

ROUND AND ROUND WE GO: SPEECH INTELLIGIBILITY AND LOCALIZATION FOR IN-THE-ROUND SPACES

E Magloire
B Harrison
L Dellatorre
B Cardenas

Charcoalblue, New York, USA
Charcoalblue, London, UK
Charcoalblue, London, UK
Charcoalblue, New York, USA

1 INTRODUCTION

The directional characteristics of the human voice are a key theatrical effect in spoken word drama. Maintaining good speech intelligibility and believable localization as performers move and reorient on the stage is a formidable design challenge for in-the-round spaces. This paper provides an overview of the design strategies for this theatrical form, using some notable precedents. Referencing Chicago's newest in-the-round theatre, named the Ensemble Theater, at Steppenwolf Theatre, we present the acoustic goals and the architecturally integrated design for speech acoustics. It also includes the key design parameters, the traditional and non-traditional intelligibility metrics used to evaluate the design options and the completed construction, along with subjective impressions of artists and audiences related to speech intelligibility and localization.

2 HISTORY OF IN-THE-ROUND THEATRES

Theatre in-the-round is a unique form of theatrical experience, where the stage is at the centre with seated audience around it. This configuration allows performers to be viewed from all sides, creating an intimate, immersive and interactive theatrical experience.

The history of in-the-round storytelling is as old as storytelling itself. The architectural origins are ascribed to the Greeks who extended the stage area into the centre of the plan with concentric seating around. Roman amphitheatres utilised similar geometrical principles.

From the Medieval area, there are records of circular moats, embankments and scaffolds used for performance with the most notable example from Cornwall, England. (Southern) Pageant wagons and similar travelling troops continued the practice of gathering around a stage, but none of these have left architectural remnants.

Elizabethan Inn-yard theatres evolved into the “wooden O” theatres which were circular (or polygonal) in plan. These theatre forms are generally interpreted as having the stage at one side and the audience not fully in-the-round. There remains, however, some speculation over the evolution of the Elizabethan raised stage. It is suggested that the placement of audience and actors may have been more fluid than we often assume, with actors playing in the centre of the room and audience sat on stage.¹

It's not an exaggeration to say that in the next few centuries, there are few if any examples of in-the-round vantage points influencing theatre architecture. However, the practice of audience sitting on stage continued.² It was David Garrick who reformed the operation of Drury Lane, in the middle of the 18th Century, and ceased the practice of selling on-stage tickets.¹

Walter Gropius offers the next, yet unrealised, architectural example of the in-the-round stage in his “Total Theatre” concept from 1926.³

While fixed-format in-the-round theatres, as we recognise them today, didn't emerge until the middle of the 19th Century, it is not suggested that in-the-round performance didn't occur until then. At modest scale, a raised stage and loose chairs could be a perfectly adequate in-the-round theatre.

Theatre educator and director Glenn Hughes designed the earliest example of the in-the-round theatre architecture for the University of Washington in 1940 (Seattle, Washington, USA). The Penthouse Theatre, which now bears his name, seats 160.⁴

Margo Jones, a major proponent of the in-the-round format, founded Theater '47 in Dallas, Texas, as the first modern non-profit professional resident theatre in America.

Shortly after, Arena Stage in Washington DC was founded in 1950 and operated out of a re-purposed venue until its fixed-form in-the-round theatre was completed in 1960. Designed with 800 seats, the theatre reduced to 680 seats during a substantial refurbishment in 2010.

In the UK, Stephen Joseph founded a theatre in-the-round in Scarborough, North Yorkshire in 1955. The company that now bears his name has proceeded to operate in-the-round, now in a 404-seat venue named The Round.

The predecessor company to Broadway's Circle in the Square Theatre was founded in 1950 and operated an in-the-round theatre at 5 Sheridan Square in New York. The company's subsequent flexible in-the-round venue (often operating in thrust format) was constructed in 1972 at Paramount Plaza at Broadway and 50th Street in Manhattan. It has been a substantial influence, offering non-proscenium staging on US theatre's biggest platform.⁵

Between 1971 and 1976 the disused cotton trading hall at the Manchester Royal Exchange was converted into a 7-sided theatre in-the-round for the company that would eventually take the name of the building. The theatre is a notable, futurist, free-standing conception which stands within the 1874 hall (which was substantially re-built following war damage).⁶

Venues such as the 600-seat @sohoplacetheatre in London and the Ensemble Theatre at Steppenwolf Theatre Company in Chicago (the focus of this paper) both opened in 2022 and are a testament to the enduring relevance of the in-the-round format to global theatre.

3 ACOUSTIC AND DRAMATIC INTERSECTIONS

The acoustic challenge of in-the-round theatres is, very plainly, that the human voice is directional and in the in-the-round format actors face away from a portion of the audience at all times. The acoustic condition is complicated by our reliance upon visual cues which we use for speech intelligibility. When the body blocks the action, context clues are removed. When faces cannot be seen, facial expressions and mouth movements cannot be used to fill in the gaps in comprehension.

3.1 Scale

The dramatic presentation of in-the-round work was initially mounted at studio-theatre scale. The close proximity of actors and audience (irrespective of orientation) allows good speech intelligibility. Additionally, the relatively small dimensions and room volume (and therefore favourable loudness and limited reverberation) are also favourable to the acoustic conditions.

As in-the-round theatres scaled up to larger, commercially-viable capacities and incorporated production values consistent with "proper" proscenium theatres, rooms got bigger to accommodate larger stages, more audience, and more technical equipment. The acoustic conditions of these rooms are more challenging at this scale.

3.2 Scenic design

The dramatic presentation of theatre work in-the-round has very specific scenic conditions. Sets must be minimal to avoid blocking the action from certain vantage points. Whereas in proscenium theatres (and especially opera) the floor can be a reliable source of early reflections, often the clutter of small-scale scenic elements and furniture on in-the-round stages can obscure useful reflections from the

floor. Often the floor is used itself as a scenic element and the material used (grass, sand, carpet, etc) can limit the acoustic utility of the floor.

The other scenic potential is the “ceiling” or suspensions above the stage. Flying of small objects is often necessary, and access to all parts of the over-stage areas required. The only overhead obstructions are generally the bridges or catwalks which are kept to minimum width. The height required to fly objects out of view and to provide safe access for technicians demands a high-ceiling height---and therefore architectural surfaces which are far away from actors’ voices and listeners’ ears.

3.3 Audience distribution

The configuration of the audience over multiple levels with shallow overhangs can allow larger capacity spaces to exist in a relatively smaller footprint. For example, the Manchester Royal Exchange in a 20m x 20m footprint has roughly the same seating capacity as the Fichandler Stage at Arena Stage in a 27m x 29m footprint. Manchester Royal Exchange has two balconies; the Fichandler is all on one level.

4 ACOUSTIC CONDITIONS AND OBJECTIVES

The following acoustic conditions and objectives are used to inform design decisions for theatres—and especially for the in-the-round format.

4.1 Loudness and Intelligibility

The proximity from actor to audience (source to receiver) is a key factor to achieving adequate loudness. Architectural reflections which contribute to the impression of the direct sound are critical to achieving intelligibility.

4.2 Realistic localization

With the audience fully surrounding the stage, it can be difficult for audience members, who rely on visual cues, to determine which actors are speaking, potentially affecting their engagement, and understanding of the performance. The strategies used to support the direct sound through architectural reflections have the potential to affect the perception of localization. Ceiling reflections may be especially useful in aligning the reflected energy with the direction of the sound source.

4.3 Preserving naturalness

The unamplified aesthetic is an important artistic component of realism in drama. The position of an actor needs to be able to be tracked aurally. An actor who turns away from the audience should sound different. The essence of the spatial sound is a continuing study, and especially what the boundary is between what sounds natural and what sounds artificial.

4.4 Actor projection

With audience members surrounding the stage, actors should intend to project their voices effectively to reach all corners of the theatre without straining or compromising vocal quality. While this similarly applies to end-on (or proscenium) theatres, in in-the-round theatre actors may instinctively adjust their voices to serve only those in their field of vision (and not those behind).

4.5 Sound amplification

The use of sound amplification for musicals, for sound effects, vocal effect, and underscoring also has an impact on the unamplified design. The loudness of amplified sound from overhead sources very near to the listener may make it difficult for the listener to quickly adjust to the loudness and

spatial qualities of the amplified voice (e.g., it may be difficult to understand the unamplified “book” in a musical after an amplified song).

4.6 Directivity of the human voice

The directivity of the human voice refers to the way in which sound energy is distributed when someone speaks. In general, the human voice exhibits a combination of both omnidirectional and directional characteristics, which vary with the frequency of sound.

At lower frequencies (below approximately 1kHz), the human voice tends to be omnidirectional, that is to say that voice is emitted relatively evenly in all directions. Therefore, there is only a small difference between the speech level in front and at the back of the head. At higher frequencies (above approximately 1 kHz), the directivity of the human voice becomes more distinct. This is also the range of frequencies (1kHz to 4kHz) where there is most sensitivity in human hearing.

These higher frequencies also carry a substantial amount of information, including the consonant, plosive, and liquid sounds that allow us to discern between words that sound similar. Such directionality creates a significant difference in speech levels between the front and back of a speaker. At 2 and 4kHz, levels behind a speaker are substantially reduced compared to 500Hz and 1kHz⁷.

As a reference, a difference of 10dB is subjectively perceived as half the loudness.

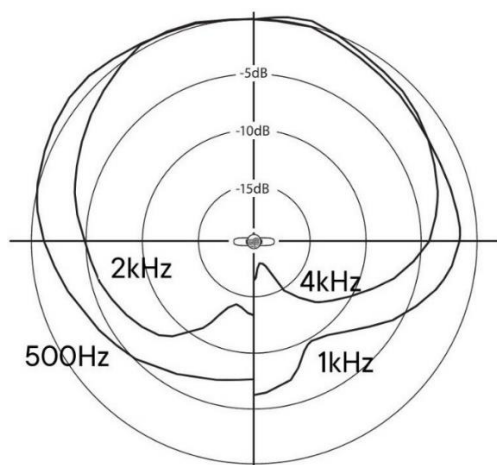


Figure 2 Polar diagram for a speaker⁷

This demonstrates the effect of the orientation of the head with regards to speech direction and loudness. The orientation of an individual's head obviously impacts the level of speech as it reduces when a speaker faces away from you.

It is important to note that the directivity of the human voice can vary among individuals due to anatomical differences, vocal projection technique, and personal habits (as well as stage direction).

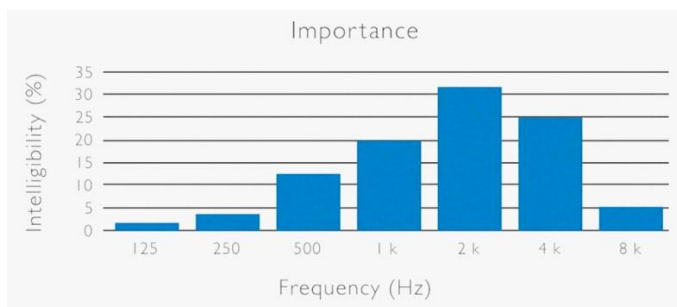


Figure 2 French and Steinberg⁸ – relative importance of difference frequencies to achieve good speech intelligibility.

French and Steinberg⁸ noted that not all the frequencies contribute in the same way in achieving good speech intelligibility. The next table shows that 2kHz might be one of the most important frequencies regarding perceived speech intelligibility with the sound of most consonants being between 2 and 4 kHz.

6 ENSEMBLE THEATRE, STEPPENWOLF THEATRE COMPANY, CHICAGO

Under the artistic leadership of Anna Shapiro, the Steppenwolf Theatre Company developed a new building to replace its previous rehearsal room-cum-theatre known as the Upstairs Theatre with a new 408-seat in-the-round theatre. The concept for the theatre was established in mid-2016 with its first show premiering just under six years later. The architect is Adrian Smith + Gordon Gill Architecture with Charcoalblue as theatre consultant and acoustician.

The design includes a large trapped stage of 7m x 10.7m and six rows of concentric seating and a perimeter circulation. Four vomitory entrances lead to stage.

An overhead network of catwalks is provided with a technical grid over an area just larger than the dimension of the stage. A control room is provided at catwalk level.

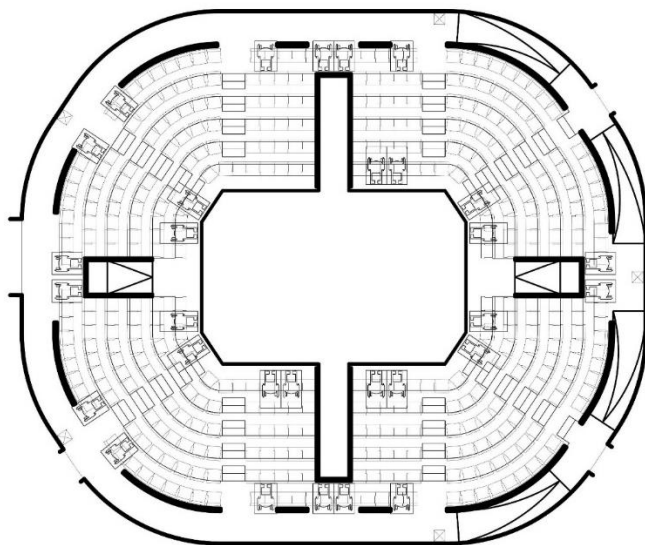


Figure 3 Plan of the Ensemble Theatre, Steppenwolf Theatre Company, Chicago (drawing by Charcoalblue)

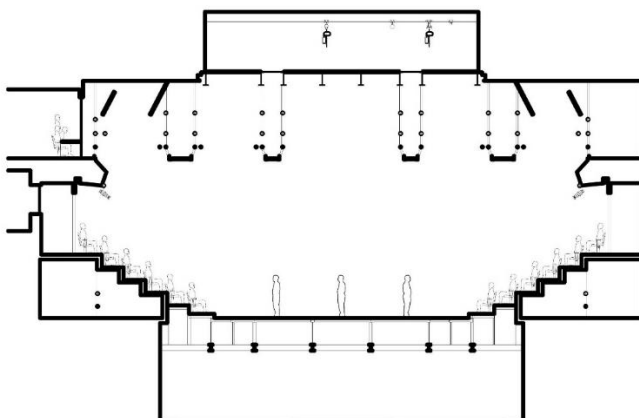


Figure 4 Section of the Ensemble Theatre, Steppenwolf Theatre Company, Chicago (drawing by Charcoalblue)

8 ROOM ACOUSTIC DESIGN FOR THE ENSEMBLE THEATER

8.1 Size and acoustic volume

Typical benchmarks for acoustic volumes of spoken word theatre spaces tend to be less than 7m³ per person while excluding some the stage house volume (perhaps both overhead and in the wings). The calculated acoustic volume for the Ensemble Theatre is 10m³ per person, including all overstage volume (but excluding the trapped space below stage).

The design of the Ensemble Theatre has intentionally excluded the overhead volume next to the flying space and above the seats in order to avoid having excess and detrimentally acoustic volume where not needed for technical requirements. This volume is largely empty, containing only structure and ductwork. The perimeter soffit created is approximately 10m above the stage to accommodate the flyloft, with 6.6m clear below the catwalks.

8.2 Reductions to “effective volume”

The perimeter circulation at the back of the seating rake contributes about 7.5% of the total volume to the room. A free-standing inner wall was integrated into the back of the seating to provide a closer sound reflective surface to listeners and to exclude that volume from the room as much as possible. The audience-facing side of this wall is provided with an irregular sound-scattering pattern providing diffusion in the horizontal plane.

8.3 Sound absorption strategy

With the exception of the flyloft, the only fixed sound absorbing architectural finish in the auditorium is on the side of the free-standing wall (described above) that faces away from the stage (circulation side), preserving a sound reflective surface facing stage. All other surfaces, except the seats themselves, are minimally sound absorptive in the mid-frequency range. Multi-layer gypsum board construction is used throughout the interior; the low-frequency sound absorption it provides was generally welcomed.

The walls and ceiling of the flying space are covered entirely with sound-absorbing thick sound-absorbing finishes. (The high ceiling reflection would have been as late as 60ms after the direct sound; it was determined that any potential up-down reflections between stage and a sound-reflective ceiling would be better off being controlled.)

8.4 Early sound reflections

The acoustic design includes sound reflective surfaces that produce early arriving sound reflections (up to 85ms), which are necessary when actors face away from the audience. These surfaces are in areas coincident with the strongest directivity of the human voices.

The surfaces, providing cross room reflection paths, include:

1. Two perimeter sound reflectors at the lighting bridge level: the outer surface is optimised for a centre stage source; the inner surface is optimised for stage-edge sources on the opposite side of the room from the sound-reflective surface serving the audience rake near to the actor (but who is facing away from them).
2. Front of the technical level: The vertical face of the soffit is optimised for actor positions mid-stage and nearer to the sound-reflective surface serving audience seated behind the actor on the far side of the room.
3. Corner at underside of the technical level (and top of perimeter wall at back of seating rake): These surfaces are optimised for stage-edge sources near to the sound-reflective surface serving the stage area behind the actor and the first few rows of seating on the opposite side of the room.

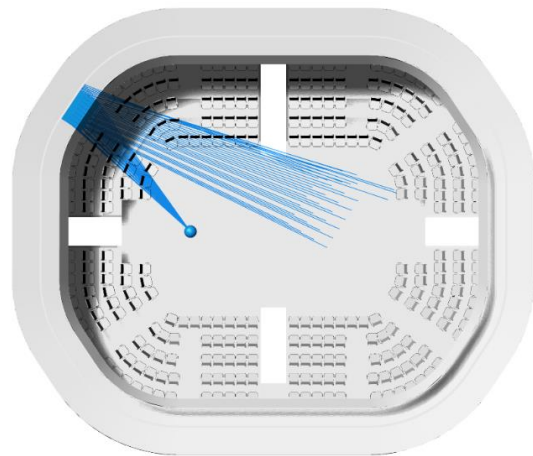
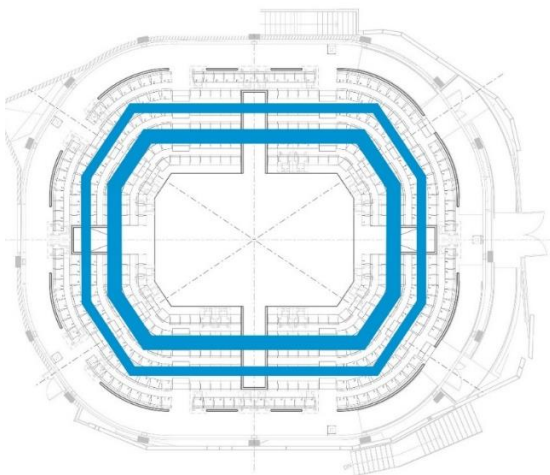
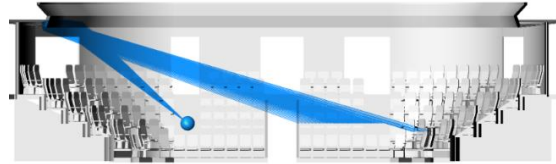
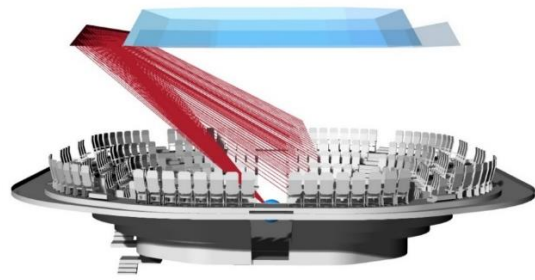


Figure 1 Perimeter sound reflectors (in section and in plan). (Charcoalblue acoustics ray tracing algorithm in Rhino3D-Grashopper)

Figure 2 Soffit optimised for mid stage actor (in section and in plan). (Charcoalblue acoustics ray tracing algorithm in Rhino3D-Grashopper)

Ray tracing analysis of first-order sound reflections from these surfaces indicates arrival times between 30ms and 70ms after the direct sound (with greater time delays for higher order reflections). Given the size of the room and distances to sound reflective surfaces, many of these reflections are after the conventional time windows for sound reflections which are assumed to contribute to speech intelligibility.

Experience in other rooms of similar scale suggests that arrival times in this range, or even later, may contribute to the clarity and intelligibility of the spoken word.

9 ROOM ACOUSTIC MEASUREMENTS

Charcoalblue performed room acoustics measurements to evaluate and analyse the acoustic impulse response (IR) of the Ensemble Theater in accordance with BS EN ISO 3382¹⁰. An acoustic impulse response is a measurement that is key to architectural acoustics, because it contains information on how sound, produced at a certain location stage, arrives at the audience. The information contained in this measurement includes strength and temporal information of the direct and reflected sound. The analysis of the IR allows to determine the acoustics characteristics of a space and helps identify potential acoustic issues.

As reference, the measured unoccupied reverberation time (using the omnidirectional source, per BS EN ISO 3382), are below.

Average Unoccupied Reverberation Time, T30, seconds (avg. of 25 positions)							
63Hz	125Hz	250Hz	500Hz	1kHz	2kHz	4kHz	8kHz
1.1	1.0	1.0	1.0	1.0	1.0	0.9	0.8

The main objective of the measurements was to evaluate the performance of the room for its specific theatrical use where the voice is not amplified. In order to do this, measurements were taken not only with the use of an omnidirectional source, as required by the ISO 3382, but also with a small directional loudspeaker, NTI Audio TalkBox.

With a sound directivity comparable to that of the human voice, the loudspeaker has a 4-5" cone and is designed for speech intelligibility / STIPA measurements.

Using the directional loudspeaker, a sine sweep was played through the directional loudspeaker, located in multiple positions on stage and with different orientations / head rotations. Recordings were conducted, using an Earthworks M30 microphone at various seating positions around the theatre.

All of our measurements of the Ensemble Theater were in an unoccupied auditorium.

From the IR, we estimated objective acoustics parameters, which through research, have been related to various aspects of subjective acoustic impressions¹¹. For the purpose of speech intelligibility, we have focused on the acoustics parameter D50, which relates to the ratio of early sound energy, within the 0-50ms time window, to the total sound energy.

$$D_{50} = \frac{\int_0^{0,050} p^2(t) dt}{\int_0^{\infty} p^2(t) dt}$$

This ratio plays an important role in establishing a room's character with respect to clarity of speech, reverberance, and spaciousness.

Definition, denoted D50: Definition parameter measures the early-to-total sound energy ratio and is related to speech intelligibility. Higher values indicate better clarity and speech intelligibility. In other words, a higher value signifies that a greater amount of sound energy arriving very soon after the direct sound than at later times. This early energy (within 50ms from direct sound arrival time) helps build clarity and strength of sound.

The Table¹² below correlates speech intelligibility with D50 values.

Speech intelligibility	D50, %
Excellent	86 to 100
Good	67 to 86
Fair	39 to 67
Poor	17 to 39
Bad	6 to 17

Another aspect from which the presentation of results differs from the requirements of BS EN ISO 3382 is that for this particular exercise the interest was to evaluate D50 at different speech frequencies from 125 to 4000Hz instead of presenting data as a 500-1000Hz average.

Since the natural voice is increasingly more directional with higher frequency the D50 associated with different head directions was expected to be more similar at lower frequencies rather than higher frequencies.

The following graph shows the measured D50 (Definition) in the Ensemble Theatre.

1. The front (0 degree): the loudspeaker faces the seat where the microphone, picking up the signal played by the loudspeaker, is located.
2. The back (180 degrees): the loudspeaker faces away from the seat where the measurement is conducted. This was to reproduce the effects of actors facing away from seat (or showing their back to seated audience).
3. Sides (90 and 270 degrees): face left and right with respect to the listener. (At high frequencies the D50 value related to these two directions are not the same due to the asymmetry of the room referenced to the loudspeaker angle.)

The graph shows “excellent” values of D50 in the “facing” condition and “fair to good” values in other source orientations. These results correlate with high level speech of intelligibility.

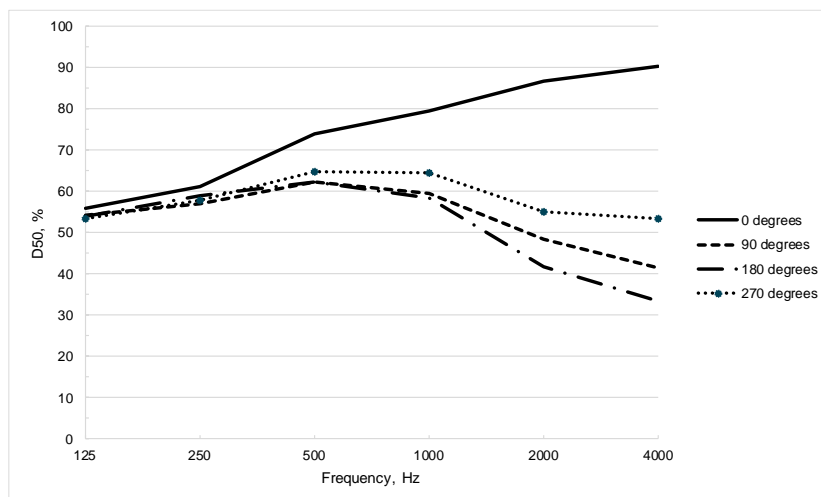


Figure 3 Definition D50 measured in the theatre, average of different head directions.

Given the relevant arrival times of the reflections from the sound-reflective surfaces, as designed, the relevant values of Clarity indices, C50 and C80, were also explored. The Clarity parameters measure the ratio of early-to-late sound energy given the noted threshold in millisecond. The value is reported in decibels. Higher values indicate better clarity and speech intelligibility. While Clarity metrics are generally used in the context of music spaces, the extended time range is the standard metric is especially relevant.

$$C_{t_e} = 10 \lg \frac{\int_0^{t_e} p^2(t) dt}{\int_{t_e}^{\infty} p^2(t) dt} \text{ dB}$$

The following graphs show the measured C50 and C80 (Clarity) in the Ensemble Theatre for “facing” and “facing away” conditions, using the simulated speech source.

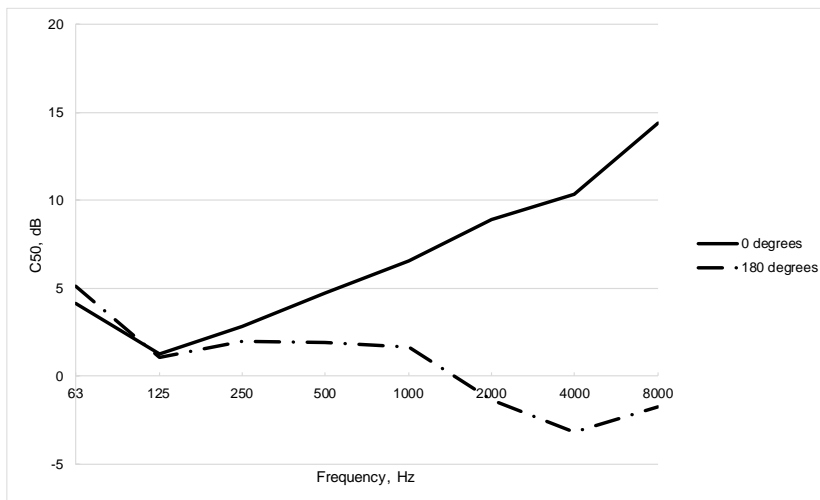


Figure 8 Clarity, C50, measured in the theatre, average of different source head directions.

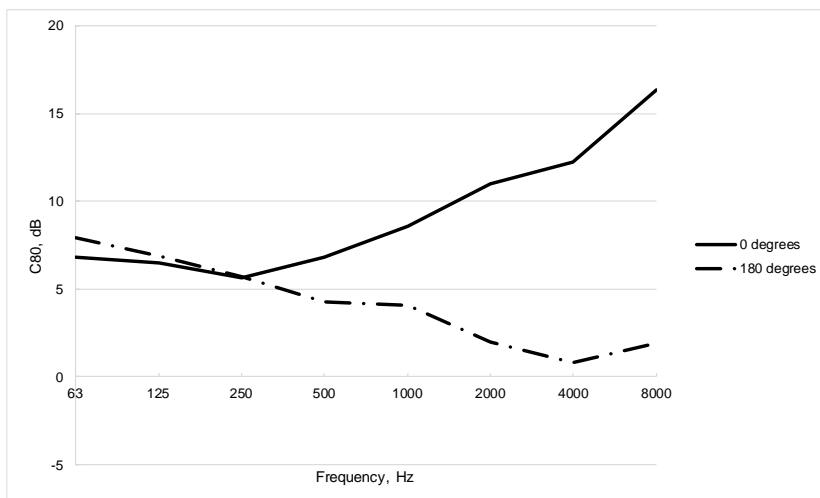


Figure 9 Clarity, C80, measured in the theatre, average of different source head directions.

By definition, the C80 values will be larger than the C50 as a larger portion of the impulse response is being integrated (with a smaller denominator, unlike Definition which references the whole impulse response). The overall trend of the C50 and C80 values are generally the same, with an approximately 4dB shift at mid- and high-frequencies.

10 SUBJECTIVE PERCEPTION

The reception of the Ensemble Theatre has been excellent and especially from the acting company. The coordination of acoustic, technical, and architectural design has been proven to allow flexible technical use of the overhead scenic zone without substantial impact from the sound-reflecting surfaces.

The audience of Steppenwolf are being offered in-the-round staging and direction on a regular basis for the first time. Actors, directors, and audiences are all adjusting to new expectations for the new format. Some audience comments have been received that some dialogue can be difficult to hear when actors are facing away. The predominance of comments have been received from older listeners. While no correlation has been possible, it is suggested that those commenting are more likely to be affected by age-related hearing loss. The frequencies of age-related hearing loss have substantial overlap with the frequencies identified in our study where Definition and Clarity degrade as actors turn from a “facing” orientation to 90-degrees of axis to “facing away.” The alignment of directivity and hearing loss issues underscores the need to enhance our understanding of how architectural acoustics can support both situations.

11 CONCLUSIONS

The Ensemble Theater for the Steppenwolf Theatre Company has been greatly influenced by precedent venues, their key architectural features, and listening evaluations. The key design interventions include an acoustic volume that is sufficiently small and adequate to support the loudness of the voice, and also include sound reflecting surfaces that are architecturally integrated and yet provide early arriving sound reflections necessary to help achieve speech intelligibility.

As standard measurements with the omni-directional sources have not been useful in evaluating the directional effects of speech in-the-round rooms, this study has attempted to explore the use of standard parameters using directional sources. However, the perceived importance of the “early, but later” energy in large in-the-round rooms is not able to be well analysed with typical time windows. Further work on energy-ratios and time windows is suggested to allow deeper evaluation of rooms of this scale and function.

12 REFERENCES

1. Joseph, S. (1963) *The Story of the Playhouse in England*. London: Barrie and Rockliff.
2. Rosenfeld, S. (1939) *Strolling Players and Drama in the Provinces 1660-1765*. London: Cambridge University Press.
3. Gropius, W. and A.S.Wensinger, eds. (1961) *The Theater of the Bauhaus*. Baltimore and London: The John Hopkins University Press.
4. <https://www.historylink.org/File/3710>
5. <https://archives.nypl.org/the/21843>
6. Fraser, D. (1988) *The Royal Exchange Theatre Company. An Illustrated Record*. Manchester: Royal Exchange Theatre Company Ltd
7. Moreno, A. and J. Pfretzschner (1978) 'Human Head Directivity in Speech Emission: A New Approach', *Acoustics Letters*, 1:78-84.
8. N.R. French & J.C. Steinberg (1947) 'Factors Governing the Intelligibility of Speech Sounds', *JASA* vol. 19, No 1.
9. W.T. Chu, A.A.C. Warnock, Detailed Directivity of Sound Fields Around Human Talkers, Research Report (National Research Council of Canada. Institute for Research in Construction); no. RR-104, 2002-12-01.
10. BS EN ISO 3382-1:2009 Acoustics: Measurements of room acoustic parameters – Performance spaces.
11. Kuttruff, H. (1991) *Room Acoustics*, 3rd Edition, Elsevier Applied Science.
12. Marshall, L.G. (1994) 'An Acoustic Measurement Program for Evaluating Auditorium Based on the Early/Late Sound Energy Ratio', *Journal of the Acoustical Society of America*, 96: 2251-2261.