FORMA. Representing Space:
Performance, Documentation, and Immersive Archive

Diego Schiavo

Introduction

How does a particular director, choreographer or stage designer use the space? How does a performance differ when it is staged in different places? What constants can be found in different performances that take place in the same environment?

Each performance is nothing other than a particular spatial organtion and exists in a particular space for a particular time, in which it enters into a one-to-one relationship with the space that contains it, defines it, and is defined by it.

The spatial dimension (both visual and acoustic) and its organization are central themes of the performing arts of the last two centuries, but when we try to objectively analyze the relationship between live performance and space, we find that to date there is no audiovisual documentation material capable of reproducing the dynamic three-dimensionality of scenic works.

With "FORMA. Representing Space: Performance, Documentation, and Immersive Archive" – born from the ALMAIDEA research project conceived by Enrico Pitozzi within the Department of Arts of the Alma Mater Studiorum University of Bologna – we have set ourselves the scientific goal of remedying this deficiency and creating a prototype that allows us to explore the possible forms of documentation of the space of a live performance, the first step through which we can trace the coordinates of a new form of archiving live performance based on unpublished and innovative material. The whole project aims to develop a method for documenting the (visual and sonic) space of compositional forms, highlighting aspects that are not yet documented and therefore cannot be investigated.

The prototype (and thus the resulting archive track) is conceived in this first phase of development for scholars and as teaching material to provide an innovative tool capable of stimulating new reflections and

Mimesis Journal, 13, 2 | 2024, pp. 65-78 https://doi.org/ 10.13135/2389-6086/9933
ISSN 2279-7203

opening up new fields of investigation by restoring space to its central importance in understanding the compositional process.

Because there is no protocol for the three-dimensional filming of live shows, it was crucial to define one to establish the technical solutions and methodology for the creation of the prototype, thus laying the foundations for the creation of a new archive in which to view the show's documentation. This is also the opportunity to reflect on the documentation of live performances and the study of their creation processes, in an unprecedented way and in line with current technological developments.

The Subject of the Investigation: Space and Performance

To adequately document the spatiality – around which the scientific urgency of this project revolves – the person setting up the technical documentation apparatus must be able to understand the creative process of the artist with whom he is collaborating and thus precisely define the entire compositional process of the work. Doing so starts from the characteristics of the physical site, where the performance takes place, the topographical and geometrical space (and, therefore, also the resonance planes of the same) in which the performers and the compositional rules operate, so as to differentiate the performance space from the audience area and the urban context in which they are located, highlighting all the different types of interactions.[1]

A show can be staged end-on or in-the-round, or the audience can be immersed in the scene and share the space with the performers. For example, the performance could take place in an "all'italiana" theater and be performed on a proscenium stage, or it could cross this boundary and also take place in the audience or, again, involve the space of the boxes.[2] The same applies to the sound space, which can be frontal or surrounding, both on stage and in the audience, allowing different listening levels that then have to be represented in the documentation.

Each of these alternatives would change the space and thus also the technical requirements for documenting everything that happens in the "performance space." The identification of this space is therefore of fundamental importance, because it strongly influences the choice of recording

[1] The public space, the theater space, and the urban context that contains them can be the subject of further documentation. See the later subsection, "Analyzing the Data."
[2] Think, for example, of the show *Boris Godunov* by Fura del Baus, which involves the entire theater: stage, audience, and boxes included.

techniques and, therefore, the re-creation of the creative process that takes place in this space. If the identification of the area to be filmed and recorded is complex, the film and sound directors must coordinate with those responsible for the staging so as to define it appropriately.

The possibility of documenting the spatiality of a work implies corporeality in an equally crucial way: The live performance is the place where audience and performer meet, and it is this space that determines the nature of their shared experience (McAuley 1999, 26). Within this space, the body of the performer (whether a singer, instrumentalist, actor, dancer...) is an entity that simultaneously acts and is acted upon. And in the body, in its relationship to space, in its posture, in its position, and in its ability to produce sound, we find another key component of what has to be recorded and preserved (Pitozzi 2015 and 2016a). When we refer to the subjective component in acting practice, we are dealing only with the senses and: "[...] even the choice – of a gesture, of a meaning attributed to that feeling, of a certain body movement, etc. – although it is necessary to give the action a certain objectivity, this is nevertheless the moment in which the subjectivity of the performer manifests itself" (Giraldo 2013, 263). Of course, this applies not only to acting but to any performing artist in general.

The object of the documentation is therefore not only the space in its physical and acoustic complexity, but also the corporeality of the performer acting in it.

The Point of View

In audiovisual products, the compression of three-dimensional space into a two-dimensional surface confines the viewer to a single viewing angle. It is not possible to move around in a filmed scene, and it is not possible to view the images produced on the screen from angles other than that of the camera. In his "Film as Art," Arnheim (1957, 11-16) already notes that cinema, despite its ability to capture movement and present a sequence of images, inevitably flattens the three-dimensionality of the world to the two-dimensional surface of the screen. The point of observation that a film proposes is a single one, that of the director, embodied first by the camera and then by the editing. The director thus not only orchestrates the events of the production, but also reduces the multiplicity of perspectives to a single possible point of view. This structural limitation becomes even clearer in the attempt to create an effect of depth in 3D cinema because, as Walter Murch said in an interview with the American film

critic Roger Ebert (2011): "3D films remind the audience that they are in a certain 'perspective' relationship to the image."

In the world of gaming, on the other hand, many technical solutions have emerged to enhance the home entertainment experience by making it more immersive. The advent of 3D audio-video systems and the use of viewers and various multi-angle features have given gamers a seeming freedom of viewing (and sometimes movement) that movies never had.[3] This shows how home gaming technology has broken through and overcome the internal rigidity of the audiovisual sector. This could be an excellent starting point to suggest which recording method should be used.[4]

The Multiplication of the Gaze

After a careful analysis of different types of immersive capture technologies and inspired by the multi-angle PIP features in games, it was decided to base the prototype on multi-camera capture. Multi-camera (or multi-angle) recording refers to recording an event with multiple cameras from different positions at the same time. With this technique, multiple angles can be captured simultaneously, resulting in a richer and more dynamic recording that is particularly useful in situations where there is a lot of activity happening at the same time. It is a standardized, consolidated, and readily available technique that does not overly burden the stage or the performers. It offers the possibility of filming at a high number of frames per second and has an excellent resolution.[5] It therefore offers the best conditions for accurately documenting the positions, facial expressions, and overall physicality of the performers.

The use of a standard technology enables modular film configura-

[3] Various forms of view multiplication are being tested in games, such as split screen, PIP modes (picture-in-picture), spectator mode, replay with free camera, open world virtual reality, and others.

[4] In Greuter, Mulvany, and Myers 2023, an approach and even an engine from the video game industry is used in a similar way. The results are extremely interesting when it comes to the immersive archiving of a show. But the archived show was already inherently immersive and the audience shared the space of movement and experience with the actors. The same technical solution would not have been as effective with a show developed in front of the audience. Further, the extremely sophisticated and advanced recording techniques did not conform to the principle of minimal invasiveness to the set and performers that was adopted in the planning phase of this work.

[5] The standard frame rate for videos in Italy is 25 frames per second (fps). Prosumer models easily reach 120 fps. Higher-end devices reach 960 fps, which is necessary to significantly slow down the playback speed while maintaining a smooth image.

tions that meet the requirements of different staging possibilities and the ability to arrange multiple cameras on extremely irregular perimeters. In light of recent developments in information technology (IT), the recorded data, precisely because they conform to a standard, also offer great opportunities for artificial intelligence processing[6] to meet various research needs and integrate seamlessly into future IT innovations.

The Point of Listening

While on the visual level it is always clear (with varying degrees of complexity) what should be framed and what should not, it is not quite so clear in the audio field. During a performance, several sound spaces with completely different characteristics and purposes coexist, and each of them offers ideal points of listening that are completely different from each other. It is therefore crucial to identify these points of listening to decide what should be recorded and how (Pitozzi 2016b, 256-288).

In creating a meaningful listening experience, we mean by "point of listening" the spatial relationship between the listener and the other sound objects around that listener, that is, as Chion says, "the spatial position from which we can say that we hear a sound at a certain distance from us" (Chion 2009, 485). The ideal point of listening is an average point that is representative of all possible points of listening of a particular room: In a hi-fi system, the ideal point is in the middle between the two loudspeakers. In a cinema, it is the central seat in the auditorium.

During a live show, different sound spaces and, therefore, different points of listening exist side by side. Consider, for example, a dramatic text performed by actors in an Italian theater without pre-produced music or vocal amplification. In the audience, the voices from the stage are heard together with the sounds of their movements. Added to this are the sounds produced by the audience. Everything is subject to the natural reverberation produced by the acoustics of the venue.

On stage, however, the perception of sound is completely different: Due to the different acoustics, the sounds are drier, and the reverberation comes from the side of the audience, as do the noises that the audience itself produces. Some sounds that the audience can barely hear, if at all, are much clearer on stage, such as the prompter's voice or the sounds of the other actors and technicians behind the scenes.

[6] See later section entitled "Future Developments."

We are therefore dealing with two completely different acoustic experiences.

Let us now consider the case of a show with pre-recorded music and sound design. The amplification system in the auditorium is designed and calibrated to meet the specific requirements of the director and the narrative. Unlike the spy monitors on stage,[7] their purpose is to provide the performers with the elements required for their performance,[8] and these elements do not necessarily match the sound intended for the audience.[9] Since the sound is pre-recorded, it exists independently of its distribution in the room and is acoustically different from what the audience hears. It is therefore a third case that is different from the others. And if a recording were made for a later television broadcast of the show, the sound that would be sent to post-production[10] would be that coming from the mixing desk with separate channels for voices, sounds, music, and audience overhead microphones.

Therefore, four sound spaces coexist in the same show:

1. an acoustic space perceived by the audience;
2. an acoustic space that can be heard from the stage;
3. a pre-recorded soundtrack without voices, sounds, or noises; and
4. a soundtrack coming from the mixing desk and containing the pure voices, the pure sound tracks, and the tracks with the miking of the sound heard in the audience.

To investigate and analyze how the acoustics of a place affect the performer, space 1 above is certainly preferable. However, this is quite different from what the audience hears in the room, which is represented by acoustic space 2. If you want to study the sounds and music as they were intended by the composer, it is more appropriate to listen to the original soundtracks as in sound space 3. In the end, the stated sound in sound

[7] The loudspeaker system positioned and calibrated to serve the technicians and performers.
[8] Some sounds may be missing or balanced differently, or there may be service sounds such as metronomes and counters that should not be played in the audience but are used by the performer to coordinate the performance.
[9] If, for example, the studio soundtrack were to be transferred and recorded in the room and then compared with the original track, it would be found that the newly recorded track contains the reverberation of the environment and the sounds of the audience, elements that are not present in the original track.
[10] The assembly and optimization of the audio and video material that follows the filming phase.

space 4 is perhaps the most complete, balanced, and pleasing yet, at the same time, the most artificial and least like the live action.

As with the point of view, also with the point of listening it is necessary to take into account the complexity of the sound spaces that coexist in a live show and try to capture them in their entirety so that the archive user can make a complete and thorough investigation. The best approach, therefore, seems to be to capture and archive all four sound spaces when they are present.

The Technical Protocol

Because live shows vary greatly, the technical solutions for their documentation must be carefully considered. To do justice to this diversity, it is important that recording methods allow the recording operator a certain degree of autonomy. The material produced must not be too heterogeneous, because it must be comprehensible, archivable, retrievable, and usable in the database. The development of a technical protocol shared by all those who produce material for an archive is a fundamental step in its creation.

Based on the typical case of a stage placed in front of the audience, the minimum configuration of the video recording consists of four camera points: one frontal, two lateral (one on the right and one on the left) and one from above.[11] For an in-the-round performance, the minimum configuration increases to five cameras, with one added on the fourth side.

This ideal minimum configuration should be adapted according to the needs of each event: The number of cameras must be sufficient to cover the entire stage area in all its parts. Each shot must capture the entire scenic space from the intended angle. Multiple cameras can be placed side by side to capture longer or more complex spaces,[12] avoiding distorting or uneven optics.

Filming should be as non-invasive as possible so as not to change the show. The lighting should also not change as a result of filming compared to normal staging, because the sensitivity and performance of the human eye and camera sensors are completely different. Consequently,

[11] Minimum recording format: H264 or H265 codec; 1920 x 1080; 25 fps; recommended: RAW, ProRes 444, ProRes 422; 4096 x 2160; 60 fps.

[12] It is necessary to provide a technical data sheet with a map to understand the actual layout of the audio and video recording systems.

minimal adjustments may be necessary to avoid falling below a certain quality standard of the image, which would affect the outcome of the documentation.

With regard to recording sound, there are infinite possibilities to use different microphone techniques to obtain a detailed multi-channel image of the two main listening points that make it possible to understand the sound spaces of a live show (the stage and the audience). The recordings[13] made by a professional sound engineer must be differentiated by acoustic space and have a minimum format of quadraphony (or double stereo) for each space. It is advisable, if available, to also acquire the sound from the mixer in the format in which it is to be broadcast,[14] as well as the original sound design and the music tracks.

Acquisition

Once we had defined the scientific purpose – documenting the visual and acoustic space of a live event – and the technologies and practices to be used during the recording phase, we moved on to the field-testing phase. The recording took place on November 25, 2023, in the "stanza delle pratiche" of Palazzo Malagola in Ravenna. The participants were the choreographer and dancer Simona Bertozzi and the author, actress and singer Mirella Mastronardi. Simona Bertozzi performed two fragments from the 2016 work *Anatomy* with a sound composition by Francesco Giomi, while Mirella Mastronardi performed a series of songs from her repertoire in Aramaic and Farsi, followed by a passage from the texts of Yiannis Ritsos, from a project that the actress herself is developing with the Studio Azzurro team.

The intrinsic differences among the three sections of the work (a dance performance, a staged reading, and a song) provided the opportunity to experiment with the documentation and analysis of three different uses of space – the fragmentation of choreographic gesture in the dance performance, the exploration of acoustic space and mood during the vocal chant, and the relationship among movement, gesture, and voice in the actress's work.

The three sections took place in the same room, which they shared with the audience, who could stand on three of the four sides of the room. The

[13] Minimum recording format: .wav 48 KHz, 24 bit; recommended: .wav 96 KHz, 32 bit.
[14] Stereo if it only comes from the LR channels of the mixer, 5.1 if it is transmitted into the room in this format, and so on.

Panasonic Lumix S5 MKII recording cameras were positioned on the same three sides, with another lighter camera (dji osmo pocket3) on the central wooden beam. The presence of the cameras limited the performance space, because shorter focal lengths would have led to excessive perspective deformation. As the audience and performers shared the same space, the auditorium and performance space were acoustically combined so that only a single ambient recording had to be made. Three-directional microphones[15] were used in a coaxial position to the cameras along with an Ambisonic-A format multi-capsule microphone[16] near the zenith camera to maximize definition for each possible view and achieve good definition for the Dolby Atmos mix.

After the recordings were completed, 4 video files - one for each machine point - and multi-channel audio files for each performance were created. In the post-production phase, the videos were aligned and colour-corrected (on a technical level) using DaVinci Resolve to make the different angles appear natural and uniform. The sound synchronized with the video was levelled between the different microphones and mixed in both multichannel and Dolby Atmos formats, using Avid ProTools, to reproduce the spatiality of the location as faithfully as possible and to accommodate different possible playback contexts – from stereo headphones to professional multichannel systems.[17]

Starting a Prototype Archive

Data Entry and Cataloguing

After the recording and post-processing phase, all materials must be catalogued. First, the data entry interface is programmed – using Next.js as the framework and Typescript as the programming language[18] – to catalogue the files of each show through a series of metadata.[19] The metadata of the files contain a variety of information such as the name of the performance and an identifier for the cameras (right camera, front,

[15] Model Sennheiser MKH 416.
[16] Model Rode NT-SF1.
[17] Both formats provide compatibility with lower order formats up to mono.
[18] This is a typed version of Javascript. The same methods are also used for the development of user interfaces.
[19] Metadata are data that describe other data and provide information about properties, structure, formatting, and meaning. They become a set of data or resources required for organizing, categorizing and facilitating the search and use of digital resources.

left, top, front left audio, back left, and so forth).[20] Some information specific to this type of material has been added, such as the date and location of the recording and the technical details of the performance such as the names of the performers, production, company, director, authors, set, music, costumes, and lighting.

Special Contents

The artists' video interviews can be integrated as interactive overlay cards to specific timecodes to provide a further analytical and educational tool. Furthermore, there is a folder for additional content related to the main performance, which may contain the following: Rehearsal material, interviews, backstage, and text files with director's notes, script, credits, and information about the artists. In addition, information about the physical location where the performance takes place (be it a theater or other architectural space and the geographical area in which it is located) can be stored: Plans of the space, maps, or geotags tracing the location of the performance in the urban context, sound recordings of the acoustic environment,[21] impulse responses of the audience and stage reverb and associated maps, and 3D scans of the stage space, audience and technical environment.

User Platform

As for the user interface,[22] the goal is to create a server-side rendered (SSR) Web site that utilizes single page application (SPA) principles for dynamic content updates and provides a comprehensive full stack application experience.[23]

Querying the database via the search field (with title, actor, author, director, location, or other previously entered metadata) loads the videos with the corresponding criteria. Selecting a video starts playback of the performance. It is possible to select the viewing angle and switch from

[20] Based on the international Dublin Core standard, the system adopted by producers, authors and rights holders to provide tools for accessing digital resources.

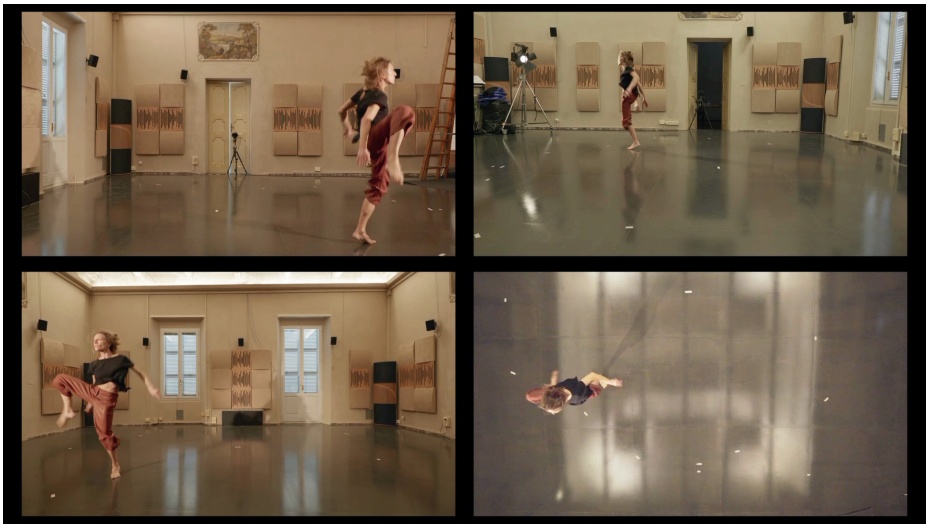
[21] Think of open-air theaters by the sea or in a city park: Although hypothetically alike in form, they are embedded in different urban and soundscapes.

[22] Compatible with the HTML5 standard, so it can be used with virtually any device and browser.

[23] SSR Web sites render the initial page on the server, improving performance and SEO; SPAs load a single HTML page and dynamically update content as the user interacts with the app; full-stack applications combine client-side and server-side code within a single project, leveraging Next.js's API routes.

one to another. The interface also allows the user to change the listening point, fast-forward and rewind playback, export snapshots, and log in to specific content linked to the video. The ability to navigate through the performance allows the positions of the performers, their actual use of the space (both physically and acoustically), and their posture and muscle tensions to be closely examined from every angle, opening up new possibilities for investigation. See Figure 1.

Figure 1

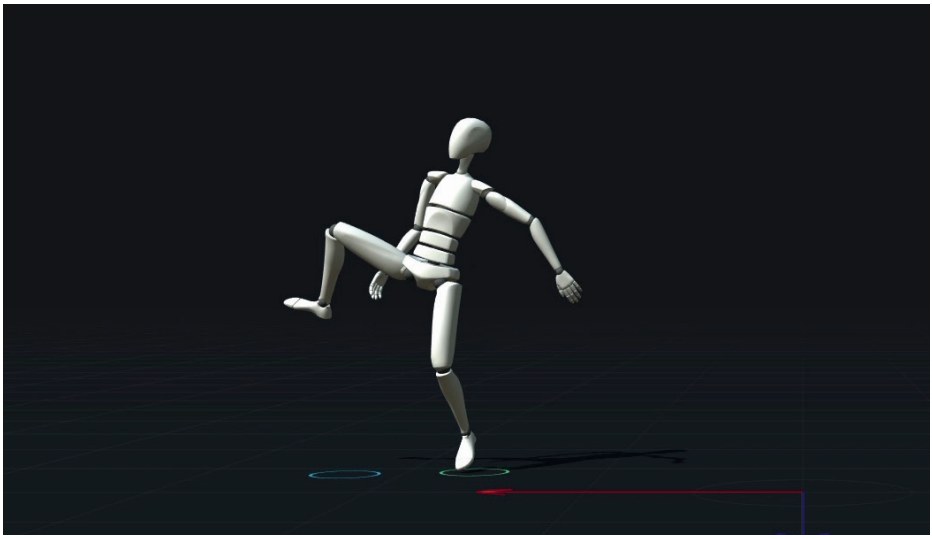


Analyzing the Data

The export of images related to the same moment and viewed from different angles, as well as the synoptic viewing of the images, makes it possible to create a navigable 3D model of this moment of the performance with the help of AI. Again thanks to AI analysis, it is possible to follow the diagram of a performer's movement in space and time or to summarize the performance in a motion tracking of an abstract model. See Figure 2.

It is also possible to choose the acoustic space you want to immerse yourself in (stage, audience, sound recorded in the studio) and move through the sound space with headphones or play the sound through a multi-channel hi-fi system.

Figure 2



Future Developments

The prototype shown here is intended to represent the first proposal for a new method of documenting an artist's creative process, based on the audiovisual reconstruction of the spatial coordinates around which the documented work is created. Placing the emphasis on the prototype, as we do here, implies the first concrete act of defining an immersive archive of live shows, capable of producing and collecting navigable material whose "active" use allows one to move from different angles within the representation to observe, listen to, and study aspects that would be (rather) elusive in a "classical" recording, thus renewing the processes of analysis of the works.

The creation of the technical protocol is the instrument with which the beta test phase is initiated. In this phase, different service providers can use the protocol to record different shows in different locations, thus expanding the catalogue and effectively starting the creation of the first immersive archive of live broadcasts.

This trial phase is important for testing and perfecting both the technical protocol itself and the functionality of the data input and the user interface.

With a view to the future development of this project, the innovations offered by artificial intelligence can be used to test new data-processing

techniques, such as converting the footage of a live show into a 3D version that is fully navigable via a viewer that allows the grantee to move around the performance space while playing the archived performance.

Works Cited

Arnheim, Rudolf. 1957. *Film as Art*. Berkeley: University of California Press.

Bardiot, Clarisse. 2017. "Organiser et conserver la mémoire de l'éphémère: les capsules de MemoRekall." *Culture et Musées* 30, no. 2: 159-174.

Baudry, Jean-Louis, and Alan Williams. 1974-1975. "Ideological Effects of the Basic Cinematographic Apparatus." *Film Quarterly* 28, no. 2: 39-47.

Bazin, André. 1967. *What Is Cinema?* Edited by Hugh Gray. Translated by Hugh Gray. Berkeley: University of California Press.

Bleeker, Maaike, ed. 2017. *Transmission in Motion: The Technologizing of Dance*. London: Routledge.

Chion, Michel. 2009. *Film, a Sound Art*. New York: Columbia University Press.

Dolby Inc. 2015. "Dolby Atmos Specifications." Dolby Professional. https://professional.dolby.com/siteassets/cinema-products–-documents/dolby-atmos-specifications.pdf.

Ebert, Roger. 2011. "Why 3D doesn't work and never will. Case closed." Movie reviews and ratings by Film Critic Roger Ebert. https://www.rogerebert.com/roger-ebert/why-3d-doesnt-work-and-never-will-case-closed.

Gill, Tony. 2008. "Metadata and the Web." In *Introduction to Metadata: Pathways to Digital Information*, edited by Murtha Baca, 20-37. Los Angeles: Getty Information Institute.

Giraldo, Mabel. 2013. "La pratica attoriale in Richard Schechner: dalla sensazione all'essere. Sulle tracce della fenomenologia di Merleau-Ponty." *Itinera* 5: 256.

Greuter, S., Mulvany, G., and Myers, M. 2023. "The immersive archive." *International Journal of Performance Arts and Digital Media* 19 no. 3: 405–423. https://doi.org/10.1080/14794713.2023.2223747.

McAuley, Gay. 1999. *Space in Performance: Making Meaning in the Theatre*. Ann Arbor: University of Michigan Press.

Pitozzi, Enrico. 2022. "From Imagination to Preservation of Movement: Digital Tools for Contemporary Dance." In *The Oxford Handbook of Dance and Memory*, edited by Susanne Franco and Marina Nordera. Oxford: Oxford University Press.

_____. 2016a. "Aux bords de l'audible : une étude des théâtres du son." Edited by Yana Meerzon. *Recherches sémiotiques Semiotic Inquiry* 36 - La sémiotique du son : vers une architecture de l'acoustique et de l'auralité dans le théâtre post-dramatique, II nn. 1-2: 223–254.

_____. 2016b. "Bodysoundscape. Perception, movement and audiovisual

developments in contemporary dance." In *The Oxford Handbook of Sound and Image in Western Art*, edited by Yael Kaduri, 256-288. Oxford: Oxford University Press.

_____. 2015. *Choreographic composition of Cindy Van Acker*. Macerata: Quodlibet.