Acoustic Comfort for Spaces Used by People with Cognitive Impairment: A Starting Point for the Application of Acoustic Event Detection and Sound Source Recognition Systems

Federica Bettarello – University of Trieste, Italy – fbettarello@units.it
Marco Caniato – Free University of Bozen-Bolzano, Italy – marco.caniato@unibz.it
Giuseppina Scavuzzo – University of Trieste, Italy – gscavuzzo@units.it
Andrea Gasparella – Free University of Bozen-Bolzano, Italy – andrea.gasparella@unibz.it

Abstract

The AED (Acoustic Event Detection) and SSR (Sound Source Recognition) systems are increasingly used in projects involving home automation, security, help for the visually impaired or for the elderly who want to pursue projects of independent living. The application of such technologies can also become a valid reference in the case of subjects suffering from cognitive, not just physical, deficits (Down syndrome, autism, etc.). In these cases remote assistance systems can represent a strong support both for the people who take care of them (parents, specialized personnel), and also for people with cognitive disabilities to pursue projects of independent living. Based on the study of the peculiarities that the internal spaces hosting activities for people with various cognitive deficits must have, criteria for acoustic comfort and internal design have been optimized for certain types of living, working and resting spaces. The aim of this work is to understand how the indoor sound field of dedicated rooms may affect boundary conditions for the installation of AED and SSR and how the specific interior design for special spaces may influence the speech intelligibility and clarity of these rooms.

1. Introduction

Acoustic Event Detection systems (AED) are designed to detect anomalous acoustic events and were developed to assist video surveillance systems (Pagin, 2016). They are based on the peculiarities of a sound signal that can be composed of very different time or spectral characteristics, different amplitude levels (consequently, different signal-to-noise ratio) and different duration. The detection algorithms are based on pre-processing techniques necessary to detect potential alarm signals, such as the sensitivity of the detection algorithm (i.e. the ability to detect irregularities whose level is more or less close to the background noise) and the adaptability to sudden or slow changes in the signal (with respect to the variability of the background noise) (Dufaux, 2001; Tronchin, 2013). These systems are mainly used for "security" in outdoor environments or large environments (garages, airports, stations, etc.).

Sound Source Recognition systems (SSR) are more recent devices, which have allowed the passage from the request for help through a "button device" to the immediate sending of a distress call through the use of one's own voice or the recognition of a specific event. These technologies are becoming increasingly popular today, especially thanks to the development of home automation techniques within the Smarthome, i.e. those spaces that use a home controller to integrate the various home automation systems (Robles and Ki, 2010; Tronchin and Coli, 2015).

Sensors can provide information about a person's posture and movement or detect a fall. Smarthomes are therefore useful for measuring a person's activity or for helping people with, for example, cognitive or physical disabilities in their daily activities (Fleury et al., 2008).

There have been many projects on acoustic recognition algorithms over the years (Chang and Chang, 2013; Foster et al., 2015; Janvier et al., 2012 ; Lecouteux et al., 2018 ; Yong and Kean, 2014; Zhang et al., 2015). The systems involved use one

Part of
Pernigotto, G., Patuzzi, F., Prada, A., Corrado, V., & Gasparella, A. (Eds.). (2020). Building simulation applications BSA 2019.
bu,press. https://doi.org/10.13124/9788860461766

or more microphones, or are even embedded in robots that allow the subject to be followed very closely. As far as voice command is concerned, these systems are able to analyse and recognize many characteristics, including: intensity, timbre, rhythm, level. This makes it possible to distinguish whether the sending of an alarm may or may not be necessary.

Such systems therefore also allow non-autonomous people to maintain control of their environment and activities, their health, their well-being and their sense of dignity (Portet et al., 2011). Many of the studies presented so far see the elderly as the main stakeholders, in particular at that stage of their life when they are still able to look after themselves but need help to manage some difficult tasks or any situations of risk, danger and accident (Yoshioka et al., 2012).

The problems encountered during the implementation phases of this type of device are related to the background noise (Biagetti et al., 2011) and internal reverberation conditions of the environments (Piana et al., 2014; Yoshioka et al., 2012). It has been shown, for example, that recognition performance decreases significantly as soon as the microphone moves away from the user's mouth (e.g. when positioned on the ceiling) due to increased reverberation and the influence of background noise. The results under such conditions are optimal up to conditions of T_R=0.5 s and with the presence of speech and classical music in the background but not for the noise of technological systems or something similar (Lecouteux et al., 2011). Reverberation can be modelled as additive interference (Kinoshita et al., 2013).

The aim of this work is to understand how the indoor sound field in dedicated rooms may affect boundary conditions for the installation of AED and SSR and how the specific interior design for special spaces may influence the speech intelligibility and clarity of these rooms.

2. Materials and Methods

A case study was chosen in order to understand acoustic proprieties of buildings dedicated specifically to people with cognitive impairment. The AUDI! Project had some stakeholders; one of them was Progettoautismo FVG Onlus, an association from the Friuli Venezia Giulia region (ITA) that supports people with autistic syndrome, Pervasive Developmental Disorders or Asperger's Syndrome and their families.

The selected rooms for possible application of the acoustic sensors were: (i) four different individual therapy rooms (speech therapy, psychomotricity, etc. Fig. 1), (ii) a soft room (multisensory room for the well-being of children and adults with autism, Fig. 2), (iii) an atelier (art workshop, Fig. 3) and (iv) a bedroom in the residential apartments (Fig. 4).

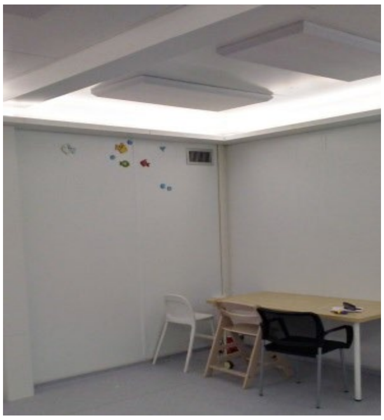


Fig. 1 – Box for individual therapy

Fig. 2 – Soft room

Fig. 3 – Atelier

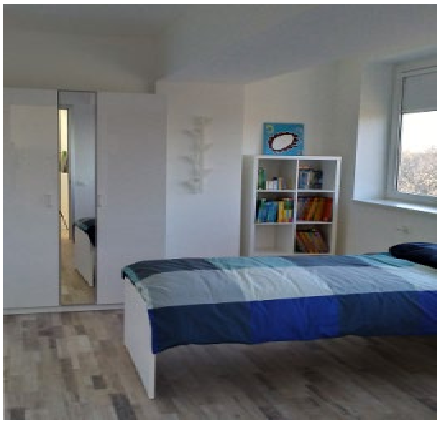


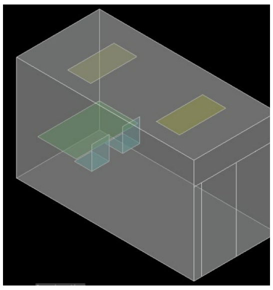
Fig. 4 – Bedroom

The rooms for individual therapy are of different sizes, rectangular in shape, with an entrance door and no outdoor window, with very few interior furnishings (a table and two chairs, few toys and adhesive wall), with ceiling lights and sound absorbing panels. Some therapy rooms also have glass behind them so that relatives or other staff can attend the therapy session. They are made of a prefabricated structure consisting of pre-painted sheet metal/insulating material/pre-painted sheet metal type and are located in an area of the building which is particularly acoustic insulated from external noise and noise coming from inside the other parts of the building. These spaces have to be very quiet, because this is one of the particular needs for users. Here a single operator works with a single patient and it is of paramount importance that nothing may disturb the quiet relationship between operator and patient during the therapy. For this reason, the particular need for voice recognition devices is to recognize a help request from the operator, for example, through the recognition of a keyword, said without alteration of the voice (i.e. without screaming) and without making the patient understand the meaning of this request.
The soft room is located inside the day-centre, equipped with impact resistant layering on the walls (up to about 2.10 m height), resilient flooring and soft cubes arranged in the space, useful both for playing and limiting patients in particular moments of venting and crisis. As for the use in the latter conditions, in which only the patient is left in the room while the operator monitors from outside through a window, the recognition device can recognize events (screams, blows) and activate a process of help request, i.e. alert an additional operator and simultaneously trigger a process of audio-visual support inside the box.
The atelier (or art room) is a very large room, used for collective works, where inside you can find more guests and more operators. Here too, the particular need for voice recognition devices is to recognize a help request from the operator, for example, through the recognition of a keyword, said without altering the voice (i.e. without screaming) as described before.
A typical room in the residential apartments is 9 or 12 m^2 (hosting one or two patients) and has big outdoor windows and a private bathroom. Guests can spend the night alone, while the operator remains in a neighbouring room. The use of the acoustic recognition device becomes helpful to the operator in the next room as they can be instantly warned.
All simulations were performed using 3D acoustic simulators in order to obtain the indoor acoustic field distributions.

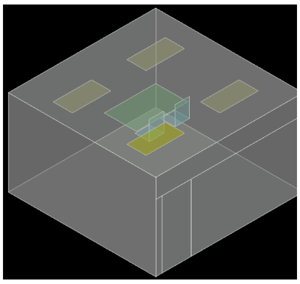
3. Results and Discussion

The applicability of a sound recognition device within a room is related to acoustic parameters such as reverberation time, clarity, definition and background noise. The performance of a speech recognition system in fact decreases significantly with the increase of the distance between the microphone and the source that emits the signal to be recognized, due to the increase in number of reflections and the influence of background noise.
Optimal conditions are represented by T_R=05÷0.7 s, clarity C_{50}>3 dB and definition D_{50}>70% (Lecouteux et al., 2011; Farina and Tronchin, 2013; Kinoshita et al., 2013).
Reverberation has a dispersive effect on the characteristic sequence of speech. The energy ratio between the direct sound part and the first reflections and late reflections is represented by parameters C_{50} and D_{50} and is highly correlated to the performance of speech recognition.
In order to understand the real applicability of the sound recognition systems within the analysed spaces, a campaign of on-site acoustic measure-

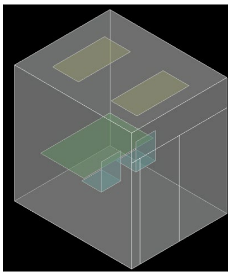
ments and simulations with acoustic predicting software was performed. The on-site acoustic measurements were conducted with impulsive technique in accordance with standards (ISO 3382). The measurements made it possible to analyse the acoustic characteristics in terms of internal reverberation and were used for the calibration of the raytracing model. Fig. 5 shows the images of the rooms recreated for the acoustic simulation program and Fig. 6 shows the results in terms of reverberation time in octaves of frequency, which demonstrate a good agreement between the in situ and simulated results.



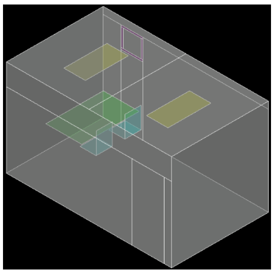
a) Therapy room n. 1



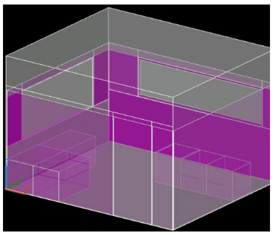
b) Therapy room n. 2



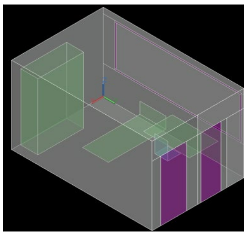
c) Therapy room n. 3



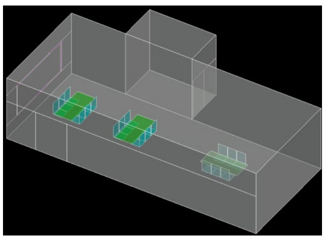
d) Therapy room n. 4



e) Soft room

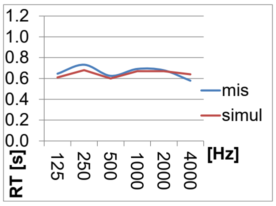


f) Bedroom

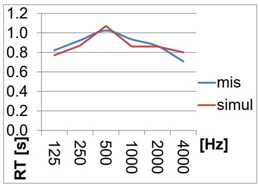


g) Atelier

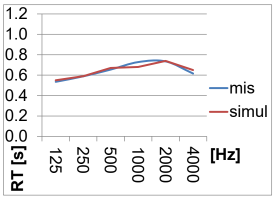
Fig. 5a–g – Images of the rooms recreated for the acoustic simulation program



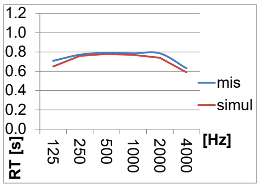
a) Therapy room n. 1



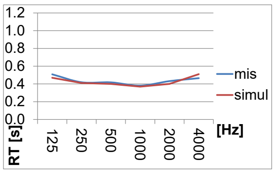
b) Therapy room n. 2



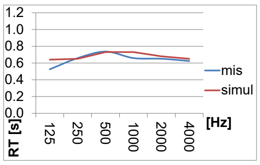
c) Therapy room n. 3



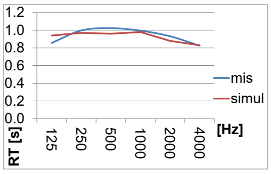
d) Therapy room n. 4



e) Soft room



f) Bedroom

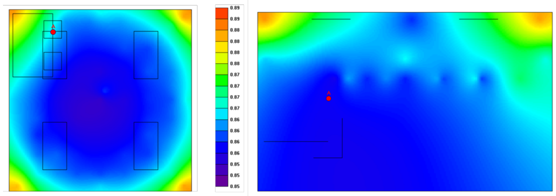


g) Atelier

Fig. 6a–g – Reverberation time octaves frequency comparison: in situ results VS simulated once

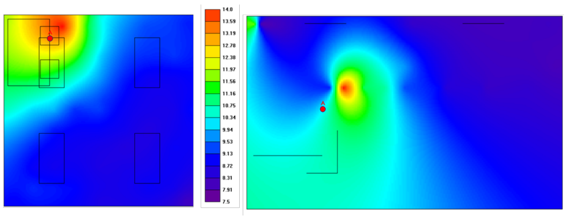
The results obtained show that therapy rooms number 1 and 3 have optimal acoustic conditions for the use of speech recognition devices inside the rooms, while therapy rooms 2 and 4 (larger than the previous ones) need further sound absorbing treatment. The spatial distribution of the reverberation time parameter inside therapy room n. 2 (Fig. 7) shows how the corner positions of the room do not represent the optimal points of dislocation for the devices. It is therefore better to choose more central ceiling positions, setting the sensor suspended (not adjacent to the ceiling).

The spatial analysis by means of acoustic simulation of parameters C_{50} and D_{50} (Fig.s 8 and 9) also shows that the optimal spots for the acoustic recognition devices are a location as near as possible to the subject emitting the keywords. This means that two possibilities are imaginable: (i) defining the doctor-patient position inside the room and then fixing the recognition device at a single point, or (ii) installing several devices spatially covering most movements of the doctor and patient during therapy.



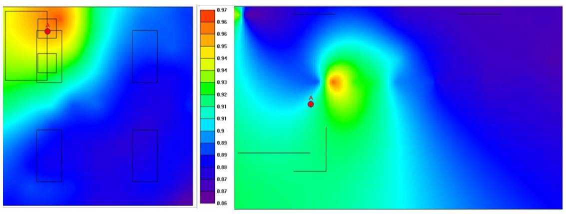
a) Top view b) Front view

Fig. 7 – Simulated RT for therapy room n. 2: spatial distribution at 1000 Hz (graduation range from dark blue = 0.8 s to red = 0.9 s)



a) Top view b) Front view

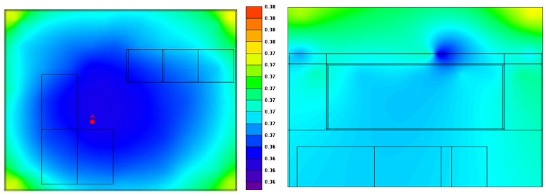
Fig. 8 – Simulated C_{50} for therapy room n. 2: spatial distribution at 1000 Hz (graduation range from dark blue = 7.5 to red = 14)



a) Top view b) Front view

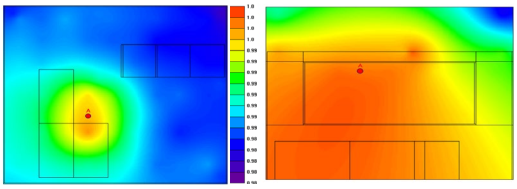
Fig. 9 – Simulated D_{50} for therapy room n. 2: spatial distribution at 1000 Hz (graduation range from dark blue = 0.85 to red = 1)

The same results obtained for therapy room n. 2 and 4 can be applied to the atelier: the larger dimensions of the room certainly require the application of several devices in order to reach and cover all the indoor space.
The soft room has low values of reverberation time (Fig. 10) because of the "soft" furniture intrinsically composed by sound-absorbing materials like polyurethane foams. Spatial distribution of Clarity and Definition (Fig. 11) are quiet good and it can be seen that a single sound suspended recognition device placed centrally on the ceiling can be sufficient for this type of room, and with the presence of one single user inside at a time.



a) Top view b) Front view

Fig. 10 – Simulated RT for soft room: spatial distribution at 1000 Hz (graduation range from dark blue = 0.3 s to red = 0.4 s)



a) Top view b) Front view

Fig. 11 – Simulated D_{50} for therapy room n. 2: spatial distribution at 1000 Hz (graduation range from dark blue = 0.9 to red = 1)

For the bedroom, the reverberation, clarity and definition are optimal for the application of sound recognition devices; at least one must be provided inside the bathroom and one should be provided inside the room. Positions of angle and proximity to the ceiling should be avoided. Obviously, however, the dimensions and the furnishings, optimized for the typical user, make the positioning of this device more "selective". For example, the presence of a wall fully equipped with windows, the location of the cabinets, the position of the bed and the presence of desks should be known in advance in order to choose the correct positioning of the devices.

4. Implementation

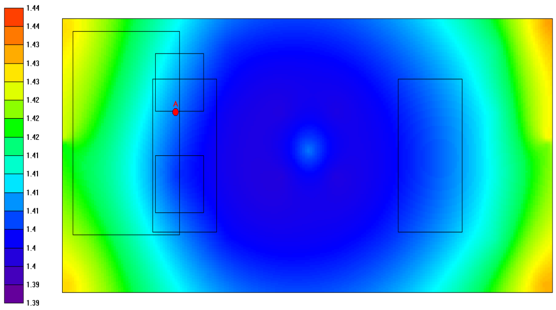
The procedural guidelines (Designing guidelines, 2008) for the design of spaces for people with cognitive disabilities such as autism, provide already optimal conditions for the introduction of assistive systems such as acoustic recognition devices.
More important is the need to have simple and relaxing spaces: linear and well-proportioned geometries are positively evaluated, which guarantee good sound reverberation and good acoustics, without sources of noise that may cause distraction or discomfort (it is better to choose radiant heating systems on the ceiling or floor and natural or forced ventilation systems with acoustic attenuators).

People with autism are extremely sensitive to stimuli and because of the difficulty of filtering foreground and background information are often able to perceive details that normal people may not notice. The fundamental concept of "simplicity" for living spaces leads to a reduction to a minimum of the furnishing accessories in the rooms (Balisha, 2017; De Giovanni, 2015; Schrank and Ekici, 2017) and, above all, to the avoidance of the insertion of suspended soundproofing elements (buffers, curtains, etc.).

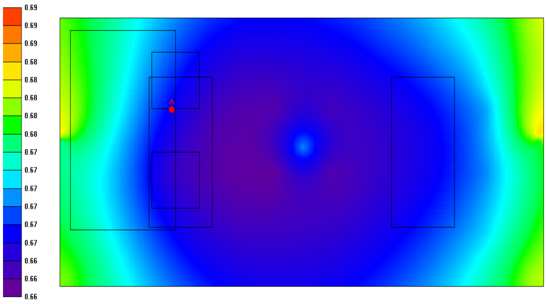
Rooms with long reverberation times, with acoustically highly reflective surfaces (i.e. those with large volumes and hard surfaces), are particularly unsuitable for many types of children's needs: some children with autism, for example, will find the room distressing. Children with hearing problems may also find the noise painful because it is amplified by their hearing aids.

All this means that the internal acoustic qualities of the room should be provided by the finishes of the walls, floors and ceilings. Moquette, for example, is soft and sound-absorbing, but difficult to clean. It is better to choose a linoleum flooring, perhaps choosing models with certified sound absorption coefficient values measured in accordance with the standard ISO 354. The walls can be covered with micro-perforated sound-absorbing panels, taking care not to alter the continuity of the wall itself. It is necessary to avoid creating elements that capture the eye and attention. Alternatively, furnishing accessories with sound absorbing elements on certain strategic points can be considered (the floor above the wardrobes or the lower part of the desk table, for example).

Fig. 12 shows some examples of the comparison results of the obtained acoustic simulations, starting from the studied rooms without furniture and sound absorbing elements, and then gradually implementing the necessary sound absorbing units in order to optimize the internal parameters of reverberation, definition and clarity. The procedure makes it possible to define in advance the installation positions of the necessary sound recognition devices, depending on the use that is made of the room and the operating needs required from time to time.



a) Without sound absorbing materials (graduation range: from dark purple 1.39 s – to red 1.44s)



b) With optimized sound absorbing materials

Fig. 12 – Comparison of simulated RT for therapy room n. 1: spatial distribution at 1000 Hz (graduation range: from dark purple 0.66 s – to red 0.69 s)

Specifically, for room therapy n. 1, the sound absorption measurements were performed using only sound absorbing panels on the ceiling in order not to alter the simple and clean look of the room and without the insertion of elements of possible distraction for the patient. Considering the possible positions in the space of the therapist and patient, the optimal positions for the sound recognition devices are along the longest sides of the room, at a height of about two meters from the ground.

For the bedroom, the sound absorption measurements were performed both through the insertion of sound absorbing panels on the ceiling/high part of the side walls and through the use of interior furnishings (Fig. 13). Here, considering the presence of large glass surfaces (for guests it is in fact better to choose as many sources of natural light as possible) and the need to insert furniture at full height (in order not to create discontinuity in the geometries, which can disturb the attention of the guest), it becomes advisable for the positioning of recognition devices to have a free wall on which to install one or two elements (depending on the size of the room) without obstacles that limit the scope of the operation (Fig. 14).

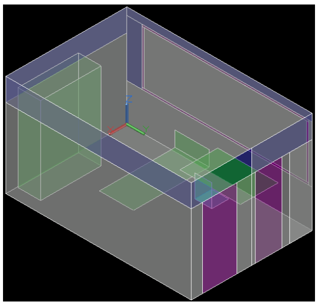


Fig. 13 – Example of optimized internal design for bedroom

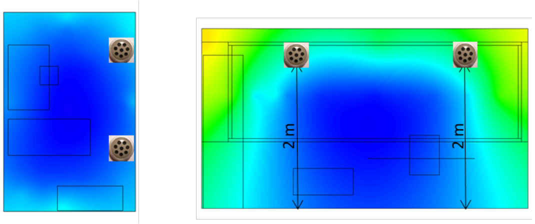


Fig. 14 – Example of optimized positions of recognition devices for bedroom (the room's internal colors represent the spatial distribution of the reverberation time parameter at 1000 Hz)

5. Conclusion

In this study, the results of the three-dimensional acoustic simulation of the spaces used by people with cognitive disabilities are presented, in order to understand the effective applicability of acoustic systems for remote voice monitoring and control.

It has been shown that interior design is very influential on acoustic fields and for this reason it may or may not allow the installation of AED and SSR technologies.

The inclusion of acoustic sensors, especially if combined with additional intelligent sensors monitoring and adjusting the environmental conditions (e.g. temperature, humidity, and acoustics) will allow a life as independent and autonomous as possible to people with cognitive disabilities, ensuring privacy and security. This technology will also make it possible to remotely deduce the state of people through centralized architectures by collecting data from a set of sensors deployed in their living environment.

Acknowledgement

The research was founded by University of Trieste in the framework of the ESF POR 2014-2020, Axis 3 of the Autonomous Region of Friuli Venezia Giulia, and ASTER system engineering smart technologies. The project is included in the Unit "HEaD Higher Education and Development" (CUP. FP1619892003). This research was also founded by EFRE 2014-2010 1095 E21@NOI CUP D56C18000180009.

Nomenclature

Symbols

AED	acoustic event detection
SSR	sound source recognition
TR	Reverberation time [s]
C_{50}	Clarity (dB)
D_{50}	Definition [%]

References

AA.VV. 2008. *Designing for disabled children and children with special educational needs Guidance for mainstream and special schools*, BUILDING BULLETIN 102, England.

Balisha, U. 2017. *Autism friendly design*. Didapress Dipartimento di Architettura, Florence University.

Biagetti, G., P. Crippa, L. Falaschetti, S. Orcioni, C. Turchetti. 2018. *Speaker Identification in Noisy Conditions Using Short Sequences of Speech Frames* Springer.

Chang, C., Y. Chang Y. 2013, “Application of Abnormal Sound Recognition System.” *Proceedings of 9th International Conference on Information, Communications & Signal Processing (ICICS).*

De Giovanni, G. 2015. *Architecture for authism. Esempi Di Architettura* Vol. 2, N. 2 ed. Edizioni Scientifiche.

Dufaux, A. 2001. *Detection and Recognition of Impulsive Sound Signals.* PhD thesis, Institute of Microtechnology University of Neuchatel Switzerland.

Farina, A., L. Tronchin. 2013. "3D Sound Characterisation in Theatres Employing Microphone Arrays." *Acta Acustica United with Acustica* 99(1): 118-125. doi:10.3813/AAA.918595

Fleury, A., N. Noury, M. Vacher, H. Glasson, J.-F. Serignat. 2008. "Sound and speech detection and classification in a Health smart home." *Proceedings of 30th Annual International Conference of the IEEE Engineering in Medicine and Biology Society.*

Foster, P., S. Sigtia, S. Krstulovic, J. Barker, M. D. Plumbley. 2015 "Chime home: a dataset for sound source recognition in a domestic environment." *Proceeding of IEEE Workshop on Applications of Signal Processing to Audio and Acoustics*, October 18-21, New Paltz, NY.

ISO. 2003. *ISO 354: Acoustics -- Measurement of sound absorption in a reverberation room.*

ISO. 2009. *ISO 3382-1 Acoustics — Measurement of room acoustic parameters — Part 1: Performance spaces.*

Janvier, M., X. Alameda-Pineda, L. Girin, R. Horaud. 2012 "Sound-Event Recognition with a Companion Humanoid." *Proceedings of 12th IEEE-RAS International Conference on Humanoid Robots.*

Kinoshita, K., M. Delcroix, T. Yoshioka, T. Nakatani, et al. 2013. "The reverb challenge: a common evaluation framework for dereverberation and recognition of reverberant speech." *Proceedings of IEEE Workshop on Applications of Signal Processing to Audio and Acoustics*, New Paltz, NY, October 20-23.

Lecouteux, B., M. Vacher, F. Portet. 2011. *Distant Speech Recognition for Home Automation: Preliminary Experimental Results in a Smart Home.* Laboratoire d'Informatique de Grenoble, GETALP Team.

Lecouteux, B., M. Vacher, F. Portet. 2018. "Distant speech processing for smart home: comparison of ASR approaches in scattered microphone network for voice command." *International Journal of Speech Technology.* https://doi.org/10.1007/s10772-018-9520-y

Pagin, D. 2016. *Localization and recognition of sound events by means of microphone arrays.* Master's thesis in Computer Engineering, University of Padua Department of Information Engineering.

Piana, E., P. Milani, N. Granzotto. 2014. "Simple method to determine the transmission loss of gypsum panels." *Proceedings of 21st International Congress on Sound and Vibration 2014, ICSV 2014, vol. 5: 3700-3706*; Beijing; China; July 13-17 2014.

Portet, F., M. Vacher, C. Golanski, C. Roux, B. Meillon. 2011. *Design and evaluation of a smart home voice interface for the elderly: acceptability and objection aspects.* Springer, London.

Robles, R. J., T. Ki. 2010. "Applications, Systems and Methods in Smart Home Technology: A Review." *International Journal of Advanced Science and Technology* 15.

Schrank, S., D. Ekici. 2017. *Healing Spaces, Modern Architecture, and the Body*. Routledge.

Tronchin, L. 2013. "On the Acoustic Efficiency of Road Barriers: The Reflection Index." *International Journal of Mechanics 7*(3): 318-326.

Tronchin, L., V. L. Coli. 2015. "Further Investigations in the Emulation of Nonlinear Systems with Volterra Series." *AES: Journal of the Audio Engineering Society* 63(9): 671-683. doi:10.17743/jaes.2015.0065

Yong, L., C. Kean. 2014 "Automatic classification of impact sounds based on time-frequency representation." *Proceedings of The 21st International Congress on Sound and Vibration*.

Yoshioka, T., A. Sehr, M. Delcroix, K. Kinoshita, R. Maas, T. Nakatani, W. Kellermann. 2012. "Making Machines Understand Us in Reverberant Rooms [Robustness against reverberation for automatic speech recognition]." *Signal Processing Magazine*, IEEE: 114-126.

Zhang, H., I. McLoughlin, Y. Song. 2015. "Robust sound event recognition using convolutional neural networks." *Proceedings of ICASSP 2015 - 2015 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP).*