving the audio for students in stereo mode.

Another possible improvement was suggested by the teaching staff: that we should use headphones at the students' side in order to try to reproduce the equalization heard from the singer in a theatre. This idea came from the need to train students for the real situation on stage, were most of the times they cannot hear themselves as clearly as they can in small rooms. This usually causes confusion for new singers at their first professional auditions.

The scenarios were:

- E1: non equalization, stereo listening.
- E2: headphones with equalization, stereo listening.
- D1: non equalization, Dolby 5.0 simulation listening.
- D2: headphones with equalization, Dolby 5.0 simulation listening.

For this prototype, 72 questionnaire entries were recovered; the teachers' side agreed on the sound improvement with the Dolby 5.0 simulation, as we can confirm with the following pie chart, where 50% of the answers qualified the amount of improvement as "high", and 8,33% qualified it as "very high". Our best choice on descriptive statistics would be range and mode (Mogey, 2009), that are 4 and 3,25. Qualitative fields distinguished the Dolby 5.0 simulation as a "meaningful contribution to improving the sound" and a "better timbre", although some of them complained because the piano and the voice where not properly balanced.

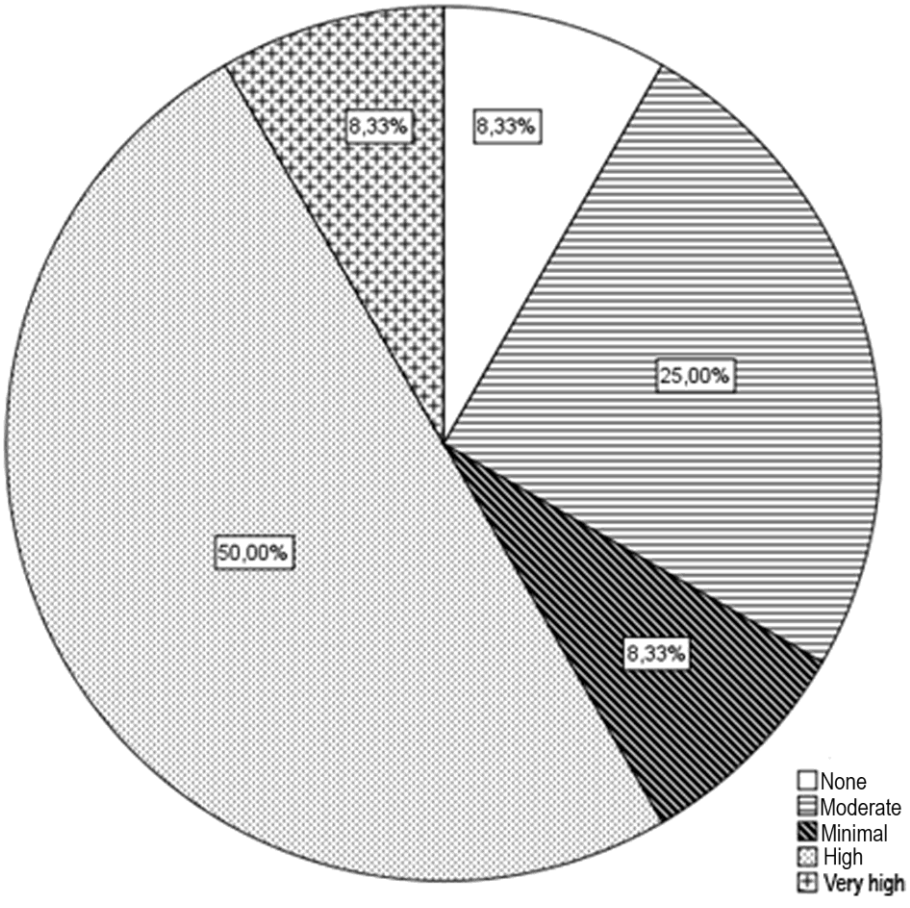

Figure 14. Percentages of frequency for the DolbyVSStero construct.

On the other hand, the use of headphones at the students' side was not accepted at all by the singers; we can see on the next figure that 61,54% of the students' side entries qualified the usefulness of the headphones as "none" or "a little". Here the range is 2 and the mode is 2,15, clearly tending to "a little". In the qualitative study of the comments from students, we find a lot of complaints: one of them put the problem in singing technical language: "when I sing I listen to the external acoustic and to my proprioceptive sensations, inside my head...headphones blocked my proprioceptive sensations, denying me the use of my previous knowledge in singing...". An interesting result is that a professor found that the students got better results when using the headphones.

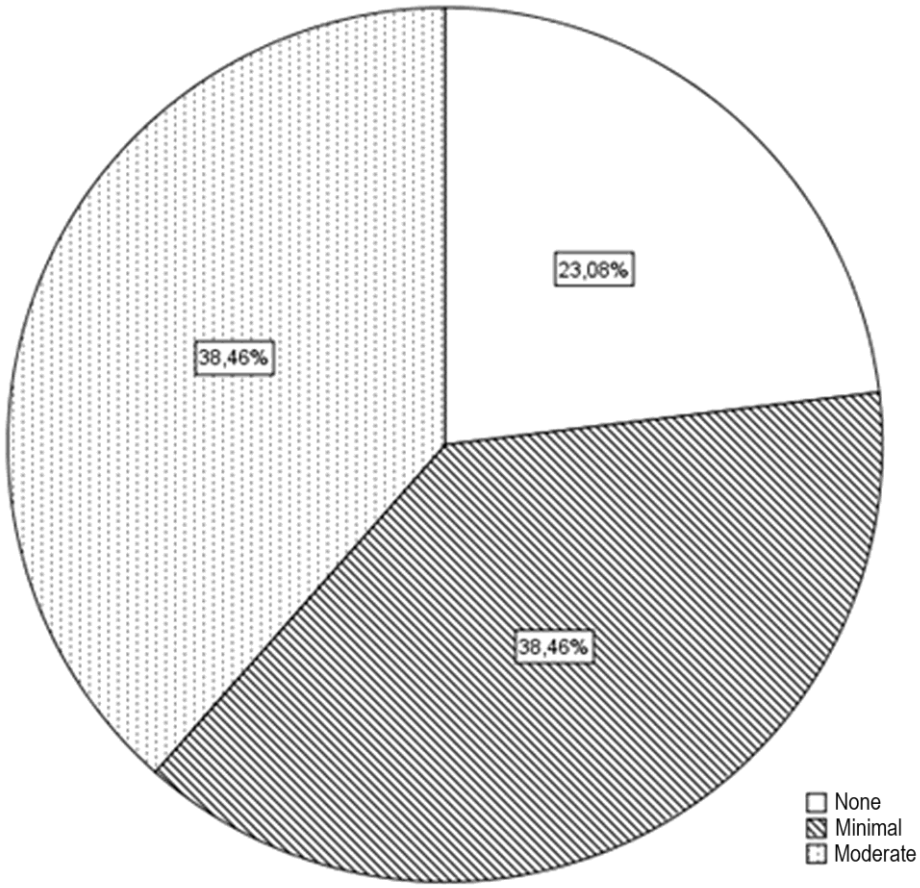

Figure 15. Percentages of frequency for the Equalization VS Non Equalization construct.

4.1.4.3. The subwoofer of Dolby 5.1 simulation and the HD for video: unnecessary expending

The next time that the multidisciplinary management team got together to design the following steps, the teachers' excitement with the possibilities of technology was all too evident. Was it possible to recreate a Dolby 5.1 simulation? Would the sound quality improve? We decided to review such matters, including the high-definition video, with the reduced staff so that we could have more time to try and retry some technical

configurations of the mixer table, or the video transmission settings, or with the microphones' and speakers' physical position. No questionnaire entries were recovered, and the result was evaluated by experienced musicians at each side of the transmission.

Regarding the use of the subwoofer, no substantial improvement could be heard by the professional musicians on the sound perceived, not even when changing the position of the corresponding transmitting microphone and speaker. With the HD video the result was also negative: the transmission got compromised, and although we were supposed to have 1GB/sec. available, and estimated to need 990 Mbps./sec., we had to drop some frames from the video so that the sound would not suffer cuts.

4.1.4.4. Acoustic echo cancellation: no conclusive results

Regarding the sound obtained, the Dolby 5.0 simulation increased a common complication in remote transmissions: the feedback echo, i.e., the echo caused by the sound emitted by the speakers that enters into the microphone during the transmission, and is retransmitted, so that it is heard as a second signal at the remote site. This is not an issue when working in an auditorium or large rooms, but in a small room it is very likely to become large signal duplication. When we used the stereo configuration, it was easier to prevent this echo, by situating both speakers at some distance from the piano microphone. The tie microphone from the professor does not cause much of this effect. However, the Dolby simulation required that the microphones and speaker would be surrounding the whole room, so it was impossible to prevent the signal from being introduced into the microphone.

Two solutions are usually used for this problem in videoconferences: the acoustic echo canceller (AEC) devices and the push-to-talk solutions. The latter consists in opening and closing the microphone, usually with a button, or in the case of our artifact when

the audio engineer controlled the volume at the mixer table. In this prototype we got a fine AEC to test the first solution, and tried several configurations of the settings for the AEC, mixing console and transmission properties. However, we could not find a configuration that would meet our requirements; these solutions are used mostly for speaking, not singing, and sometimes for our purposes we found the sound slightly changed, affecting the timbre, for instance. It must be said that we could not find a working configuration, but it may exist; as we had to decide whether to include this device or not, we finally chose not to.

4.2. The case study

Next pages describe the process and data from the case study; it was a project sustained by the Sabadell's City Council as part of the City of the Music. Sabadell is a city nearby Barcelona that has a long history as part of the Opera circuits and festivals. The City Council and the Opera School from the City of Music are looking forward to have available a high bandwidth enough to start a formal master degree supported by this telepresence artifact.

4.2.1. Who, when, how

"*Opera eLearning project included a case study focused on remote learning of lyric singing, which has delivered 14 masterclasses, 4 of them between cities in Spain and 10 between Sabadell (nearby Barcelona), and Amsterdam. As part of the project, prior to this masterclasses several pilot tests has been developed inside Spain in order to define the proper audio and video settings. A few lessons learned from the pilot stage were the definition of a minimum standard for the microphones and speakers, to dismiss the possibility of using headphones, to confirm the absolute requirement of use a piano accompanist at student´s site due the communications delay, and how the immersive*

sound emulating the space on the rooms could be optionally used to improve the listening of the teacher, although stereo configuration allows to deliver the class as well. The system is intended to be used soon in a opera singing graduate course, and also includes a Web Platform with the common tools (forums, documents, messages) that won´t be discussed beyond in this paper, but is an important part of the whole blended learning proposal of Opera eLearning project. The recorded sessions compressed in MPEG-4 format are between the published documents for further student's revision.

At the masterclasses two teachers, ten professional opera singers as students and six pianists participated. Two kind of repertoire has been reviewed, opera and lied singing. In each session there were observers at free demand, which included students, and until five evaluators. Evaluators were opera singing school directors, orchestra or chorus directors, or audio engineers. All evaluators have wide experience (over 25 years) and works nowadays at the more recognized institutions related to opera from Catalonia, in Spain. " (Rojas-Rajs et al., 2009. p.2)

Testing profile	**Detailed roles**	**Activity**	**Output**	**Site & Date**
				Sabadell, 12/12/08
Professor	1 Soprano	Deliver the class, give feedback about the experience.	Open interviews, rating questionnaires	
Technician	2 Audio 2 Video 1 Network 2 Testing Coordination	• Record and mix • Executive production	• Recordings • Reports • Rating questionnaires	
Evaluators	1 orchestra director, pianist 1 singing professor 1 audio	Attend and express their professional opinions	Committee reviews, rating questionnaires	

	engineer			
Observers	Non planned, counted on rating statistics	None	Rating questionnaires	
				Valencia, 12/12/08
Student	1 Tenor 2 Soprano 1 Bass	Receive the class, give feedback about the experience.	Free comments, rating questionnaires	
Pianist	1 Accompanist	Support the class, give feedback about the experience.	Free comments, rating questionnaires	
Technician	2 Audio 1 Video 1 Network 1 Testing Coordination	• Record and mix • Executive production	• Recordings • Reports • Rating questionnaires	
Evaluators	1 singing professor 1 chorus director 1 music professor	Attend and express their professional opinions	Committee reviews, rating questionnaires	
Observers	Non planned, counted on rating statistics	None	Rating questionnaires	
				Sabadell, 25 & 26/03/09
Professor	1 pianist, Lied professor	Deliver the class, give feedback about the experience.	Open interviews, rating questionnaires	
Technician	2 Audio 2 Video 1 Network 2 Testing Coordination	• Record and mix • Executive production	• Recordings • Reports • Rating questionnaires	
Evaluators	1 singing professor 1 chorus director 1 singing school director 1 audio opera house responsible 1 audio engineer	Attend and express their professional opinions	Committee reviews, rating questionnaires	
Observers	Non planned, counted on rating statistics	None	Rating questionnaires	
				Amsterdam, 25 & 26/03/09
Student	4 Soprano 2 Tenor	Receive the class, give feedback about the experience.	Free comments, rating questionnaires	

Pianist	4 piano students 1 piano professor	Support the class, give feedback about the experience.	Free comments, rating questionnaires	
Technician	2 Audio 2 Video 1 Network 2 Testing Coordination	• Record and mix • Executive production	• Recordings • Reports • Rating questionnaires	
Evaluators	1 singing professor 1 orchestra director 1 singing school director	Attend and express their professional opinions	Committee reviews, rating questionnaires	
Observers	Non planned, counted on rating statistics	None	Rating questionnaires	

Table 10. Testing profiles and detailed roles.

4.2.2. Tools and materials

Next slide shows a brief schema of the tools and materials used for the masterclass telepresence environment. Please keep in mind that the figure shows a Dolby microphone, but actually we used five microphones distributed in the room, besides other microphones for the pianist. Sound from the microphones goes to the mixer, which sends a line for the professional audio recorder and another signal to the analog-digital converter, whose audio signal as well as the video camera video signal are sent in separate channels to a computer that communicates with the remote site.

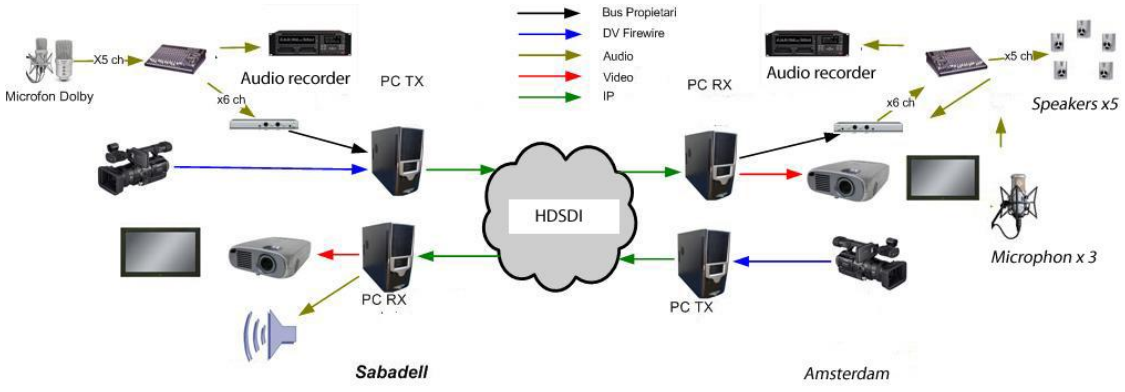

Figure 16. Audio and video transmission schema.

All telepresence classrooms were installed in single rooms, either large classrooms in a university building, auditoriums, and even a TV studio. Nevertheless, we had no control on the acoustics of the rooms, and a minimum requirement of ambient noise < 40 dB had to be previously tested at each site. For opera singing it would be preferable to have a reverberation time from 1 to 1.5 sec. (Beranek, 1962), but this was not a requirement for the rooms. The masterclass zone included the piano, lighting, one or two plasma monitors (50" minimum, HD 1920 × 1080i30 & PAL 720 x 576), speakers, and one or two video-cameras. Teachers delivered the masterclass sitting at the piano, while singers stood next to it, so each camera had to be accommodated to take a mid-shot of the teacher at the piano and a wide-shot of the singer including the piano or guitar accompanist. Video-cameras were connected to a computer (double processor at 2.40 GHz, 2 GB RAM, Windows Operating System) through the IEEE 1394 port for the video transmission.

The network engineer and the audio engineer had a specific space, including their equipment. When experts acting as evaluators or observers attended, we had a separate area for them. Within large auditoriums this was easy to organize, but difficulties arose when trying to set up the telepresence room in large classrooms. The next figure shows the implementation of the student's side for the international masterclass. Detailed pictures of the equipment used in this map may be reviewed at Appendix B.

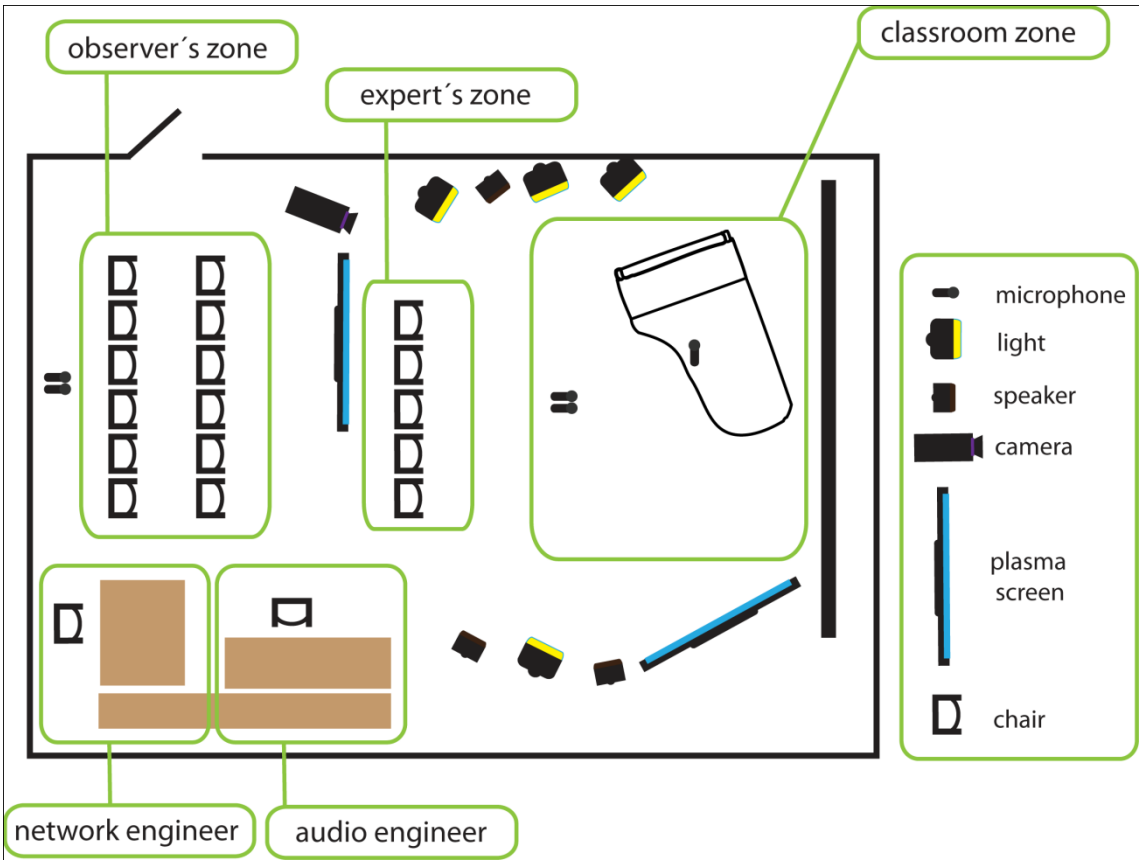

Figure 17. Map of the Prototype VI classroom at student's side.

Regarding audio and video equipment, only professional condenser microphones were used. The Stereo system requires two speakers, while the Dolby 5.0 simulation requires five of them, as well as the same number of microphones recording and sending signal from the students' side. It is important to underline that we did not use any Dolby compression or codification. Instead, we sent the whole audio signal to the remote site.

The simulation wanted to emulate the surrounding experience obtained in fine Dolby auditoriums using five channels of recording, transmission and reproduction of the sound. All microphone inputs entered into a professional audio-mixer at each side. The mixer sent the audio channels to a professional computer audio-card with an AD/DA (24 bit/96kHz) converter and recorder, installed on a computer (double processor at 3.16 GHz, 2 GB RAM, Linux Operating System) for the selected software audio

transmission and reception. Since the audio signal arrives prior to the video signal, we used the mixers' delay function to coordinate sound and image by delaying the sound.

Next tables show a brief description and then the following table describes with more detail the materials used.

Tools and materials for the case study Opera e-LEARNING (Quick view)				
Number, goal and date	**Students' side: microphones, speakers, display**	**Teacher's side: microphones, speakers, display**	**Communication software, video and audio resolution, AD/DA converter**	**Required bandwidth**
National and International Masterclasses 03/11/09	3 Hypercardioid condenser microphones 20 to 20.000 Hz frequency response 2 Hypercardioid condenser microphones 20 to 20.000 Hz (piano) 80" plasma monitor 2 Separate speakers	Tie clip microphone 2 Hypercardioid condenser microphones 20 to 20.000 Hz (piano) 80" plasma monitor with speakers 5 Separate speakers	DVTS Ultra-videoconferencing DV 720 x 576 PAL 25fps 32kHz, 16 bits PCM 16 bit AD/DA at 48kHz 5 channels	30 Mbps + 1 Mbps per audio channel = 35 Mbps in each direction

Table 11. Tools and materials for the case study Opera e-LEARNING (Quick view)

Detailed audio materials for the case study Opera e-Learning		
		Observations
Microphones	Professor, pianist (tie clip) • Omni directional • *Frequency response* 20 Hz - 20 kHz • *Dynamic Range* 144 dB • *S/N ratio re 1 Pa (A-weighted*) • S/N ratio, re. 1 kHz at 1 Pa (94 dB SPL): 68 dB(A) • *Equivalent noise level (A-*	Microphone's distribution at the space in the recording room allows obtaining meaningful results in the spatial sensation at the reproduction site. Emissions room at Sabadell was too small.

	weighted) 26 dB (DIN 45412) Singer • Hypercardioid • *Frequency response* 20 Hz - 20 kHz • *Dynamic Range* 130 dB • *Equivalent noise level A-weighted)* 11 dB (DIN 45412) • S/N ratio, re. 1 kHz at 1 Pa 83 dB(A) PIANO • Condenser, Omnidirectional • *Frequency response* 18 Hz - 20 kHz • *Dynamic Range* 141 dB	
Speakers	• Auto amplified acoustic boxes • *Frequency Range* 60Hz to 55kHz ±3dB • *Sensitivity pink noise* 115 dB-SPL @ 1W, 1m	Pink noise adjustment allows improving sound reproduction at the reception's room.
Audio digital sound mixer	• Mic Input (Head Amp): 16 (XLR) • Line Input: 4 OMNI in • Outputs: 12 OMNI Out • Digital Inputs:2TR IN DIGITAL (AES/EBU, IEC-60958) • Digital Outputs:2TR OUT DIGITAL (AES/EBU, IEC-60958) • BUS: 8 mix buses, 8 AUX, Main ST Bus	
Noises		There are unknown noises when transmission rates are close to the channel limit. Saturation should be avoided in the speakers in order to

		allow the proper dynamic of singing voices.
Acoustic Echo Canceller	Acoustic Echo Canceller (AEC) Si-400 Symphonix	Testing results are no conclusive for the stereo configuration. At the multichannel Dolby 5.0 simulation configuration there is no improvement shown.

Table 12. Detailed audio materials for the case study Opera e-Learning

A crucial material for the transmission is the network; each of the transmission software used has requirements on the jitter (is the undesired deviation from true periodicity of an assumed periodic signal in electronics and telecommunications, often in relation to a reference clock source, in this case the variation of time between IP packages), delays and losses. Using uncompressed formats implies that any alteration may lead to severe issues in communication, that we have managed first that all removing frames from the video signal, and giving priority to the sound signal. The recommendation is to have available a controlled network from point to point of level 2, where there are no routers (level 3 equipments) involved.

High bandwidth required depends on the application used, and is showed in next table.

Transmissions software	**Required bandwidth**	**Recommended bandwidth**
DVTS	30Mbps	40Mbps
DVTS + Ultravideoconferencing	42Mbps	50Mbps
Ultragrid	990Mbps	1Gbps
Ultragrid + Ultravideoconferencing	990Mbps	1Gbps

Table 13. Required bandwidth by transmission software.

Our testing labor under networks of level 2 has been transparent and with no issues at all. Unfortunately, connection to Amsterdam was using a shared environment of level 3

at some point of the communication, after the networks that we could control at level 2 (i2cat in Spain and Surfnet in Netherlands). That left us in an awkward position when we start to have communication cuts at the masterclasses on March 26th. Previous tests has shown a stable network, that with a 90 Mbps has 53ms of delay, 0.12 ms of jitter (almost neglectable) and a losses percentage of 0.1% (is meaningful since 1% at the audiovisual level). Such values where stables on March 11th at the quality assurance tests, and March 25th for the first set of masterclasses, but the 26th we got several cuts, that are shown in the next figure.

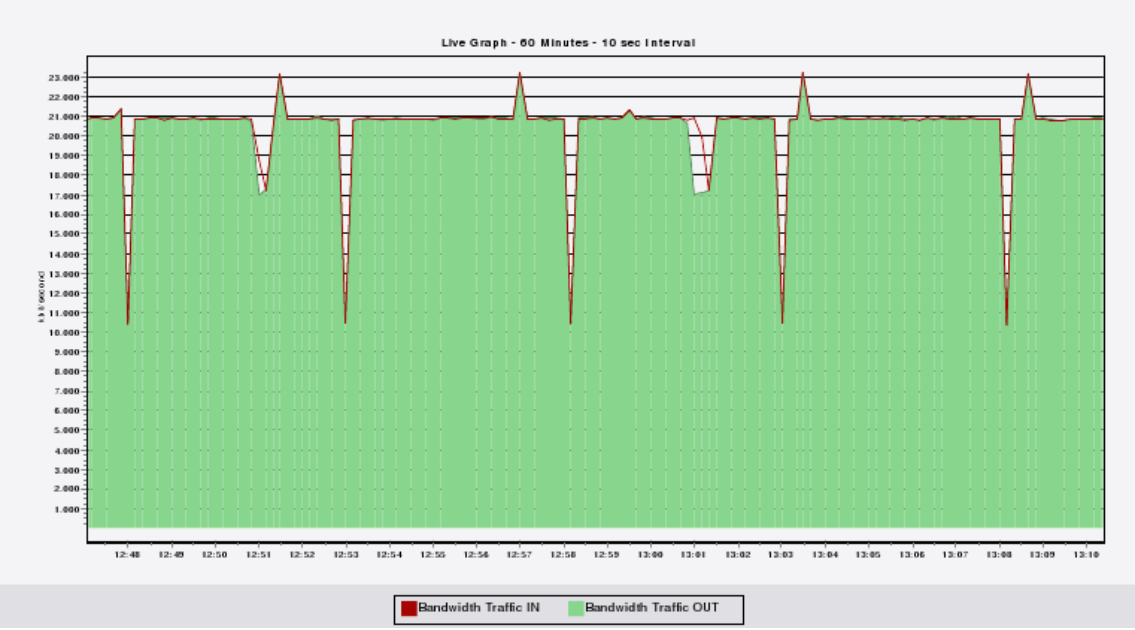

Figure 18. Communication cut-offs statistics on March 26th of 2009 between Amsterdam and Sabadell.

That has been so far the only issue with the network experienced; therefore the conclusion is not to use level 3 networks, although this increases the expenses.

4.2.3. Team and activities

Besides the rough description of activities given in the "testing profiles and detailed roles table" (Table 10), we now shall review the detailed activities and outputs generated, as well as the participation of the institutions involved on each team. In the

technician area we also have to include the hired personal, and will be distinguished as external with the name of the company or institution in parenthesis. In terms of the job developed, the teams that appear grouped as "technicians" in the Table 10 now will be split in different specialties that have a good amount of job for each one. The previous testing mentioned used to long at least 4 hours before the beginning of the masterclass, and for the international masterclass we also had a previous test the day before.

Team	**Institution**	**Activity**	**Output**	**Site & Date**
				Sabadell, 12/12/08
Professor	City of the Music	Deliver the class, give feedback about the experience.	Open interviews, recorded sessions, rating questionnaires	
Evaluators	City of the Music and guests	Attend and express their professional opinions	Committee reviews, rating questionnaires	
Audio engineering	Extern (Sonostudi[5])	• Previous testing • Record all the channels separately in DAT format • Mix the channels emitted and recorded a mixed version • Adjust volume on remote signal received, or increase or decrease volume on professor's microphone avoiding echoes and feedback	• Separate channels recordings • Mixed recording • Channels description list	
Video recording	LAM	• Previous testing • Record all the sessions, including audio • Send input to network engineer	• Recordings • Recording cassettes description list	
Network	I2cat	• Previous testing • Receive audio and video signal, execute A/D conversion and send it by the transmission software • Measure network status • Adjust transmission parameters if necessary	• Network status statistics • Record of transmission adjustments	

[5] Sonostudi is a well known audio provider in Barcelona www.sonostudi.com

Coordination	LAM City of the Music	• Host the session • Apply rating questionnaires • Apply open interview • Register assistance, general performance matters and comments	• Compilation of rating questionnaires • Official report on session	
Multimedia	LAM	• Compile recorded sessions for Web Site after the session	• Online videos	
				Valencia, 12/12/08
Students	City of the Music	Receive the class, give feedback about the experience.	Free comments, rating questionnaires	
Evaluators	City of the Music and guests	Attend and express their professional opinions	Committee reviews, rating questionnaires	
Audio engineering	Extern (Sonostudi)	• Previous testing • Record all the channels separately in DAT format • Mix the channels emitted and recorded a mixed version • Adjust volume on remote signal received, or increase or decrease volume on professor's microphone avoiding echoes and feedback	• Separate channels recordings • Mixed recording • Channels description list	
Video recording	I2cat	• Previous testing • Record all the sessions, including audio • Send input to network engineer	• Recordings • Recording cassettes description list	
Network	I2cat	• Previous testing • Receive audio and video signal, execute A/D conversion and send it by the transmission software • Measure network status • Adjust transmission parameters if necessary	• Network status statistics • Record of transmission adjustments	
Coordination	LAM	• Obtain facilities access and coordinate auditorium tech staff • Conduct team's travel from Barcelona to Valencia • Host the session • Apply rating questionnaires • Register assistance, general performance matters and	• Compilation of rating questionnaires • Official report on session	

		comments		
Multimedia	LAM	• Compile recorded sessions for Web Site after the session	• Online videos	
				Sabadell, 25 & 26/03/09
Professor	City of the Music	Deliver the class, give feedback about the experience.	Open interviews, rating questionnaires	
Evaluators	City of the Music and guests	Attend and express their professional opinions	Committee reviews, rating questionnaires	
Audio engineering	Extern (Sonostudi)	• Previous testing • Record all the channels separately in DAT format • Mix the channels emitted and recorded a mixed version • Adjust volume on remote signal received, or increase or decrease volume on professor's microphone avoiding echoes and feedback • Support network engineer for restart audio communications when issues on network affected the masterclass	• Separate channels recordings • Mixed recording • Channels description list	
Video recording	I2cat LAM Extern (lights)	• Previous testing • Record all the sessions, including audio • Send input to network engineer	• Recordings • Recording cassettes description list	
Network	I2cat	• Previous testing • Receive audio and video signal, execute A/D conversion and send it by the transmission software • Measure network status • Adjust transmission parameters if necessary (As on 26 we had some issues, remove frameworks from video transmission, and restart audio communication when required)	• Network status statistics • Record of transmission adjustments • Special report on issues happened on March 26th	
Coordination	LAM	• Host the session, with special attention to some guests of importance, like the Major of the City • Apply rating questionnaires	• Compilation of rating questionnaires • Official report	

		• Register assistance, general performance matters and comments	on session	
Multimedia	LAM	• Compile recorded sessions for Web Site after the session	• Online videos	
				Amsterdam, 25 & 26/03/09
Students	City of the Music	Receive the class, give feedback about the experience.	Free comments, rating questionnaires	
Evaluators	City of the Music and guests	Attend and express their professional opinions	Committee reviews, rating questionnaires	
Audio engineering	Extern (Pakhuis de Zwijger[6])	• Previous testing • Record all the channels separately in DAT format • Mix the channels emitted and recorded a mixed version • Adjust volume on remote signal received in order to avoid feedback or echoes	• Separate channels recordings • Mixed recording • Channels description list	
Video recording	Extern (Pakhuis de Zwijger[7])	• Previous testing • Record all the sessions, including audio • Send input to network engineer	• Recordings • Recording cassettes description list	
Network	I2cat	• Previous testing • Receive audio and video signal, execute A/D conversion and send it by the transmission software • Measure network status • Adjust transmission parameters if necessary (As on 26 we had some issues, remove frameworks from video transmission, and restart audio communication when required)	• Network status statistics • Record of transmission adjustments • Special report on issues happened on March 26th	
Coordination	LAM City of the Music	• Host the session • Apply rating questionnaires • Register assistance, general performance matters and comments	• Compilation of rating questionnaires • Official report on session	
Multimedia	LAM	• Compile recorded sessions	• Online videos	

[6] Pakhuis de Zwijger is the cultural institution where the TV studio was loaned at Amsterdam. www.de**zwijger**.nl

[7] Pakhuis de Zwijger is the cultural institution where the TV studio was loaned at Amsterdam. www.de**zwijger**.nl

		for Web Site after the session		

Table 14. Activities and teams table.

4.2.4. The Opera eLearning Website

As told before, the Opera eLearning project has included in its design the functionalities provided from a learning social network as well as a knowledge database. Although the masterclasses were punctual with no further follow-up from the teachers, so it really didn´t was part of the scope to test or include the Web into the case study, the Web tools was constructed and putted available, mainly for the consult on the recorded video sessions, scores and other data.

For a description of most of the operations available in this Web Site, as well as for a quick review on the whole project designed in UML diagrams, please refer to Appendix C.

Next slide shows the main page of the Opera eLearning Web Site.

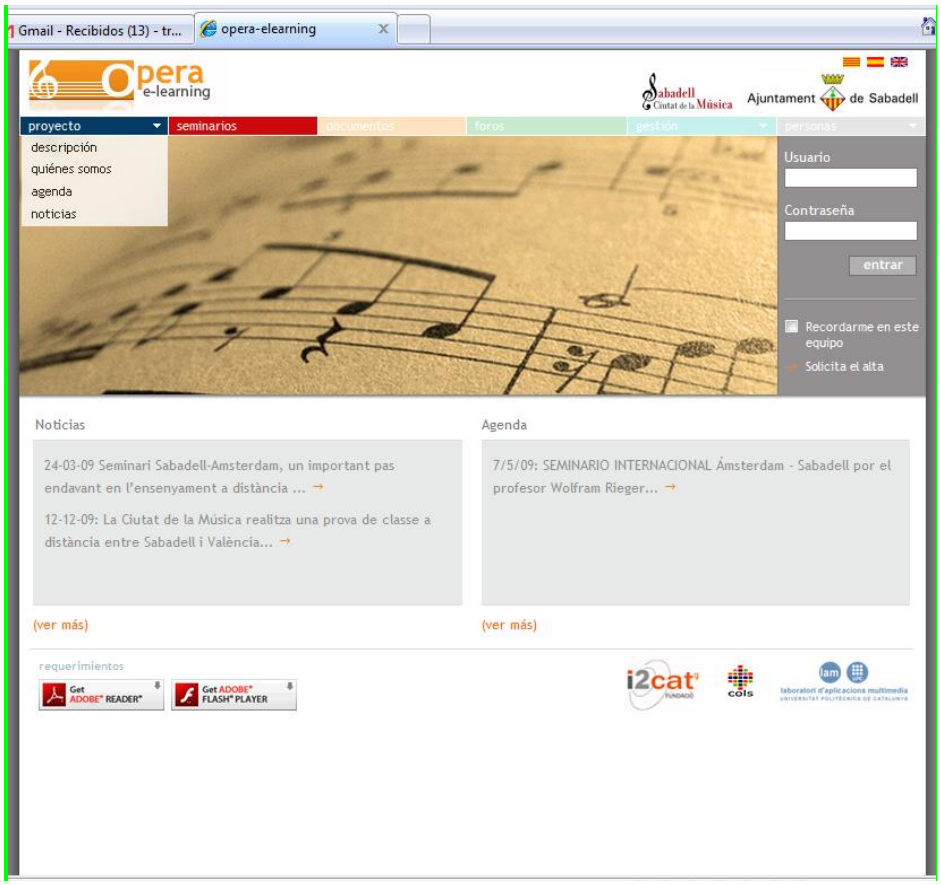

Figure 19. Main page of the Opera eLearning Web Site (www.opera-elearning.com)

The Web Site had several user´s profile: general public, students, professors, content administrators and the Web Administrator. It is allowed to publish an event linked to a calendar for the public or restricted zone, or well to describe the details of a masterclass in process, so possible students may accommodate their agenda to assist to such lessons. Students may review all the masterclasses available, including the documents associated. Teachers may publish new documents and scores, and lead forums. There is a Web Administrator that can add new forums and who is responsible to publish the encoded in MPEG4 videos of the recorded sessions.

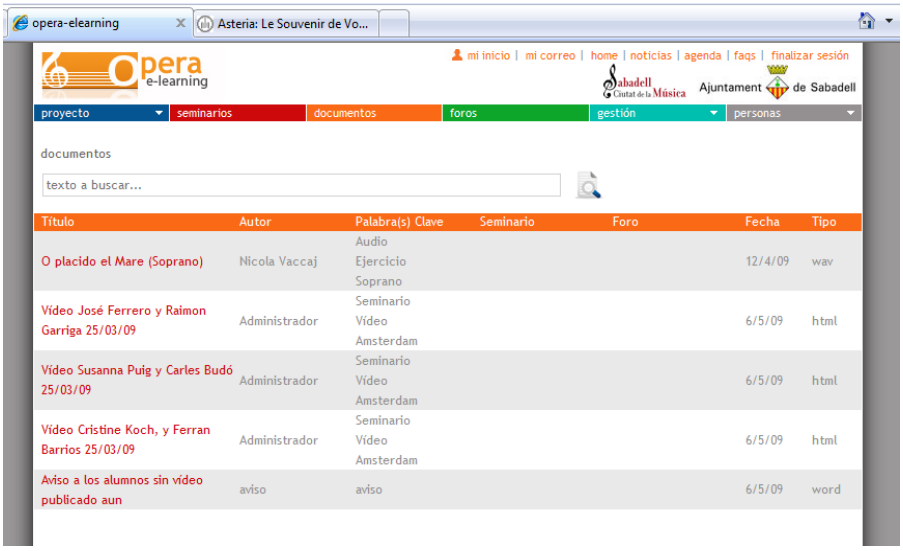

Figure 20. Documents page of the Opera eLearning Web Site (www.opera-elearning.com)

Although Web site was not tested at all and was delivered by formal reasons as part of the project, as the distance education section established is crucial as part of the whole blended learning model proposed by the Opera eLearning project.

5. Results

This chapter is devoted to describe the results obtained. It has been divided in two main parts that correspond to different techniques of approaching to results: statistic and qualitative analysis from satisfaction questionnaires and ratings, and the qualitative analysis of the recorded masterclasses. Both together points to very similar conclusions regarding the effectiveness and viability of the learning environment, and remark the issues that need work in further versions.

The hypothesis were user oriented: the model of remote singing learning room (1) was enough for the teacher to fully evaluate the student; (2) allowed the student to understand observations and emulate the teacher's sound; (3) observers could use the information received as well and (4) a high quality, immersive and space distributed sound could improve teacher's listening. The main question to be answered, at last, is prove if it was possible to deliver proper masterclass remotely with this artifact.

During all the classes we had video and audio recording; for the analysis, categories of research were defined, which are listed in the following table. Such categories can be reviewed against each of the hypothesis. The research categories defined and tuned after the pilot stage may be reviewed in the following table:

Research categories defined for Case Study		
Category	**Subcategory**	**Specific information about subcategories**
Learning objectives	Technique	• Collocation • Air Support (*Appoggio*) • Articulation and diction
	Expression	• Conducting (*piano*, *forte*, entrances, coordination) • Emotion and feeling

Sound quality objectives	Tones Dynamics Harmonics	• High, low and medium frequencies • Changes between *piano* and *forte* • Harmonics, timbre and colour
Classroom effectiveness objectives	Class support Communication fluency	• Seamless remote teaching and learning • Speaking, singing and visual communication

Table 15. Research categories defined for Case Study

Some of the subcategories appeared as part of the original goals of the case study (for instance all of the classroom effectiveness,and sound quality objectives). On the other hand, the learning objectives inherit the singing technique main characteristics that have been described in chapter 2.2.2. Briefly, these included the requirement of having a good air flow support, placing the sound in the cranium resonators (Baño, 2003) and keeping the "open throat" position (H. F. Mitchell & Kenny, 2004; Vaccaj, 1990). These central pillars of the opera technique have been mapped into the air support and tone placing constructs. Articulation and diction are of fundamental importance, not only for the public's proper understanding of the sung lyrics –something which sometimes is not at all possible in opera – but also as a technical means of obtaining the right mouth- and throat-position that lead to the specific tone and color (Baño, 2003). A passionate, deep song in the German language sounds different from, and obtains another texture compared to, a lighter and less heavily pronounced song in the Italian language. In fact, this is the reason why in traditional opera technique teaching books the text for students was written in Italian (Vaccaj, 1990): consonants with strong pronunciation (like / *r* /) tend to stop the air flow. Achieving the right pronunciation of a foreign language is one of the greatest challenges for a singer. This is the reason why we included an Articulation and diction construct.

Finally, emotion, feeling and conducting are subcategories that relate to the higher education level (Karlsson & Juslin, 2008), but they are remarkable characteristics in the testing of the transmission and the conception of the Opera eLearning environment as a telepresence artifact; in other words, with these constructs we can test the ultimate ability to communicate something complex: human emotions.

The questionnaire was administered to teachers, pianists, evaluators and students as part of their participation. Only teachers and some of the pianists were paid for their services, students attended the classes for free. Observers and technical staff present in the room were requested to answer the questionnaire on a volunteer basis.. We recovered 195 entries from all the masterclass sessions, including qualitative and quantitative data. Five point Likert-Scales were used, with 1 (one) indicating the lowest ("none") and 5 (five) indicating the highest (''very high") degree of satisfaction for each of the questions asked. The subcategories were mapped into constructs for the questionnaire. The achievement of Learning objectives was not asked with Likert-Scales to all the participants, but only commented by teachers and evaluators with free text fields and studied from the conversational analysis of the recorded sessions. All the questionnaires had a free text field for the Classroom effectiveness category, which allowed another qualitative discourse analysis. The questionnaire was filled after each masterclass when that was possible, or after a set of three classes when needed. The following pie diagram shows the composition of the universe sample in percentages.

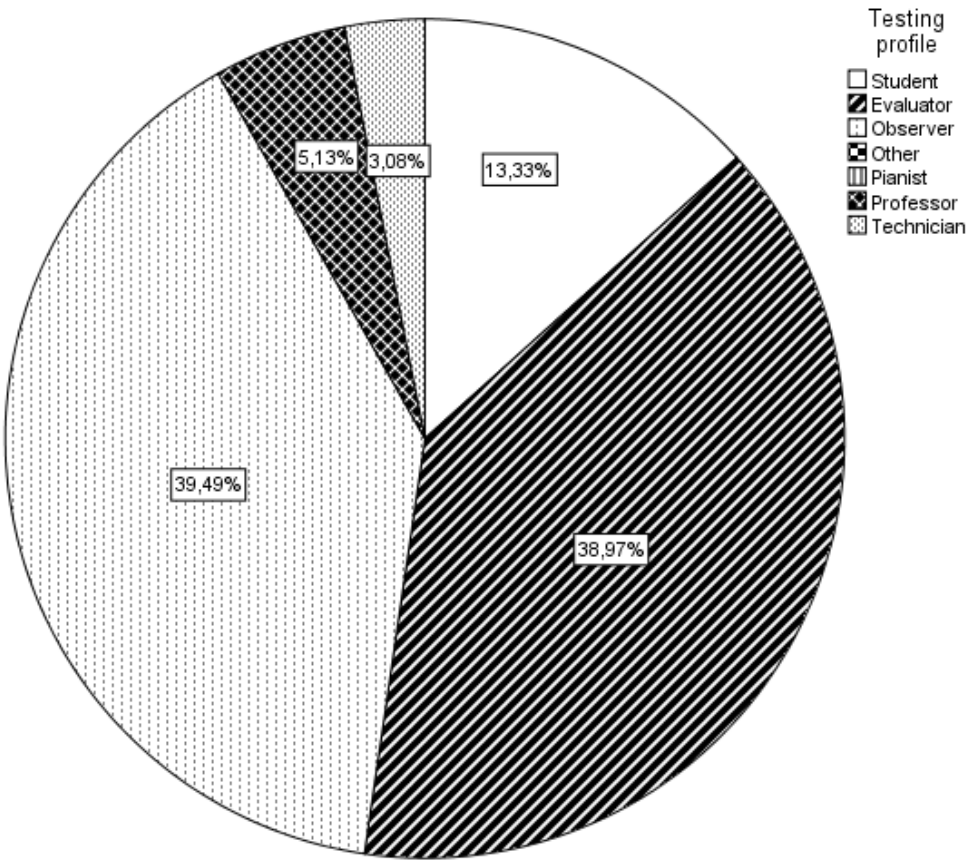

Figure 21. Percentages of testing profiles participants in rating questionnaires.

From the categories, the sound quality objectives and the classroom effectiveness have been measured by the satisfaction questionnaires, while the learning objectives have been commented by the teachers in open interviews, and are also analyzed from the qualitative perspective in the session's recordings. Therefore we applied the conversational analysis to the learning objectives, and the users experience rating methods as well as the conversational analysis to the sound quality objectives and the classroom effectiveness categories.

Given the use of the Likert scales, it is quite important to properly manage the obtained data, since: "*data collected are ordinal: they have an inherent order or sequence, but one cannot assume that the respondent means that the difference between*

agreeing and strongly agreeing is the same as between agreeing and being undecided" (Mogey, 2009). Basically, the difference with numeric data is that when applying statistical analysis we must choose those methods that give as a result one of the ordinals used in our scale, and not a value that is not included in it (for instance, in our scale, if we got a median of 2.5 we would not be able to understand that number as "half minimal and half moderate"). We can summarize "*using a median or a mode (not a mean); the mode is probably the most suitable for easy interpretation*"; while we can "*express variability in terms of the range or inter quartile range (not the standard deviation)*" (Mogey, 2009). For the graphic display of results we have found it easier to interpret the percentages in a pie diagram, which shows at a glance the percentage of frequencies, and variability. The questionnaires' data has been analyzed with SPSS analytical software. For statistical purposes, the answers given by pianists were grouped with those of the students, since they were receiving instructions from teachers and closely shared the space with singers.

For each of the categories we will first go through the results obtained from the statistics on satisfaction ratings, followed by a discussion on the qualitative research on the recorded video material. For the discourse analysis we will turn to conversational analysis, searching for traces of institutional analysis (Heritage, 2001), and exploring the characteristics of this adjacency pair of teacher-student in the masterclass environment. "*An adjacency pair is an example of turn-taking in the production of utterances where the speaking of the first utterance (first turn) provokes a responding utterance from the second turn, for instance 'Hi, how are you?' is followed by 'Fine, and you?'. The adjacency pair is the minimal pair of the conversation*."(Centro Virtual Cervantes, 2012)

In education, an institutional conversation with a very clear order priority, adjacency is predetermined, as teachers must talk most of the time and should be respected in his or her opinions.

Other examples of institutional talk where studied at the systematic literature review process in fields like medicine, religion, academic and casual talk (Barbour, 2010; Boehringer & Wolff, 2010; House, 2010). A study about the worker-boss adjacency pair shows how certain argumentative conducts and their subsequent responses preserve the institutional hierarchical relationship (Argaman, 2009). Samples found in education always show the prevalence of teacher as leader of the conversation, which is part of the type of relationship of this pair (Escudero et al., 1999; House, 2010)

Our conversation analysis will look for specific sequences that correspond to some of the research categories, and study their organization and content, giving priority to some terms or word selection as explained in the methodology chapter. For instance, we are interested in learning how the teacher corrects the singer's tone placing, and mainly if this happens fluently through the use of the telepresence artifact.

5.1. Results of the satisfaction rating questionnaires

5.1.1. Sound quality objectives

High, medium and low frequencies are the pillars of the decomposition of sound for equalization and transmission. These sound quality objectives are extremely basic, and at the same time it is very easy to give privilege to, or obtain more bright for one of these frequencies. This made necessary to review how they had been heard.

The next figures show the results obtained as a pie diagram for each of the subcategories of the different frequencies' sound quality objectives, briefly discussing the descriptive statistics, as well as adding some of the written answers in open text fields that have to do with each subcategory.

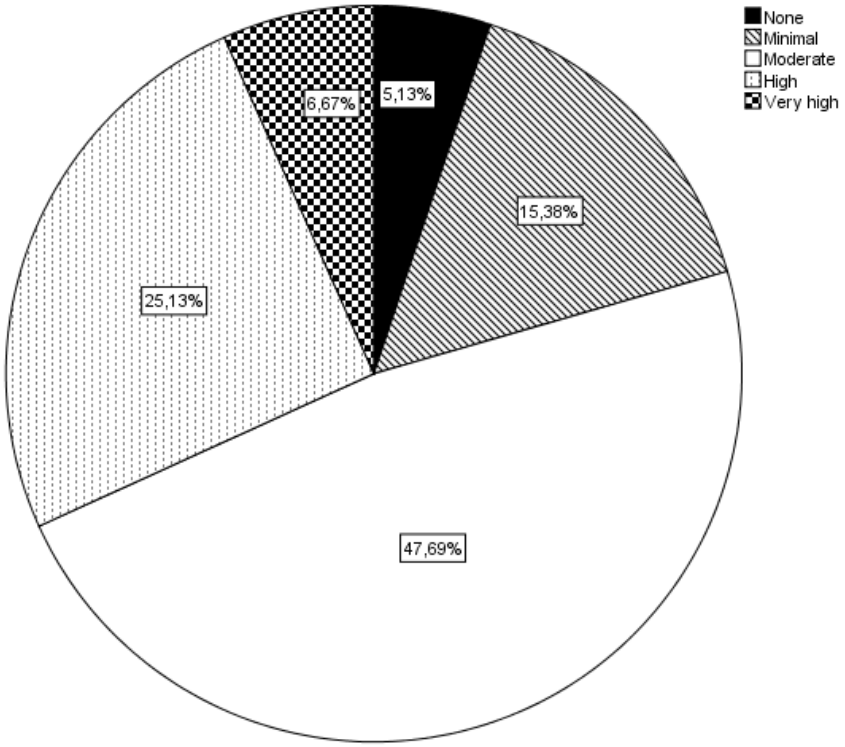

Figure 22. Percentages of frequency for the high tones construct.

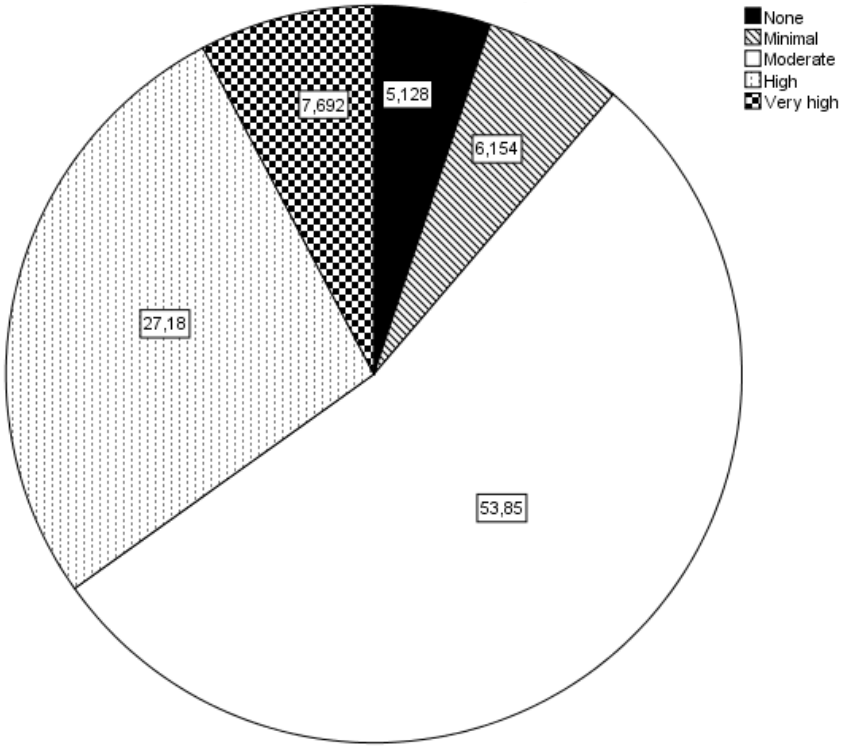

Figure 23. Percentages of frequency for the medium tones construct.

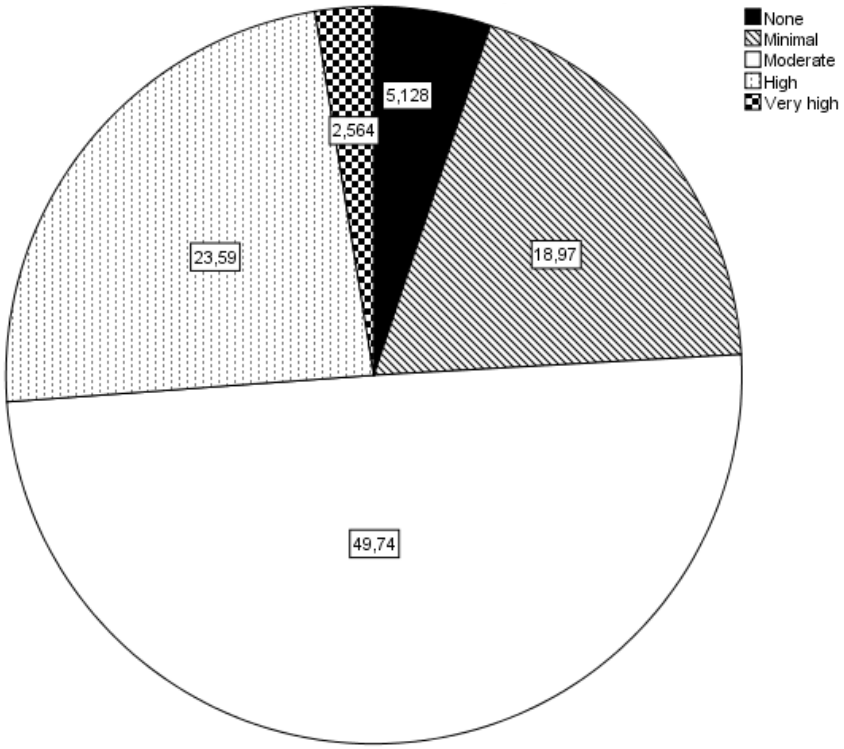

Figure 24. Percentages of frequency for the low tones construct.

For all three subcategories the median and the mode are 3, with a range of 4 that gives a measure of dispersion. A stronger manner of review dispersion is the inter quartile range (defined as the difference between the upper and lower quartiles, that is the 25% and 75% percentiles(Wikipedia, 2012)), and again the result is the same for three of the subcategories, receiving a 1, that match with the range as dispersion amount from the median of 3. 3 is the moderate value, so we may say that this construct is neither getting an excellent qualification nor is considered in bad shape.

In the open text fields we got some comments as “timbre gets cutted in the higher and medium frequencies tones”, “with the baritone is more obvious the lack of low frequencies”, “there is an artificial timbre, we got the baritone almost as a tenor”, “high tones seems short” or “low tones are at poor volume”. Several pianists and observers noticed that the pianos should be tuned at the same exact frequency with each other, since these little differences of tone are noticed by their trained ears.

A possible bias existed derived from singer’s kind of voice, that is, that we may be listening better the high frequency tones because the singer was a soprano. Unfortunately our quasi experimental sample did not have the same amount of singers from different voices, actually most of the singers were female sopranos, and we had one baritone and three tenors. Next table describes the detailed descriptive statistics by voice classification, where we find a *really lower qualification* (a range of 1 and a median of 3 for low and medium tones) for the baritone that may leads to thing that our lower tones frequencies may be improved. Next table shows the results of descriptive statistics for high, medium and low tones decomposed by the voice classification.

Voice classification		High tones	Medium tones	Low tones
Baritone	Mean	3.11	3.44	3.44
	N	9	9	9
	Standard deviation	.782	.527	.527
	Rank	2	1	1
	Median	3.00	3.00	3.00
Soprano	Mean	3.16	3.27	2.98
	N	131	131	131
	Standard deviation	.951	.901	.877
	Rank	4	4	4
	Median	3.00	3.00	3.00
Tenor	Mean	3.05	3.22	2.95
	N	55	55	55
	Standard deviation.	.911	.896	.848
	Rank	4	4	4
	Median	3.00	3.00	3.00
Total	Mean	3.13	3.26	2.99
	N	195	195	195
	Standard deviation	.930	.884	.859
	Rank	4	4	4
	Median	3.00	3.00	3.00

Table 16. High, medium and low tones descriptive statistics by voice classification

This is a line left open for further research using more singers, because it is very likely that a common equalization is not suitable for everyone. Therefore, for huge masterclasses, a previous small sound test for each singer should be arranged. In regular one-to-one closed sessions with stereo configuration, perhaps this would be unnecessary and too complex.

Tone, dynamics and harmonics are characteristics of the voice timbre and texture that may change according to the technique, emotion and feeling implied. That is the reason why it is so important for the teacher to receive full feedback on these sound qualities, while students need to receive a good example to learn from.

The following figure shows the results for the dynamics subcategory, where we have a result that is common for the three subcategories with a mode of 3, a median of 3, a rank of 3 and where the IQR is 1. These results shows a more regular outcome than in the high, medium and low tones evaluation, that is, most of the answers tend to be 3 (moderate). For each construct we have the qualitative analysis to complete the description of results.

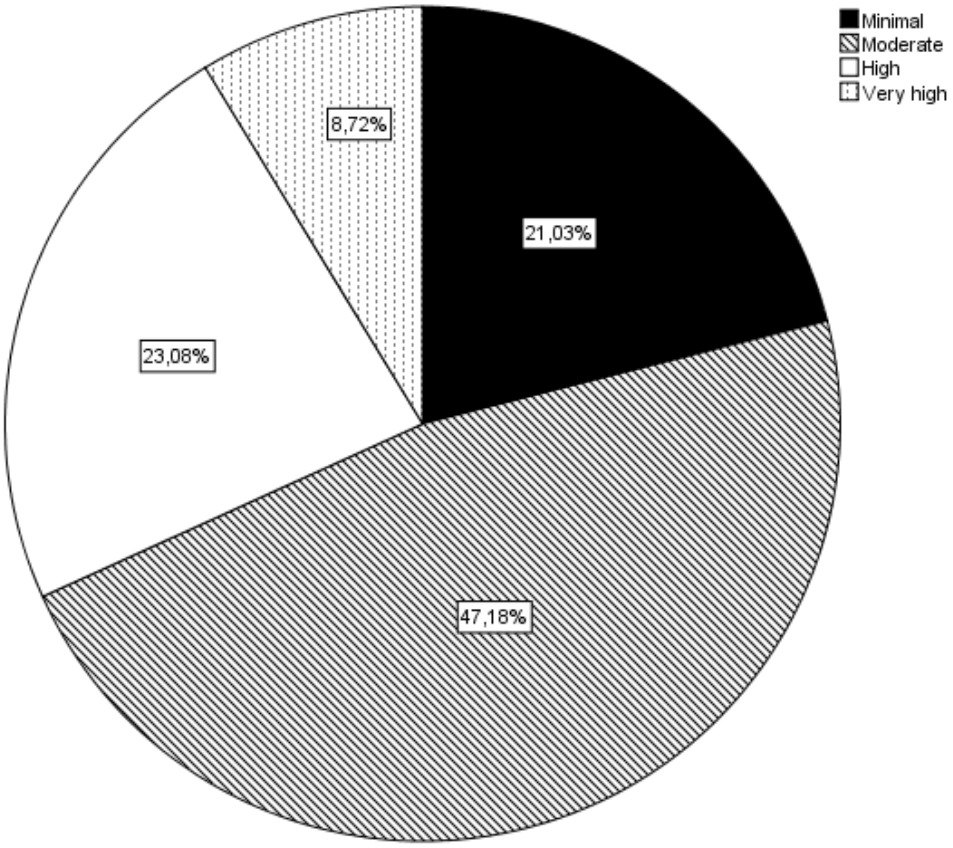

Figure 25. Percentages of frequency for the dynamics and harmonics construct.

The comments in the questionnaires about dynamics point to a better result than the one obtained with the sound objective, as previously shown. This matches the fact that we really did not count the "none" answer results for this construct; we actually only got 21.03% of the answers with the minimal value, so there is a tendency to agree on the ability to properly listen to *pianos*, *fortes*, conduction, bridges and entrances, to distinguish timbres and perceive different colors in interpretation.

Some open text field answers remarked that the amplified sound received from the teacher sometimes seemed very different from the student's sound. We do not have many written comments on dynamics, but as the video recording analysis reveals (below) a large amount of work on dynamics was done during masterclasses, which

illustrates that the artifacts allow a positive interaction for conducting purposes. Therefore, we can infer that sound dynamics were well heard.

Most of the written comments were related to the timbre and harmonics of the piano, which seem to require more attention at the mix table. A piano accompanist from the student's side says: "There is an unbalance between the quality of the teacher's voice and the piano, which is left in the background in terms of volume and quality". Clearly, there are some comments regarding voice timbre, but we received many more on the piano. Another result that was expected is that most of the written comments have segmentation for the profession of the person who answers; therefore, theater- and opera- teachers are concerned with the singers' comfort, while audio engineers recommend different devices or equalization, and the pianist tends to control the piano sound.

5.1.2. Classroom effectiveness

We have defined classroom effectiveness according to two dimensions: the seamless remote teaching and learning, that is, the practical support given by the artifact to the class under development; and fluent communication, specifically concerning the ability to properly hear and view in high quality the other part.

These are the pie diagrams for frequency percentages.

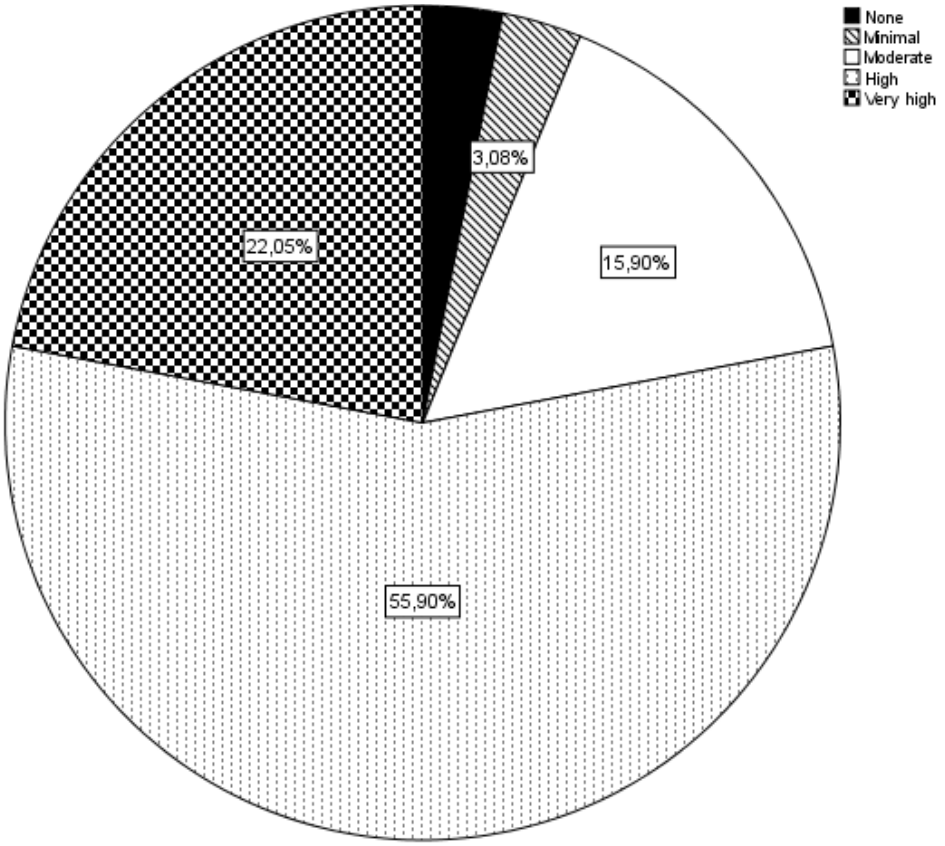

Figure 26. Percentages of frequency for the classroom effectiveness construct.

There is a clear tendency to higher values, confirmed by the descriptive statistics for the whole population that results on a mode of 4 with variance of 0.775, with a rank and a median of 4 as well. Only 22.1% of the answers were in the lowest three values (1 to 3). Students made up 13.3% of the sample, while teachers were 5.1%. Students keep the median at 4, with a variance of 0.695, while teachers have a median of 5, with 0.143 of variance.

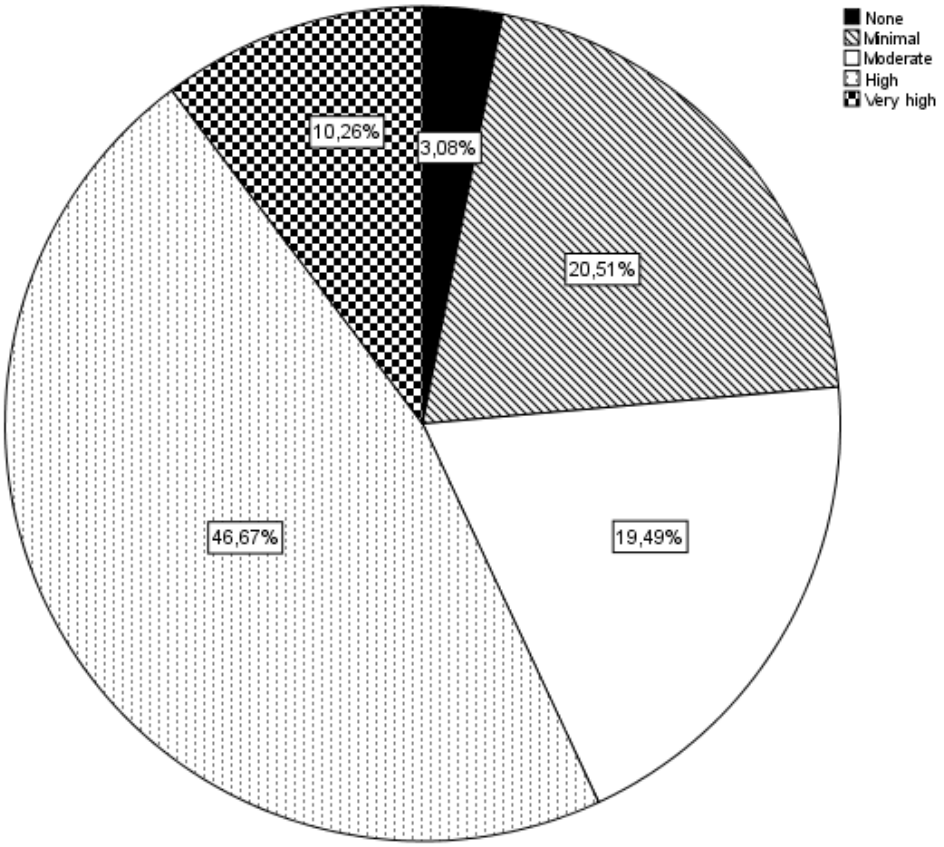

Figure 27. Percentages of frequency for the communication fluency construct.

Again we have a median and range of 4. We have a slightly larger dispersion to a lower result, if we compare the interquartile range, which is 0 for the previous construct and 1 for communication fluency. This has been influenced by a network connectivity problem we had for one set of masterclasses held between Amsterdam and Sabadell, when some classes were disrupted by lack of audio transmission (video frames can be arranged, and to the human eye if some frames are removed it is almost unnoticeable, but we cannot cut audio channels at all). We will have a section of those classes described in the recording session's account.

We can briefly analyze the data grouping by testing profile: regarding the class support construct, we find that teachers valued the class support always as high and very

high, while there are students who found that there was none. Some students only had the possibility of assisting to one class, the one with the sound cutoffs, so this must influence the result. We also have the lower results from the observers, although none for evaluators. It is quite interesting to note that evaluators tend to grade with higher qualifications because they were selected professionals whose opinions where valued for usability demonstration.

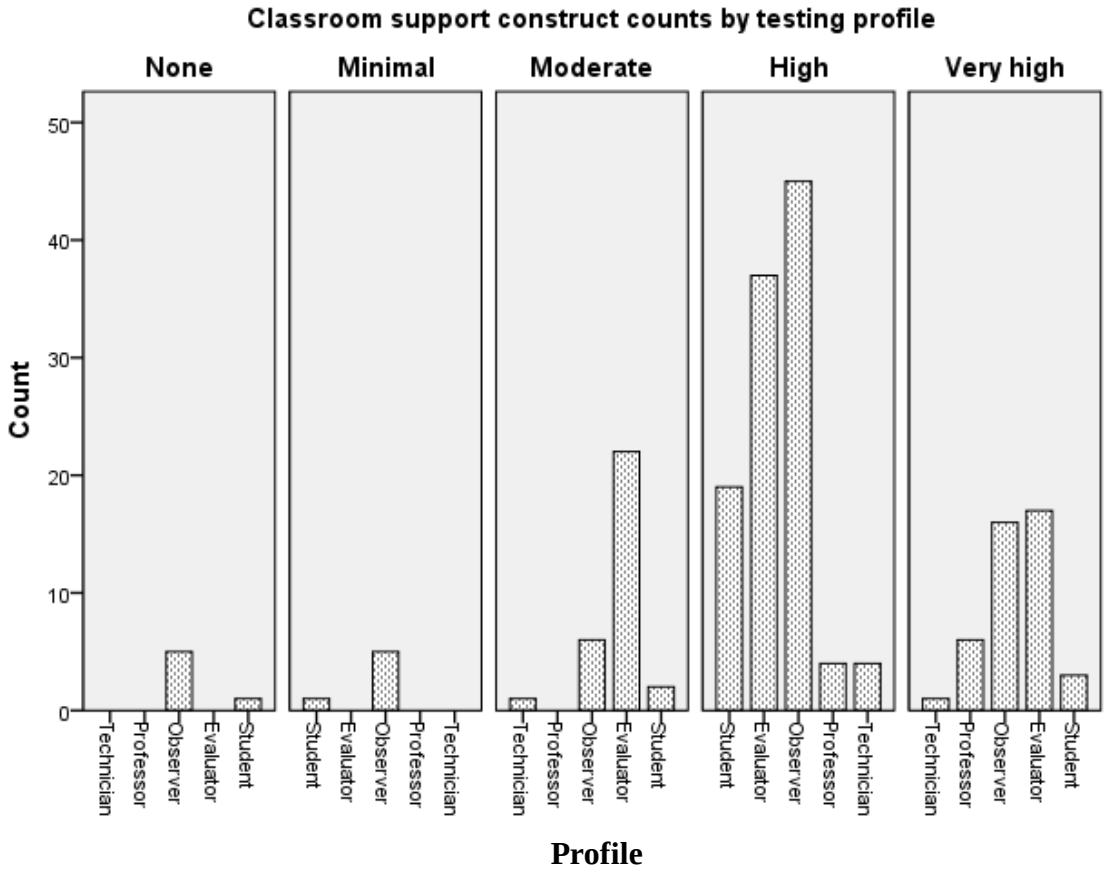

Figure 28. Classroom support construct counts by testing profile.

5.2. Qualitative study of the recorded sessions

This section summarizes the results found analyzing the available video recorded material. There were more than 30 hours of video available, although they actually correspond to 15 hours of class (both ends of the transmission were recorded). One of the executed tasks has been the video edition and compression of the classes for the website that did not include all the singers and sessions. Therefore, some videos have been analyzed in their website format, where we can see the teacher and the student at the same time in a double screen, with synchronized audio and the testing pieces eliminated from the session. When this final cut was not available, it was necessary to analyze the video from both the student and the teacher's side, since some audio sections were missing if we chose only one of the transmitters.

An expected result is that the opera classes compared to the Lied classes were more technique-oriented, simply because the subject under analysis is different. In fact, the opera technique classes included sections of vocalization that the Lied teacher did not try.

A behavior that is a common convention for teachers is to implicitly qualify the singer at first sight on first hearing them while allowing him or her to finish (or not) the first presentation song. If the teacher finds something really incorrect from some point of view, he or she will stop the interpretation and make an observation. These interruptions shape the form of the class itself, the student interprets and the teacher observes, but this initial tacit message is mutually understood and represents a convention on the conversational style of this human group. Usually teachers decide the duration of the exercises, and ask for repetitions, or dismiss the student at will; this implies that the results for the study are not equal for all sessions. We have counted the occurrences of

certain expressions corresponding to the subcategories' constructs, but the numeric values are under discussion and we can only present them as part of the analytical research developed. Since each class has a variable length, and for some students it is almost impossible to keep singing after being interrupted after two compasses, while the teacher leaves other students to finish complete pieces, we cannot estimate, for instance, how many expressions on dynamics a regular masterclass should have. Again, in the teaching of music at the higher level, this is a common situation. Classes are personalized, teachers have a different approach depending on the student. An interesting achievement of the study that was not planned is that the opera teacher decided to include in the sessions a final mutual evaluation of the experience, that is, she asked the student to tell how they had felt working in the Opera eLearning environment. This provides us with a set of comments on class support and communication constructs that, in this case were introduced as part of the class, while we had also only received spontaneous feedback on these matters from the other teacher and students. Finally, we should mention that it is very common for a teacher to speak and the student not to reply at all, but simply change their interpretation and keep trying again.

The next table shows the total numeric occurrences of observations during the classes. It has been calculated with the opera technique and Lied classes together, but we must remember that the opera lessons where only four, while the Lied lessons where ten. An observation is considered one phrase, and if it includes more than one construct, only the first one is taken into account.

Qualitative occurrences count by research category				
Category	*Subcategories*	*Construct*	*Count in Opera Technique*	*Count in Lied*
Learning objectives	Technique	Tone placing	56	2
		Air Support	38	35
		Articulation and Diction	19	72
	Expression	Conducting	3	112
		Emotion and feeling	2	43
Sound quality objectives	Tones	Medium, low and high	4	0
	Dynamics	*Piano* and *forte*	2	1
	Harmonics	Timbre and color	3	2
Classroom effectiveness objectives	Class support	Seamless remote learning	12	7
	Communication	Sight and listening	8	35

Table 17. Qualitative occurrences count by research category.

Some comments on this data are:

- Clearly, tone placing was a main objective for the opera technique classes.
- Although the number of comments made to refer to air support is quite similar, we should remember that this number summarizes only four classes of opera technique against ten of Lied singing. As we shall review in the written conversations, the Lied teacher considered air support in terms of relaxation, preparation and improvement of the color, and not necessarily as a technique issue, as did the opera teacher.
- The number of references to articulation and diction were higher in the Lied classes because there was one class where the German spoken by the soprano was not clear enough.
- The opera teacher asked for class support assessment in all her sessions, therefore these constructs show values for her sessions only.

- Cutoffs in network communication between Amsterdam and Sabadell made it necessary for the Lied teacher to comment on communication, to re-establish it and continue with the session.

In the next section we shall review some of the recorded session's samples that have been deemed more meaningful for the study. In order to make the interventions anonymous, dialogues will be shown between Teacher (T) and Student (T), removing all the names included in the conversation. It is important to mention that all dialogues were originally in Spanish, and the version offered is the author's translation.

Tone placing and air support observations

The following dialogue takes place at the opera technique classes. When teachers and students have been working together for a while, they eventually come to know each other. The singer was singing the piece, looking at his teacher while singing as usual, and he noticed that he was doing something wrong, partly because he heard something and partly because the teacher lifted an eyebrow.

S- "Mh…this was not that good…"

T- "Well, all the time…the support for the first note on the phrase is…mh…Are you nervous?"

S- "Not, but I am aware of too many things at the same time and I cannot concentrate."

(General laugh)

T- "Well said…the support for the first note of the phrase has been a little short…always aim a bit higher from the first note…"

(He starts singing)

T- "No, no, no…Stop…let's see…can you produce just the first note and maintain it for a while?"

(He sings, she does not find it good enough and sings herself, showing how, and making a gesture with the hand of lifting and sustaining in the air…he tries again)

T- "A little bit higher…" (she lifts her hand even more)

(He tries again, starts singing and stops)

S- "Now, for instance, the /u/ and the /o/ close the voice a little…and…when I get to the word "*modesto*", with the notes F and G, I do not find the space that I usually find at class…"

T- "You should place the sound rather high, and give us the /u/ and let it out now…don´t think too much…"

S- "And maybe if I…"

T- "Yes, you may soften the articulation a little bit …"

In the previous dialogue we have several observations on technique: sound placing, tone and diction. A spontaneous phrase from the singer's side establishes how he feels working within the learning environment for the first time; this happened at the beginning of the class, but half an hour later he was working almost normally, and became "committed to working through this means in other classes" as he answered in the questionnaire.

The following set of phrases took place in the Lied masterclass, and we can notice from the respiration recommendation that we have images that aim at giving feeling and color to the voice.

T- "You can better prepare the phrase, with a good respiration, as if looking at a beautiful woman that is entering the room, you have to avoid making the accent with all your body…we listen to you properly, the sound is great, you do not have to exaggerate with your body… (Singer starts singing)…No, no, stop…You thought about the beautiful woman too late…"

In this case the teacher observes tension on the body posture of the student, and provides the example of the lady to obtain a feeling that will connect the body's memory with pleasant experiences, so as to color the voice in a tender, softer and perhaps intimately passionate mood. It is also a matter of conduction, but the breathing advice is more related to air support results.

The same happens in the following mix of breathing, emotion and feeling (and even life) advice.

T- "SINGER you have to use breathing, you also will fell more joy; breathing is a fantastic help for all colors, in everyday life also, to perfectly prepare our body…"

Then (unexpectedly) the connection dropped down, and with the technical equipment we managed to keep the video signal alive by cutting frames, but the audio was lost. Teacher and student started making signs letting each other know that they could not hear and waited with a smile, while also playing a little bit through the video cameras. This class was important because this same soprano had had a class the day before in which the teacher really gave her a hard time; on this day he gave her a lot of support.

A common observation in singing classes is always: "you can finish this line without breathing again, you can do it". Other observations refer to the body posture of the singer, to relax their shoulders, stand stable and easy going, and keep the hands steady

or use them only occasionally. We have recorded several occurrences from both kinds of remarks.

The role of the pianist

The role of the pianist gave interesting and unexpected results. All the pianists situated themselves in either the students or teachers band, participating in the class from the chosen perspective. In a regular class, singing teachers usually sit at the piano playing the accompanist. For scientific rigor, it must be stated that the experience of the pianist as teacher was recorded as part of the pilot stage, so it should not be included as part of the results, if we feel fussy about case study methods. But we will include just one little piece of what happened in a session where the pianist accompanist was also a chorus director and a singing professor, older than the student. Let P be the pianist. It is a sound placing observation.

T-"It is a little bit lower, try to pick it higher."

S- (Sings)

T-"No, no a little bit higher."

S- (Sings)

T- "No, stop…"

S- (Sings, seems under stress)

P- "Try this scale (plays a scale), slowly…"

The singer plays the scale a couple of times and corrects the intonation of the song. They carry on. This is an example of an interesting participating role for the pianist that could be material for further studies or experiences.

All the other pieces of conversation come from the sessions of the case study so they can be formally included in the analysis.

We did not take into account the participation of the pianist as part of the class. Like the student, the pianist received a large amount of observations, influenced certainly by the Lied teacher being a pianist himself, although the other teacher also made remarks. Then, we should consider including a voice microphone for the pianist at the student side, too. This is actually recorded occurring several times: the student's pianist asking something to the professor, but since the microphone was devoted to the piano and was directional, his voice was not well heard by the teacher, and singers ended up repeating the questions. It is not easy to include more audio channels for the Dolby 5.0 version, but in stereo there are still plenty of places for other microphones. In this case, the LAYLA card would be the bottleneck, we have up to 6 channels. On the other hand, when pianists showed different positions for the hands, the student generally could understand what to do.

Diction and pronunciation

The worst case in these subcategories samples was with a soprano that simply could not present her first class to the Lied teacher, because he did not accept her incorrect German pronunciation. He sent her home to study her pronunciation until the following day. Even though this could sound too strict, a good pronunciation is certainly required, and the correct utterance of the vocals at least makes a dramatical difference for a phrase. The cranium resonators are different when playing an /i/ or an /o/, then timbre, harmonics, texture and color are affected. Of course, foreign languages are one of the most difficult challenges for professional singers. The German language in particular requires extensive study if your mother tongue comes from Latin roots.

T- "I have forgotten, the "tse" word, the nice "tse" word sometimes is completely right, but sometimes you introduce an /s/ with strong sound, which is not correct, "tse", /t/ and /s/, we cannot choose a more beautiful sound, but we have to do /ts/.

The next sample also includes conduction and emotion components. But it is also related to diction and pronunciation.

T- "'*Blaue augen'*[8] was very good, but '*miitet'* was too short, those words you know '*mi-i-tet'* they are always in danger of being too short… '*miiten blaugen augen'*…and this '*kornblumen'* could be with a nice /k/…and with a better breathing preparation in the breathing point… (Sight)…it is something very pleasant…"

S- (Starts singing)

T- "No, no, I am missing the good breathing."

The next sample illustrates how to properly stress the words in the singing process and help the good direction of the entire phrase.

T-"Thank you, a nice /f/…but the piano should avoid this accent…" (SINGER starts and teacher stops again) "Thank you…the word '*fremd'* is the most important one, because it describes something typical of the whole cycle, you are a stranger (here he stops and asks for help with his Spanish, so he confirms the word 'stranger'). This '*fremd'*, and the word '*eingezogen'*, they have an accent over the head, '*ein'* and '*aus'* are the two different directions…and everything is dominated by '*fremd'*… The first compass was marvelous… (SINGER starts singing). Thank you…this /se/ does not matter at all…and try to connect a little bit more the syllables, a tenser *ligato*…" (He

[8] From this moment, words in German will be written in *italic characters*.

draws an imaginary circle in the air that is a common sign in music to represent the continuous of a *ligato*).

Conduction, emotion and feeling

It is in this section where musical higher education has more intensity, and there were a large number of comments regarding these constructs. We shall go through some short comments that came from both teachers:

T- "Try to avoid this *retardando*."

T- "You entered too fast / too late / too loud."

T- "Here the piano could be *pianissimo*, more than in the dynamic point of the color."

T- "Desire as a Romantic concept requires that the music be beautiful…it is a positive desire…"

T- "Horror comes with the piano, but is neutral in the voice."

T- "You may enjoy this part more."

T- "This accent is a hint from Schumann for the next phrase."

T- "This is too correct, PIANIST, those thrills could be more like a game, with a good mood, enjoying and controlling the sound at the same time.(…) SINGER, you may also enjoy it, playing around, those are waves, waves between you both."

T- "PIANIST, do not introduce accents in the piano because SINGER always accepts the invitation…" (Laughs)

Now we shall go through a piece of conversation where we find the accents topic, a classic issue in conduction.

T- “In this part, if you want, if you do not have other idea …”

S- “I do, I do…”

T- “You may try a long /u/…”

S- “No, I do it because I want to match the piano’s accent.”

T- “Yes, but we are independent sometimes.”

S- “Of course, yes." (Laughs).

T- “Since we are talking about accents…What is an accent? How much should an accent be stressed? These are questions that only we can answer while playing and singing.”

S- “They are rhetorical questions.”

T- “No, no, no, no they are not rhetorical, they are important questions, these accent matters, sometimes an accent is light, sometimes it is strong…”

Other interruptions from the professor gave form to interesting images to support the interpretation.

T- “In this part you say the word darkness, but it is not darkness opposed to light, it is the darkness of the soul, it is a pessimist feeling. And then you are missing a change of color, because you speak of a friend and nevertheless it is a thought with little light….it is moonlight…use the German consonants, which are hard, to illustrate the darkness, it is cold, it is pessimistic…do not overuse accents, because if everything receives an accent then you lose the effect…you can avoid breathing in this line…look for the emotion, the color.”

To finish this section, the following anecdote was really amusing to witness, and caused general laughter at both sides, and remarks on duet conversation and entrances management.

T- "Very good, but in this part of the song I have the impression that SINGER is not always the guiding person, we have to be very decided, the pilot is SINGER and the copilot is PIANIST. In German we call the copilot 'the first officer', and he is not less, but in case of emergency, the pilot is always the one that decides…and always the first officer is the one that does the job (general laughs)…For instance, when this compass starts (plays the piano), when to enter is SINGER's decision, positively, and if she decides to enter too early, it is as it is… and we notice it and pretend nothing has happened and carry on, because SINGER as pilot always has the right… Of course, it is not that the piano is less valuable, it is just like with traffic rules, in case of emergency if a huge truck is driving where it does not correspond, this is what the rules are for: to react quickly."

Music history and style

In the Lied classes the teacher offered some observations that aimed at involving the student in learning about the authors that they are singing, and describing the style required.

T-You should emphasize this duet because Schumann in this part of the piece is trying to transmit an unpleasant emotion (…) here the piano and voice do not understand each other…

T- There is great danger, particularly with Strauss's music if there is too much *appogio*…probably in Strauss' time they used to do more *appogio* than we can appreciate today, but we cannot change the general taste, it has to be less like a plane

that arrives…it does not need that…when we have all those adjectives, '*rotblutigen' 'gesunden',* they could have a different color, we have to display those quick changes in emotion…

T- There is a little change in the rhythm from Schumann to Schubert with this plucky rhythms, Schumann sometimes changes the punctuation for special little moments, and this is not always easy to memorize…

Meaning of the words being sung

For a great interpretation it is always important to fully understand the meaning of the words being sung. Some examples of hints on this aspect are listed below.

T- "To whom are you singing?"

S- "To Malinconia."

T- "Yes, who is she?"

S- "A muse, a…"

T- "A nymph, how do we think nymphs are? What do they look like? Are they ugly, odious or are they kind? What do you think?"

S- "Kind."

T- "Yes, I think so, too. Let's build the feeling if you were talking to a beautiful, nice nymph, that you certainly enjoy, although she is a little bit sinister (…) and when you sing "gentile nymph", is she really gentile? And if she is not, let her be gentile if you sing…we need a little bit of desire…she does not listen to you freely, but if you sing with great emotion and desire, she will listen…

Another example of the same kind is described in the following lines.

T- "What do the lyrics say? Who is him?"

S- "He is the greatest one."

T- "He is the most handsome man; the marvelous…let us feel it in your voice…"

S- (Sings)

T- "Who is him?"

S- "He is the lover."

T- "I don't know if he is in love, but *you* are in love and this could be the difference…let's see if we understood that he is the most fantastic man in the world."

Classroom support and communication fluency

Some of the recorded conversations or expressions are related to the classroom support and communication fluency constructs. The teachers gave some indications that were only related to this artifact.

T- "The line was much better with distance from the microphone."

T- "We listen to you properly, don't make your body tense." (This is a strange feature, but most of the singers started singing very high, as if they tried to make up for the distance… just like a person who screams at the telephone because they are afraid the other will not hear them).

S- "I felt that because of the distance I had to do everything bigger, exaggerated, which is very useful for a concert as well."

Our main deficiency in sound was the echo. This echo is the result of placing such advanced technology in microphones and speakers within a small university classroom. When the teacher or student started talking it was the audio engineer's responsibility to set the volume up or down of the correspondent microphone, in a system similar to the *push to talk* method, but provided by the two engineers (local and remote). A small distraction was enough to hear a huge echo. We have recorded a large number of comments about it, for instance the Lied teacher made a joke on hearing himself as if he were a giant. Students also complained about it, as did observers. And they were all right. In the next piece of text the echo issues where contextualized by the teacher in relation to the rest of the experience.

T- "We have finished now, it is a good experience for us to work with this huge distance, from Northern Europe to nearby Barcelona…a nice little adventure, because we realize that it is possible to work this way…some annoyances we have experienced now must be corrected, but for the next time we will know how it works, so we may be more relaxed than today…"

Another deficiency occurred the only time the network fell down continuously every three or four minutes. This was an important problem because the class re-started, and was suddenly interrupted again. This issue lasted 1.5 hours, affecting three masterclasses.

On the other hand, when the artifact was successfully working communication became so fluent that in some classes teacher and students were able to make substantial improvement.

It is not necessary to include more dialogues about the comments on class support, when the opera technique teacher asked the students how they felt and if they could

work again with this artifact, they all answered positively, as they later wrote in the questionnaire. They also agreed to being open to take a formal, curricular masterclass in this manner.

An interesting fact that may be observed in each recorded session is that, after a while, teachers and students forgot that they were in a virtual space, relaxed and started working fluently and with far-reaching effects, speaking to the plasma monitor and forgetting about the rest of the people in the room, even trying to achieve some functions that are still impossible, such as playing together at the same time, or literally walking towards each other.

.

6. Conclusions

The main conclusion confirms the effectiveness of the remote artifact, which requires a large number of technological equipment and resources, but provides a framework to develop teaching lessons at high professional standards. Our development cycle finally achieved a sustainable model. Since the resultant artifact has afterwards been used in remote opera singing lessons with success, we can agree to consider that the development process accomplished the development goal. Of course, some improvements can and should be included in future versions.

Other interesting conclusions can be drawn from the assessment of the innovative and multidisciplinary research process, especially taking into account the user-oriented, innovation and usability approaches.

.

6.1. Research results

In the interaction of experts from many different disciplines, as well as in the management of a project of technological complexity, our guide was to follow the user-centered design, i.e., to define, plan, test and redesign according to the user's requirements and feedback on an iterative cycle. User-centered design provides a stable, well defined quality standard for the developing process. One useful shortcut for the management of multidisciplinary teams was to work with individuals who had more than one specialty; in this case, some of the engineers were also musicians, so the common language was easily generated, and some detailed requirements of the users where fully understood by the technical team.

We acknowledge that some of the activities developed did not become part of the final model, but this is the price of innovation: to try different options, and to choose the best of results.

From the formal scientific point of view, the hypotheses under examination have been contrasted and evaluated with the obtained results, overall having a good result on the usability dimension. Of course, the execution of previous pilot tests was fundamental in the design, and helped deliver a version of the artifact that had already proven itself usable: this is the process followed to make stable technological innovation in most fields. Remembering the case study methodology suggested by Yin, the generalization proposed is to be able to repeat the exercise, to repeat the generation of the artifact and installation for other schools or grades. We have sufficiently proven that this can be done with a wide margin of reliability. Also following Yin's approach, a very strong pilot stage supported a further solid case study research, not only because the artifact had already been tested, but because we had the chance to tune the measuring methods and tools, as can be contrasted in the first part of the empirical work section.

The research categories evaluated in the rating questionnaires where Sound Quality Objectives and Classroom Effectiveness. The former was intended to confirm the usability in the recording, transmission and reception of the sound wave, and the latter was used to measure user satisfaction with the communication and class support provided by the artifact. Although the questionnaire results point towards a generally rather good evaluation from users, this is a study oriented to distance learning, which is the cornerstone of the process: a remote class of higher education music. We are including the student as final user, but our main focus has been oriented to the teacher.

In the research categories for the distance learning dimension, that is, the learning objectives that were freely commented by teachers but not evaluated in rating questionnaires, the analysis of the recorded material illustrates the accomplishment of the following items, confirming the opinion of teachers on the viability for the remote working session:

- Tone collocation was corrected profusely, especially in the Opera technique masterclass. The student's abilities to reach the tones with precision, the proper "open throat" technique, and enough relaxation was immediately perceived by the teacher and meticulously corrected if necessary.
- Air Support has more commented occurrences, and is addressed by both teachers when required. This can easily be explained because a lack of air support produces low quality singing, therefore showing the good audio perception of the teacher.
- Articulation and diction have received a large number of observations. It should be remembered that the Lied teacher was German, so he tended to be more strict about foreign language pronunciation. Nevertheless, this large amount of observations allows us to conclude an excellent result for sound reception: the sound of consonants, which are shorter, rough and sometimes hard to amplify (like the /s/, the /r/ or the /f/, that tend to make noises through a good microphone) got proper reception. Vowels are easier to amplify and their transmission is more likely to be well completed (especially open vowels, /a/, /e/ and /o/), but when we receive so many observations on articulation and diction we can be certain that also consonants are being received correctly.
- Regarding the Conducting construct we find most of the occurrences take place in the Lied masterclass sessions. As mentioned above in Table 17, a larger

quantity of observations are expected in Lied classes compared to Opera technique lessons, since we had ten Lied and four Opera masterclasses delivered. However, the result for this construct illustrates the fluency of the masterclass, as well as the competence of the teacher as a superior grade music teacher.

- It is in the Expression, emotion and feeling construct where results should be the priority, as proof of the fine service provided by the telepresence artifact. Providing comments on emotion and feeling is a personal matter, and before they can be given, it is necessary to have established a reasonable level of personal rapport. This was a recurrent concern of singing teachers when the system was designed: Would we feel comfortable enough to discuss feelings, to reach emotional states, to obtain different colors from students? Or would this technology-aided environment create a cold, unnatural feeling? The answer to these questions provided by this study shows that a great number of emotional and feeling-related aspects were successfully discussed and put into practice by the students: being in love, feeling the presence of the beloved one, being inspired by beautiful women, all allow the involvement of feelings in the discussion process. Furthermore, speaking about existential loneliness, or feeling a stranger, or resorting to metaphors like "the moonlight penetrating the deep darkness of the soul" have been fluently worked through during the masterclasses.

We used qualitative research analysis for this last category of learning objectives, and enough evidence of the accomplishment of these learning objectives has been measured. Therefore, the conclusion is that it is certainly possible to create this personal relationship between teacher and student, which usually is developed after

a time of working together, even when seeing each other face to face. It is remarkable that most of the students in the Lied masterclass were taking classes with this professor for the first time. So, we could say that they built the personal relationship through the telepresence artifact.

6.2. Required improvements

Connectivity matters should be addressed and an available workaround should be considered. For instance, we could turn to a stereo, DVTS-only configuration that requires less bandwidth. In fact, for stereo configurations in a one-to-one class the recommendation is to use only DVTS, and therefore being able to skip the requirement of support from audio engineers, for instance.

Whether in the stereo or multichannel configuration, the echo should be addressed and provide a stable solution. In the masterclass sessions it is imperative to have a proper auditory, theatre, or TV studio to manage such sound. We cannot afford to have a professional C5 piano with a large coloratura soprano voice in small rooms: in this case we also can go back to the stereo configuration used at the beginning as a better practice, unless the room's acoustic is fantastic.

Finally, the piano sound could be improved assigning a second or a double capsule microphone, which would be mixed in real time to improve the piano's received sound.

6.3. A stable version for a whole master degree

We look forward to having the possibility to install a permanent device of Opera eLearning design in order to provide a whole master's degree course, through a long term project that would allow data collection and innovative applications. The obtained results support the viability to use the tools developed for the case study to start operations. However, for the Council of Sabadell such a master's degree is unlikely to progress until the optical fiber network arrives to the City of Music. Another possibility would be to install a remote laboratory connecting other conservatories.

From the scientific point of view, the best test possible would be to have the same teachers and classes given with and without the artifact, so as to obtain a control group. However, it is more reasonable for a university to finance such further project if all the classes have the possibility to count with the remote learning environment. We are ready to support such a project, and provide all the technical know-how.

Here we can make a brief comparison with other projects, not necessarily other case studies, but other projects on the same field. As we have already reviewed in the related work section, the American New World Symphony (New World Symphony, 2009) is delivering masterclasses from one auditorium to another using Internet2 and a set of support professionals in a similar way to the Opera eLearning project, giving successfully remote access to a blended learning approach for formal degrees. We can say that this experience could be a reference that illustrates the possibility of delivering a whole singing degree in this manner.

The project's webpage and other social networks should be the complement of these regular sessions, providing extra self-practice, documentation and communication tools services to the students.

6.4. Weakness of the study

The main weakness of this study is the limited number of cases under analysis, and the possible bias of the population used. All the teachers involved do certainly have an open mind about working with technology. It is important to keep researching with a population that uses this classroom regularly, as part of the usual curriculum. We also look forward to being able to permanently fit out laboratories with the telepresence classroom. The fact that different rooms had to be used in this study may affect some variables under analysis, and this could be another weakness of the study.

There is more data available as a result from the case study, but we have not been able to analyze it, and this constitutes another weakness. Audio recordings were not decomposed in acoustic terms –in Fourier fundamental waves, for instance– to analyze the results of transmission. Here we have a limit for this research: the time schedule required to assign priorities from all the possible research methods, and as we opted for usability and user centered design as a guide, we chose to analyze other parts of the data. It is possible that a larger research team would have been able to properly analyze all the data collected. This idea leads us to the further studies required section.

6.5. Further studies required

Here we refer to several possible pathways for further studies.

- We should have a formal study on the stereo-only configuration in a one-to-one class. There are several institutions on higher education doing this, so this would not be as innovative, but nevertheless it should be incorporated to the Opera eLearning baggage.

- The role of the pianist suggests that a microphone and perhaps another camera should be in place for this musician. If we integrated another camera, then we should have a video mix in addition to the audio mix. Another excellent idea would be to focus a third additional camera into the teacher's fingers at the piano. This opens the path to having close-up shots of the singer as well, with more than one camera.
- Small choral ensembles or duets could be assessed by a teacher with almost the same configuration. This has not been formally tested, but it was certainly tried spontaneously by singers and worked fine. This should be formally tested.
- Other musical instruments could be the taught subjects, and perhaps some adjusting of the microphones could be required. More microphones distributed in the room would certainly be required, each one specialized for each instrument or group of instruments, and the perception of the room's acoustic should be reviewed (it is not necessarily the same main frequencies that a chorus or string ensemble actually emit, and this would help decide where to put the additional microphones in the room).
- Another aspect in terms of the setting up of long-term telepresence classrooms is that intelligent agents could be in charge of functions that in this study were accomplished by technicians. This is far beyond the scope of this study at the present moment, but it is clearly an excellent research line.
- All the audio recordings are sound waves that could be decomposed in Fourier fundamental frequencies, or analyze the decomposition of the wave in their long term average spectra LTAS. Studies by Mitchell have already used the latest techniques to show that the open throat technique can be predicted by the wave sound, but that is not the case with the beauty or quality of the voice. These kind

of interesting matters could be addressed, trying to associate musical characteristics to the physical wave sounds, instead of simply comparing the emitted and received wave sounds.

6.6. Final conclusions

Music and all scenic arts are integrating new technologies, not only for production, but also for publishing and broadcasting materials. The possibilities of distance learning for the graduate level are increased with high bandwidth because of the fidelity in the transmission of sound and image. During the following years, we will be able to experiment with this remote human-to-human interaction until we make it completely seamless.

The aim of this telepresence application is not to substitute the personal tutoring that is highly valued in musical education, but to allow in some cases the remote tutoring, providing all the support necessary for a personal relationship between teacher and student to develop. This study has been carried out for the voice instrument, but other instruments could be the subject. The idea of a telepresence music learning classroom is completed with the computer-assisted tools and support that were introduced at the beginning of this thesis paper, so as to integrate a whole blended learning system (partly remote and partly presence learning) where support for daily practice is enabled, opportunities for one-to-one tutoring increase, and web platforms enhance the social interaction and information delivery.

Appendixes

Appendix A. Satisfaction Questionnaire Sample

Sr/a (Singer)

Fecha: 25 de Marzo ■ **26 de Marzo** ☐

Lugar: Amsterdam ☐ **Sabadell** ■

Perfil de prueba: Estudiante ■ Profesor ☐ Evaluador ☐
Observador ☐ Técnico ☐ Otro ☐

Profesión: cantante e-mail:
Professora d'alemany i cant, directora de coral

EFECTO DOLBY 5.0 SIN ECUALIZAR Mal
Excelente

Agudos	1	2	3	4	5
Medios	**1**	2	3	4	5
Graves	**1**	2	3	4	5
Timbre (harmónicos)	**1**	2	3	4	5
Dinámica (entre Piano i Forte)	1	2	3	4	5
OBSERVACIONES: A vegadas se sentía massa fort el so retardat d'Amsterdam en el fons.					

CLASE DE CANTO Poco Mucho

Fluidez de la comunicación	1	2	3	4	5
ELS como soporte a la clase	1	2	3	4	5
OBSERVACIONES: (+) A l´haver-hi més distancia m'obligava a fer tot el que feia mes gran, exagerant, això segur que m'es molt util a l'hora d'un concert també. (-) Estava molt insegura de com se sentía el so nostre desde Holanda, i també, just en el tema delicat de Lied, em falta un contacte més proper amb el professor, que difícilment es pot obtenir a través d'una pantalla. També es gairebé impossible poder canta tocant el Sr. Rieger. De todas formes m´ho he passat molt be i la qualitat del so M´ha sorprés gratament.					

Appendix B. Pictures of the artifact

Next slides show the following equipment, correspondent to student's site at Sabadell:

- **5 microphones**
  - 2 for room's acoustic
  - 2 for singer's voice
  - 1 for the piano
- **4 lights**
- **3 speakers**
- **2 plasma monitors**
- **1 HD video cameras**
- **Network computers and digital converters**
- **Audio processing equipment (mixer, cables, etc.)**
- **Scores stand**
- **Professional piano**

Figure 29. Picture of boxed equipment.

- **5 microphones**

Figure 30. Picture of 2 microphones for room's acoustic.

Figure 31. Picture of 2 microphones for singer's voice.

Figure 32. Picture of 1 microphone for the piano.

- **4 lights**

Figure 33. Picture of lights.

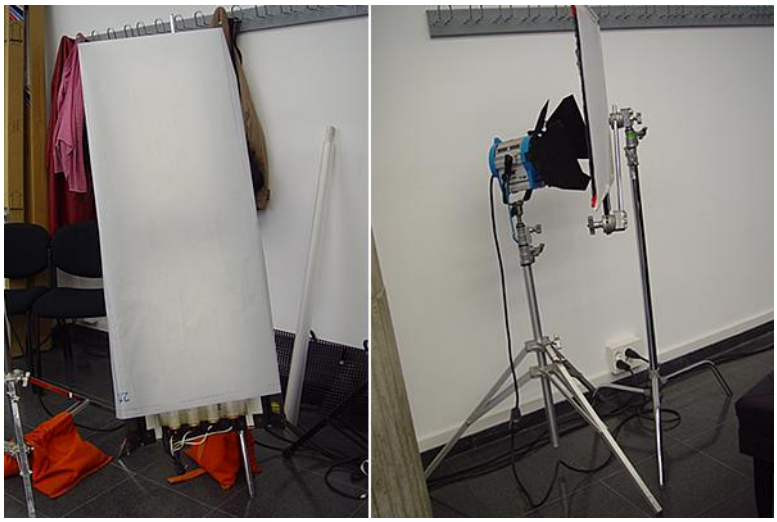

Figure 34. Picture of lights.

Figure 35. Picture of lights.

- **3 speakers**

Figure 36. Picture of speakers.

- **2 plasma monitors**

Figure 37. Picture of plasma monitors.

- **1 HD video camera**

Figure 38. Picture of videocamera.

Figure 39. Picture of videocamera.

Figure 40. Picture of video camera plugged to a computer.

- **Network computers and digital converters**

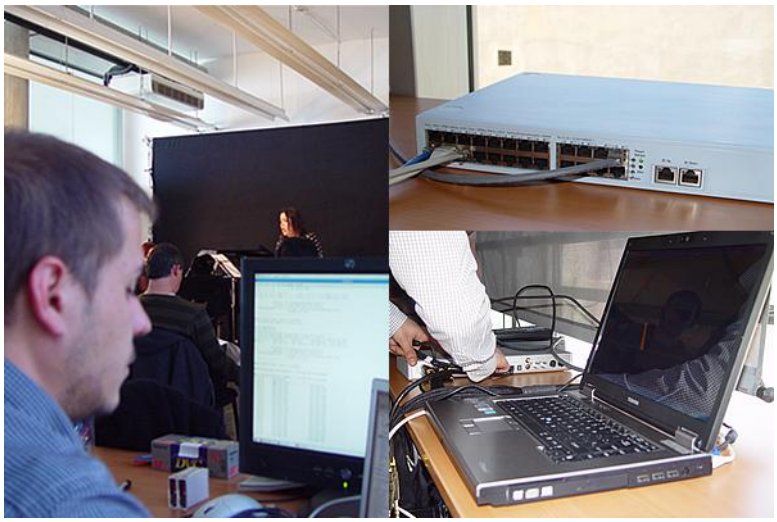

Figure 41. Network transmission computers and digital converters.

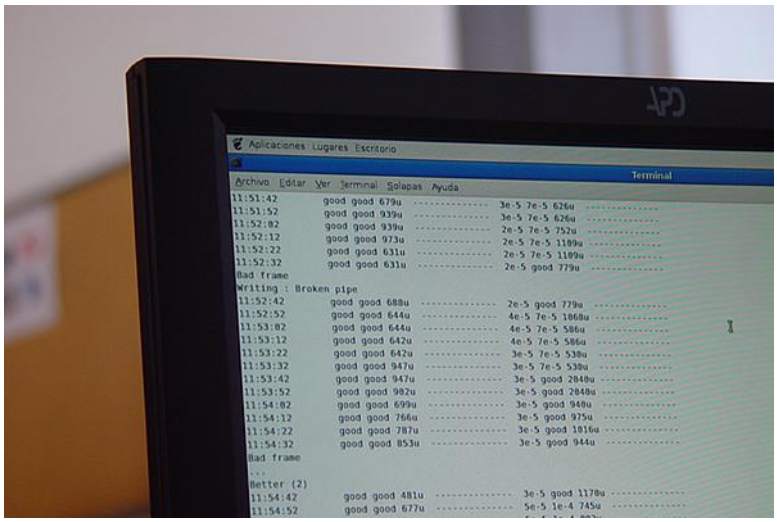

Figure 42. Picture of transmission statistics.

- **Mixer, recorder and audio equipments**

Figure 43. Picture of mixer, recorder and audio equipment.

- **Scores stand**

Figure 44. Scores stand and author leaving the room.

- **Professional piano**

Figure 45. Professional C3 piano.

Appendix C. UML Use Cases (in Spanish language).

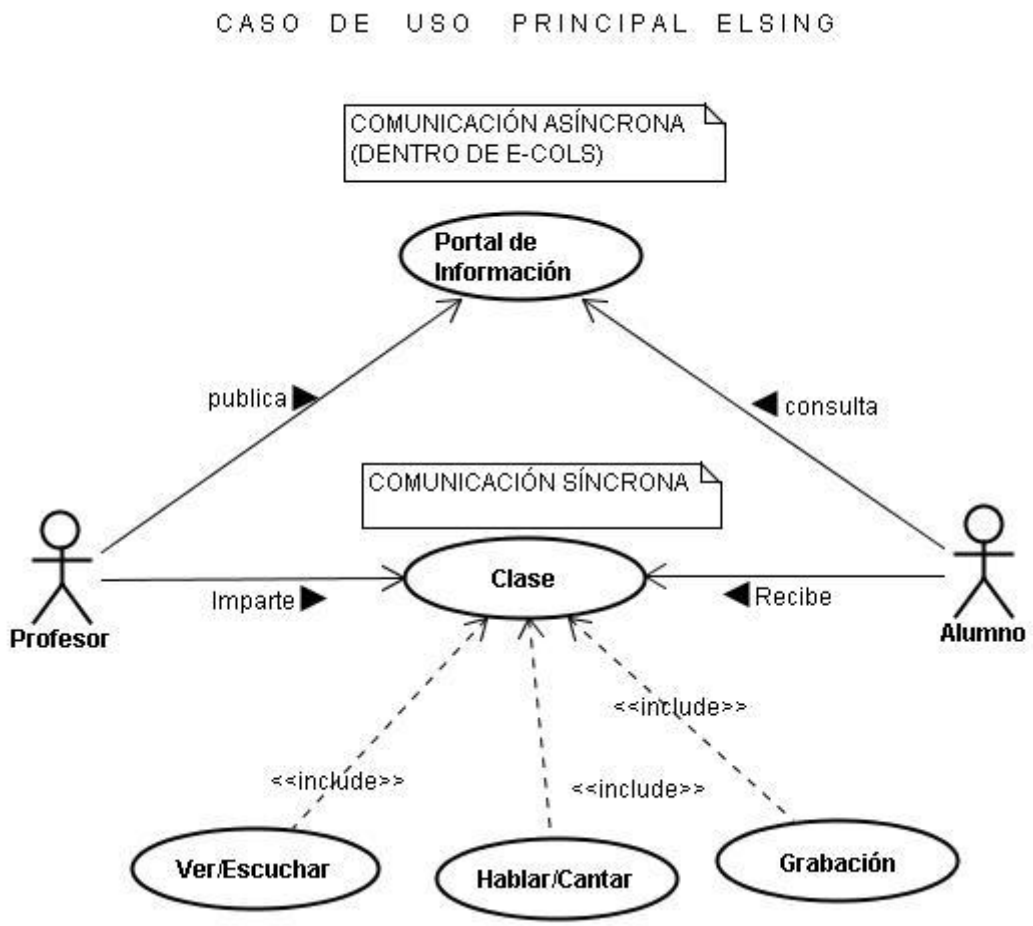

Figure 46. Main Use Case

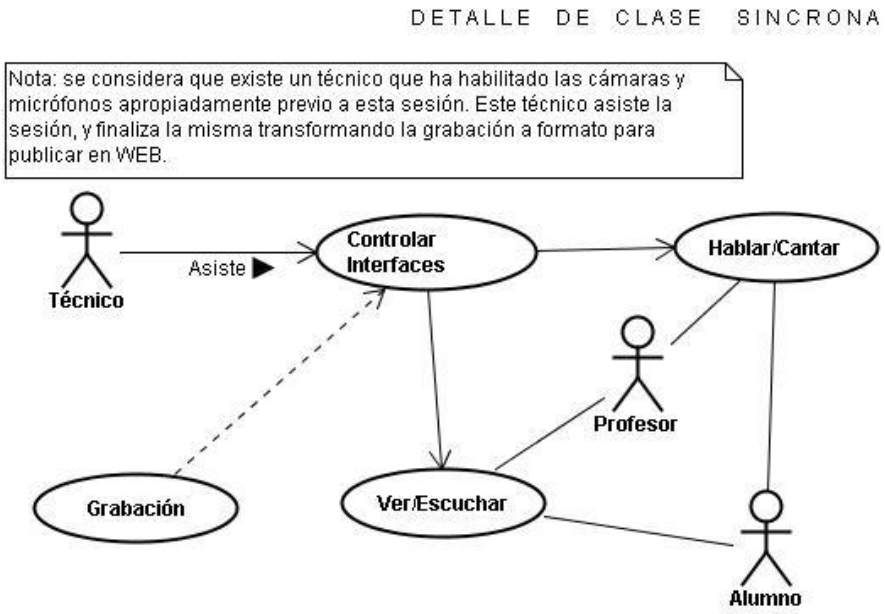

Figure 47. Synchronous class description in UML diagram

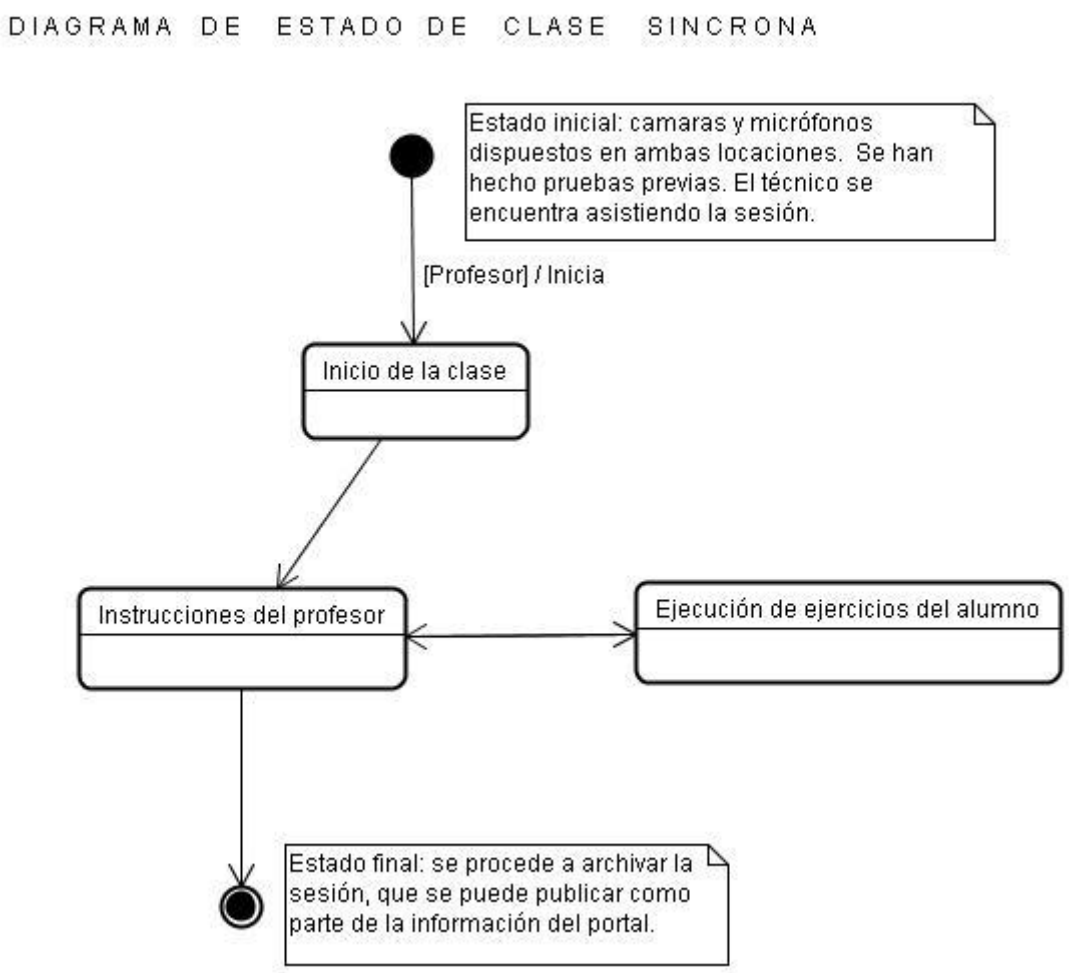

Figure 48. States diagram of the synchronous class

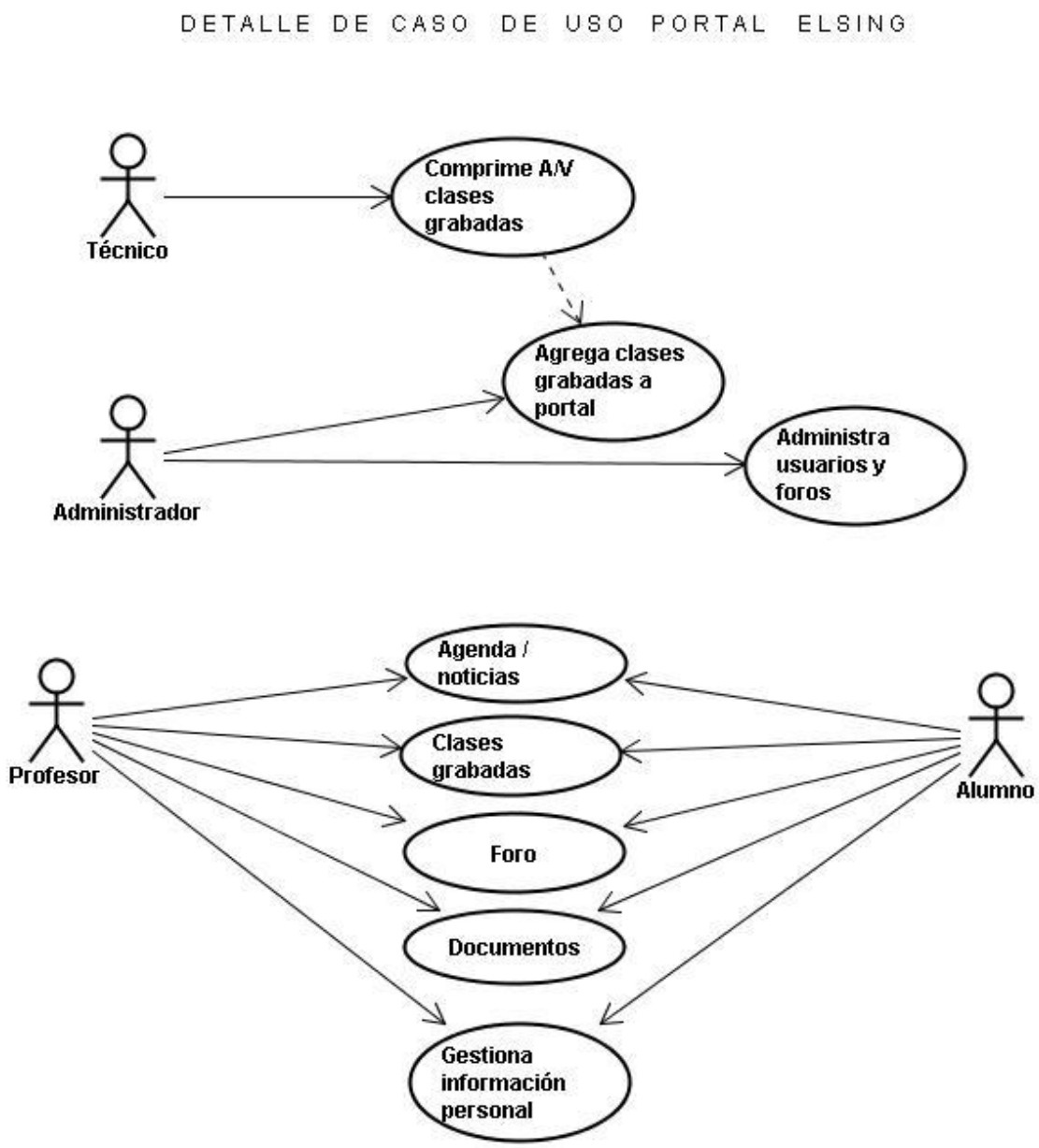

Figure 49. Web Site use UML diagram

Bibliography

Abrahmov, S. L., & Ronen, M. (2008). Double blending: online theory with on-campus practice in photography instruction. [Article]. *Innovations in Education and Teaching International, 45*(1), 3-14. doi: 10.1080/14703290701757385

Akin, D. L., Minsky, M. L., Thiel, E. D., & Kurtzman, C. R. (1983). *Space applications of automation, robotics and machine intelligence systems (ARAMIS) Phase II, Executive Summary*. Hunstville, AL.: NASA, Marshall Space Flight Center.

Alpiste, F. (2008). Modelo de formación híbrido para el diseño, producción e integración de contenidos formativos y entornos virtuales de aprendizaje. *Material docente del doctorado en Multimedia.* Barcelona: Laboratorio de Aplicaciones Multimedia

American Orff-Schulwerk Association (2011). What is Orff-Schulwerk? Retrieved 21/10/2011, from http://www.aosa.org/orff.html.

American Psychological Association. (2012). APA Style, Retrieved on 2012, from http://www.apastyle.org/

Anderson, A. J., & Ellis, A. (2005). Desktop video-assisted music teaching and learning: new opportunities for design and delivery. *British Journal of Educational Technology, 36*(5), 915-917.

Argaman, E. (2009). Arguing within an institutional hierarchy: how argumentative talk and interlocutors' embodied practices preserve a superior-subordinate relationship. [Article]. *Discourse Studies, 11*(5), 515-541. doi: 10.1177/1461445609340498

Bannan-Ritland, B., Dabbagh, N., & Murphy, K. (2000). Learning object systems as constructivist learning environments: Related assumptions, theories, and applications.The Instructional Use of Learning Objects: Online Version. Retrieved from http://reusability.org/read/chapters/bannan-ritland.doc

Barbour, J. B. (2010). On the Institutional Moorings of Talk in Health Care Organizations. *Management Communication Quarterly, 24*(3), 449-456. doi: 10.1177/0893318910370269

Barron, M. (1993). *Auditorium Acoustics and Architectural Design*. London: E & F SPON for Chapman & Hall.

Baño, F. (2003). *La Antitécnica : la impostación vocal en la ópera, teatro musical, jazz, música ligera y agrupaciones corales*. Madrid: Alpuerto.

Beranek, L. L. (1962). *Music, Acoustics and Architecture*. New York/London: John Wiley & Sons.

Bergstroem-Nielsen, C. (2011). The methodology of Godfried-Willem RAES' improvisation teaching at Conservatory of Gent, Belgium Retrieved 21/10/2011, 2011, from http://www.logosfoundation.org/g_texts/improvisation-methodology.html

Berklee School of Music. (2012). Berklee School of Music Web Page, from http://www.berkleemusic.com/

Bersin, J. (2004). *The Blended learning book : best practices, proven methodologies, and lessons learned*. San Francisco: Pfeiffer.

Bevan, N. (2001). *Cost effective user-centred design using ISO 13407*. Amsterdam: I O S Press.

Bevan, N., & Curson, I. (1998). Planning and Implementing User-Centred Design. Retrieved from doi:http://www.nigelbevan.com/papers/ucdtut97.pdf

Bevan, N., Kirakowskib, J., & Maissela, J. (1992). What is Usability? Retrieved from doi:http://www.nigelbevan.com/papers/whatis92.pdf

Boehringer, D., & Wolff, S. (2010). The Computer as a "Partner" in Institutional Talk. *Zeitschrift Fur Soziologie, 39*(3), 233-251.

Bonk, C. J., & Graham, C. R. (2006). *The handbook of blended learning : global perspectives, local designs*. San Francisco: Pfeiffer.

Booch, G., Jacobson, I., Rumbaugh, J., & Safari Tech Books Online. (1999). The unified modeling language user guide *The Addison-Wesley object technology series* (pp. xxii, 482 p.). Retrieved from http://proquest.safaribooksonline.com/?uiCode=valencia&xmlId=0201571684

Budd, T. A. (2002). *An introduction to object-oriented programming / Timothy A. Budd* (3rd ed.). Boston [etc.]: Addison Wesley.

Carnegie Mellon School of Music. (2012). Carnegie Mellon School of Music Web Page, Retrieved on 2012 from http://music.cmu.edu/

Carroll, J. M., & Kellogg, W. A. (1989). *Artifact as theory-nexus: hermeneutics meets theory-based design*. Paper presented at the Proceedings of the SIGCHI conference on Human factors in computing systems: Wings for the mind.

CCRMA. (2012). Center for Computer Research in Music and Acoustics Web Page, Retrieved on 2012 from https://ccrma.stanford.edu/

Centro Virtual Cervantes. (2012). Par Adyacente, Retrieved on 2012, from http://cvc.cervantes.es/ensenanza/biblioteca_ele/diccio_ele/diccionario/paradyacente.htm

Chen, I. L., Eckhardt, J. N., Sinkowitz-Cochran, R. L., & Jarvis, W. R. (1999). Satellite videoconferencing for healthcare workers: Audience characteristics and the importance of continuing education credits. [Article]. *Infection Control and Hospital Epidemiology, 20*(11), 778-780.

CIRMMT. (2012). Centre for Interdisciplinary Research in Music Media and Technology Web Page Retrieved on 2012, from http://www.cirmmt.mcgill.ca/

Colwell, R., Richardson, C. P., & Music Educators National Conference (United States of America). (2002). *The New handbook of research on music teaching and learning : a project of the Music Educators National Conference*. Oxford [etc.]: Oxford University Press.

Conde, J. G., De, S., Hall, R. W., Johansen, E., Meglan, D., & Peng, G. C. Y. (2010). Telehealth Innovations in Health Education and Training. [Editorial Material]. *Telemedicine Journal and E-Health, 16*(1), 103-106. doi: 10.1089/tmj.2009.0152

Creech, A., Gaunt, H., Hallam, S., & Robertson, L. (2009). Conservatoire students' perceptions of Master Classes. *British Journal of Music Education, 26*(3), 315-331. doi: 10.1017/s026505170999012x

Cummings, J. N., & Kiesler, S. (2005). Collaborative research across disciplinary and organizational boundaries. [Article]. *Social Studies of Science, 35*(5), 703-722. doi: 10.1177/03061270505535

Dalcroze Society of UK. (2011). The Dalcroze Society Retrieved 18/10/2011, 2011, from http://www.dalcroze.org.uk/indexb.html

Digital Video Transport System, DVTS (2009). Retrieved on 2009 from http://www.sfc.wide.ad.jp/DVTS/

Draper, J. V., Kaber, D. B., & Usher, J. M. (1998). Telepresence. [Article]. *Human Factors, 40*(3), 354-375.

Ellis, L., & Curless, I. (1986). A survey of video taped course offerings in higher-education. [Article]. *College and University, 61*(3), 212-217.

Escudero, C., González, S., & García, M. (1999). Resolucion de problemas en el aula de fisica: un analisis del discurso de su enseñanza y su aprendizaje en nivel medio. [Article]. *Investigações em Ensino de CiênciaS, 4*, 229-251.

ESSI-SCOPE. (1991). ISO 9126: The Standard of Reference, Retrieved on 2010 from http://www.cse.dcu.ie/essiscope/sm2/9126ref.html

European Commission, Directorate General for Enterprise. (2004). *Innovation Management and the Knowledge - Driven Economy*. Brussels-Luxembourg: ECSC-EC-EAEC.

Ferruzca, M. (2008). *Estudio teórico y evidencia empírica de la aplicación del marco teórico de "Cognición Distribuida" en la gestión de sistemas de formación e- Learning.* (Doctor), Universitat Politécnica de Catalunya, Barcelona.

Freedman, D. (1991). *Statistics* (2nd ed.). New York: London : W.W. Norton.

Garrison, D. R., & Shale, D. (1987). Mapping the boundaries of distance education: Problems in defining the field. *The American Journal of Distance Education, 1*, 7-13.

Gaunt, H. (2008). One-to-one tuition in a conservatoire: the perceptions of instrumental and vocal teachers. [Article]. *Psychology of Music, 36*(2), 215-245. doi: 10.1177/0305735607080827

Giraldo, F., Jimenez, A., Trefftz, H., Restrepo, J., & Esteban, P. (2010). Distance Interaction in Education Processes using a Telepresence Tool. [Article]. *Technological Developments in Education and Automation*, 509-512. doi: 10.1007/978-90-481-3656-8_92

Gordon, E. (1997). *A Music learning theory for newborn and young children* (1997 ed ed.). Chicago: GIA.

The Gordon Institute(2011). The Gordon Institute for Music Learning Web Page, Retrieved 21/10/2011, 2011, from http://www.giml.org/

Gunawardena, C. N., & McIsaac, M. S. (2004). Distance education. In D. Jonassen (Ed.), *The handbook of research on education communications and technology* (2nd. Ed. ed., pp. pp. 355-395). Mahwah, NJ: Lawrence Erlbaum Associates.

Gran Teatre del Liceu (2010). Cultural Network (Anella Cultural), Retrieved on 2010, from http://www.liceubarcelona.cat/

Géant2. (2009). Retrieved on 2009, from http://www.geant2.net/

Heaton, N. O. (1992). Defining usability. *Displays, 13*(3), 147-150. doi: 10.1016/0141-9382(92)90016-k

Heritage, J. (2001). Goffman, Garfinkel and conversation analysis. In M. Wetherall, S. Taylor & S.J. Yates. (Ed.), *Discourse Theory and Practice*. London: Sage.

Hernon, P. (1991). *Statistics : a component of the research process*. Norwood: Ablex.

House, J. (2010). Impoliteness in Germany: Intercultural encounters in everyday and institutional talk. [Article]. *Intercultural Pragmatics, 7*(4), 561-595. doi: 10.1515/iprg.2010.026

i-maestro. (2009). Retrieved on 2009 at http://www.i-maestro.org/

International Organization for Standardization (1991), ISO/IEC 9126:1991.

International Organization for Standardization(1999), ISO 13407:1999.

International Organization for Standardization (2004), ISO/IEC TR 9126-4.

Internet2. (2009). Retrieved on 2009 from http://www.internet2.edu/

Irby, D. J., & Hall, H. B. (2011). Fresh Faces, New Places: Moving Beyond Teacher-Researcher Perspectives in Hip-Hop-Based Education Research. [Article]. *Urban Education, 46*(2), 216-240. doi: 10.1177/0042085910377513

James, P. E. (1970). Comparison of efficiency of programmed video-tape and instruction booklet in learning to operate a desk calculator. [Article]. *Programmed Learning & Educational Technology, 7*(2), 134-139.

Johnson, G., & Skinner, M. (2009). The demands of professional opera singing on cranio-cervical posture. *European Spine Journal, 18*(4), 562-569. doi: 10.1007/s00586-009-0884-1

Jonassen, D., Davidson, M., Collins, M., Campbell, J., & Bannan Haag, B. (1995). Constructivism and Computer-Mediated Communication in Distance Education. Retrieved on April 2009 from http://www.c3l.uni-oldenburg.de/cde/media/readings/jonassen95.pdf
doi:http://www.c3l.uni-oldenburg.de/cde/media/readings/jonassen95.pdf

Jonassen, D. H., Howland, J., Moore, J., & Marra, R. M. (2003). *Learning to solve problems with technology: A constructivist perspective* (2nd. Ed. ed.). Columbus, OH.

Karlsson, J., & Juslin, P. N. (2008). Musical expression: an observational study of instrumental teaching. [Article]. *Psychology of Music, 36*(3), 309-334. doi: 10.1177/0305735607086040

Kenny, D. T., & Mitchell, H. F. (2004). *Visual and auditory perception of vocal beauty: conflict or concurrence?* Paper presented at the 8th International Conference on Music Perception & Cognition (ICMPC8);, Evanston, IL, USA.

Kenny, D. T., & Mitchell, H. R. (2006). Acoustic and perceptual appraisal of vocal gestures in the female classical voice. *Journal of Voice, 20*(1), 55-70. doi: 10.1016/j.jvoice.2004.12.002

Kitchenham, B. (2007). Empirical paradigm - The role of experiments. In V. R. Basili, D. Rombach, K. Schneider, B. Kitchenham, D. Pfahl & R. W. Selby (Eds.), *Empirical Software Engineering Issues: Critical Assessment and Future Directions* (Vol. 4336, pp. 25-32). Berlin: Springer-Verlag Berlin.

Kitchenham, B., Brereton, P., Turner, M., Niazi, M., Linkman, S., Pretorius, R., . . . Ieee. (2009). The Impact of Limited Search Procedures for Systematic Literature Reviews - A Participant-Observer Case Study *Esem: 2009 3rd International Symposium on Empirical Software Engineering and Measurement* (pp. 337-346). New York: Ieee.

Kodály Music Education Institute of Australia, K. (2011). Kodály Music Education Institute of Australia Web Page Retrieved 06/10/2011, 2011, from http://www.kodaly.org.au/

Kshirsagar, A. M. (1983). *A Course in linear models*. New York, NY etc.: Marcel Dekker.

Lange, D. N. (1971). Application of social learning theory in affecting change in a group of student teachers using video modeling techniques. [Article]. *Journal of Educational Research, 65*(4), 151-&.

Larsen, R. E. (1992). Relationship of learning style to the effectiveness and acceptance of interactive video instruction. [Article]. *Journal of Computer-Based Instruction, 19*(1), 17-21.

Larson Guerra, S. (2010). *Pensar el sonido: una introducción a la teoría y la práctica del lenguaje sonoro cinematográfico* (1a. Ed.). México D.F.: Universidad Nacional Autónoma de México.

Learning Object Metadata. (2011). Learning Object Metadata, Retrieved on 2011, from http://ltsc.ieee.org/wg12/

Learning Technology Standards Committee. (2010). *Learning technology standards committee website*, Retrieved on 2011, from http://ltsc.ieee.org/

Lustigova, Z., & Lustig, F. (2009). Remote and Open Laboratory in Science Education - Technological, Educational and Psychological Issues. [Article] *Proceedings of the Iti 2009 31st International Conference on Information Technology Interfaces*, 415-420. doi: 10.1109/iti.2009.5196119

Manhattan School of Music. (2009a). Distance Learning Web Page, Retrieved on 2009 from http://dl.msmnyc.edu/

Manhattan School of Music. (2009b). Retrieved on 2009 from http://www.msmnyc.edu/special/distancelearning/

Marcandella, E., Durand, M. G., Renaud, J., & Boly, V. (2009). Past Projects Memory: Knowledge Capitalization from the Early Phases of Innovative Projects. [Article]. *Concurrent Engineering-Research and Applications, 17*(3), 213-224. doi: 10.1177/1063293x09343824

Martins, S., & Martins, I. (2010). Videoconference in education at U.PORTO. [Article]. *Proceedings of the 9th European Conference on E-Learning, Vol 1*, 351-361.

McGill Ultra-videoconferencing, R. G. (2009). Retrieved on 2009 from http://ultravideo.mcgill.edu/

Media College. (2010). Condenser Microphone, Retrieved on 2010, from http://www.mediacollege.com/audio/microphones/condenser.html

Merril, D. (2002). First principles of instruction. Educational Technology Research and Development. Retrieved from http://mdavidmerrill.com/Papers/papers.htm doi:http://mdavidmerrill.com/Papers/papers.htm

Mitchell, H. (2005). *Defining vocal quality in female classical singers: pedagogical, acoustical and perceptual studies.* University of Sydney, Sidney. DOI: http://hdl.handle.net/2123/710

Mitchell, H. F., & Kenny, D. T. (2004). The impact of 'open throat' technique on vibrato rate, extent and onset in classical singing. *Journal of Voice, 29,* 171-182.

Mogey, N. (2009). So, you wan to use Likert-Scales? , Retrieved on 2009, from http://www.icbl.hw.ac.uk/ltdi/cookbook/info_likert_scale/index.html

Moravcsik, M. J., Dorati, A., Rosenbluth, D., & Moravcsik, F. (2002). *Musical sound : an introduction to the physics of music*. New York: Kluwer Academic-Plenum Publishers.

Morrison, S. J., & Demorest, S. M. (2009). Cultural constraints on music perception and cognition. *Prog Brain Res, 178*, 67-77.

Multimedia's Lab from Barcelona Tech. (2012). PhD Web Site, Retrieved on 2012 from http://www.i-cell.net/

New World Symphony (2009). Retrieved on 2009 from http://www.nws.org/AcademyInternet2.asp

Open College of the Arts. (2012). Open College of the Arts Web Page, Retrieved on 2012 from http://www.oca-uk.com/

Open Drama. (2012). Open Drama Web Page, Retrieved on 2012, from http://mtg.upf.edu/opendrama/

Parbery-Clark, A., Skoe, E., & Kraus, N. (2009). Musical experience limits the degradative effects of background noise on the neural processing of sound. [Article]. Research Support, U.S. Gov't, Non-P.H.S.]. *J Neurosci, 29*(45), 14100-14107.

Pierce, B. C. (2002). *Types and programming languages*. Cambridge [etc.]: MIT Press.

Polycom. (2012). Official Web Site, Retrieved on 2012, from www.polycom.com

Prevignano, C., Thibault, P. J., & ebrary Inc. (2003). Discussing conversation analysis, the work of Emanuel A. Schegloff *E-Libro* (pp. 192 p.). Retrieved from http://0-site.ebrary.com.medina.uco.es/lib/bibliocordoba/Doc?id=10032042

Purao, S., Baldwin, C., Hevner, A., Storey, V., Pries-Heje, J., & Smith, B. (2008). *The Sciences of Design: Observations on a Emerging Field*. Harvard Business School.

Rapoport, E. (1996). Emotional expression code in opera and lied singing. [Article]. *Journal of New Music Research, 25*(2), 109-149.

Ravenscroft, A. (2001). Designing E-learning Interactions in the 21st Century: revisiting and rethinking the role of theory. *European Journal of Education, 36*(2), 133-156.

Regelski, T. A. (2005). Music and Music Education: Theory and praxis for 'making a difference'. *Educational Philosophy and Theory, 37*(1). doi: 10.1111/j.1469-5812.2005.00095.x

Reigeluth, C. M. (1999). *Instructional design theories and models: A new paradigm of instructional theory* (Vol. II). Mahwah, NJ: Lawrence Erlbaum Associates.

Rigden, J. S. (1977). *Physics and the sound of music*. New York, etc.: John Wiley & Sons.

Roers, F., Murbe, D., & Sundberg, J. (2009). Predicted Singers' Vocal Fold Lengths and Voice Classification - A Study of X-Ray Morphological Measures. [Article]. *Journal of Voice, 23*(4), 408-413. doi: 10.1016/j.jvoice.2007.12.003

Rojas-Rajs, T., Alpiste, F., Lorente, P., Iglesias, F., Fernández, J., & Monguet, J. (2009). *Effectiveness of Telepresence Learning Environment for Opera Singing: First Results from a Case Study.* Paper presented at the World Conference on E-Learning in Corporate, Government, Healthcare, and Higher Education 2009, Vancouver, Canada.

Rubin, J. (1994). *Handbook of usability testing : how to plan, design, and conduct effective tests*. New York: Wiley.

Sabadell City Council. (2011). Ciutat de la Música de Sabadell, Retrieved on 2011 from http://www.sabadell.es/Ciudaddelamusica/p/ciutatmusica_esp.asp

Schank, R. (1997). *Virtual Learning, A Revolutionary Approach to Building a Highly Skilled Workforce*: McGraw Hill.

Schank, R. C., Berman, T. R., & Macperson, K. A. (1999). Learning by doing. In C.M. Reigeluth(Ed.), *Instructional design theories and models: A new paradigm of instructional theory* (Vol. II). Mahwah, NJ: Lawrence Erlbaum Associates.

Schegloff, E. A. (1988). Goffman and the Analysis of Conversation. In P. D. a. T. Wootton (Ed.), *Erving Goffman: Exploring the Interaction Order* (pp. 9-135). Cambridge: Polity Press.

Schegloff, E. A. (2007). *Sequence organization in interaction a primer in conversation analysis I*. Cambridge, UK: Cambridge University Press.

Shenai, M. B., Dillavou, M., Shum, C., Ross, D., Tubbs, R. S., Shih, A., & Guthrie, B. L. (2011). Virtual Interactive Presence and Augmented Reality (VIPAR) for Remote Surgical Assistance. [Article]. *Neurosurgery, 68*, 8. doi: ons20010.1227/NEU.0b013e3182077efd

SITOLA. (2009). Ultragrid , Retrieved 10-10-2009, from http://www.sitola.cz/wordpress/projects/multimedia-distribution-and-processing/ultragrid/en/

SPSS. (2009). SPSS Help for 11.0 version.

Suzuki Books (2011). Why Is The Suzuki Method So Popular? Retrieved 21/10/2011, from http://www.suzukibooks.com/

Stevens, S., & Warshofsky, F. e. a. (1981). *Sonido y audición* (2nd. ed.). México D.F.: Time Life Internacional de México.

The Dalcroze Society of America. (2011). Dalcroze Society of America Web Page Retrieved 05/10/2011, Retrieved on 2011, from http://www.dalcrozeusa.org/

Torres Mulas, J., Gallego, A., & Álvarez, L. (1976). *Música y sociedad*. Madrid: Real Musical.

Tukey, J. W., Hoaglin, D. C., & Mosteller, F. (1985). *Exploring data tables,trends and shapes*. New York [etc.]: John Wiley & Sons.

UK, D. S. o. (2011). The Dalcroze Society Retrieved 18/10/2011, 2011, from http://www.dalcroze.org.uk/indexb.html

Vaccaj, N. (1990). *Metodo practico di canto*. Milano: BMG Ricordi Music Publishing.

Virtual European Music School, VEMUS (2009). Retrieved on 2009, from http://www.vemus.org/

Web of Knowledge. (2012). End Note Web, 2012, from http://www.myendnoteweb.com/

Wikipedia. (2009). Microphone Retrieved November, 2009, from http://en.wikipedia.org/wiki/Microphone

Wikipedia. (2011). Music education Retrieved 06/10/2011, 2011, from http://en.wikipedia.org/wiki/Music_education

Wikipedia. (2012). Percentiles, Retrieved on 2012, from http://en.wikipedia.org/wiki/Percentile

Wiley, D. A. (2000). Connecting learning objects to instructional design theory: A definition, a metaphor, and a taxonomy.The Instructional Use of Learning Objects: Online Version. Retrieved from http://reusability.org/read/chapters/wiley.doc

Woody, R. H. (2006). Musicians' cognitive processing of imagery-based instructions for expressive performance. [Article]. *Journal of Research in Music Education, 54*(2), 125-137.

Wright, P., & Bason, G. (1983). Detour routes to usability - a comparison of alternative approaches to multipurpose software-design. *International Journal of Man-Machine Studies, 18*(4), 391-400.

Yin, R. K. (2003). *Case study research : design and methods* (3rd ed ed.). Thousand Oaks, Calif.: Sage Publications.