SDA CONSULTING | WHITE PAPER

Hearing Justice: What the Science of Attention Tells Us About Courtroom Sound

An SDA Consulting White Paper on courtroom sound system design, prepared for a judicial readership.

Abstract

Most courtrooms reinforce speech through distributed ceiling loudspeakers, a design that delivers adequate loudness but discards the spatial and visual cues the auditory system relies upon to separate one voice from many. Research on selective auditory attention, commonly termed the "cocktail party effect," indicates that this omission increases listening effort and degrades comprehension and memory, particularly among listeners with hearing loss. This article reviews that research and proposes an alternative design built on audio-visual anchoring: loudspeakers co-located with the bench, with the display used for remote participants, and with time-aligned desktop extension speakers at each bench location and at the attorney's tables. The Jury and Gallery are supplemented by time-aligned overhead speakers. The design is intended to serve the hybrid courtroom, in which in-person proceedings in the well coexist with remote arraignments and remote expert testimony.

Introduction

The integrity of adjudication presupposes that participants hear what is said. Jurors must comprehend and retain testimony; witnesses must understand questions; defendants must be able to follow proceedings against them. Courtroom acoustics are seldom examined in these terms, yet the perceptual conditions of the room bear directly on each. The growth of hybrid proceedings, in which remote participants appear by videoconference, adds a further demand: the voice of a person on a screen must be as intelligible, and as clearly attributable, as the voice of a person in the room.

The Science of Selective Listening

Cherry (1953) first described the listener's capacity to follow one talker among several. Broadbent (1958) proposed an early filter that excludes unattended input, while Treisman (1964) demonstrated that unattended speech is attenuated rather than eliminated. Bregman (1990) situated these findings within a broader framework of auditory scene analysis, in which the auditory system groups acoustic components into discrete sources using cues such as pitch, onset timing, and spatial location.

Subsequent work has emphasized the role of spatial separation. Speech is markedly more intelligible when target and competing sounds originate from different locations (Bronkhorst, 2000), and attention appears to operate on perceptual objects whose formation depends on such cues (Shinn-Cunningham, 2008). Neurophysiological recordings confirm that cortical activity selectively tracks the attended talker (Mesgarani & Chang, 2012).

Two further phenomena are central to the design proposed here. The first is the precedence effect: when the same sound arrives from several directions within a few milliseconds, listeners localize it to the direction of the first arrival (Wallach, Newman & Rosenzweig, 1949; Litovsky et al., 1999). The

second is the ventriloquist effect: when auditory and visual cues conflict modestly, perceived sound location is drawn toward the visible source (Alais & Burr, 2004). Together they imply that a voice can be made to appear to come from its talker, provided the earliest-arriving sound originates at or near that talker.

Listening Effort and Its Consequences

Intelligibility is not binary. Under degraded conditions listeners may report every word correctly while expending substantially greater cognitive resources to do so. The Framework for Understanding Effortful Listening (Pichora-Fuller et al., 2016) describes how such effort draws on limited capacity otherwise available for comprehension and memory. For jurors charged with retaining and evaluating extended testimony, elevated listening effort represents a plausible, if seldom acknowledged, threat to deliberative quality.

These costs fall unevenly. Age-related and noise-induced hearing loss impair spatial separation abilities before they impair hearing in quiet, so affected jurors may function adequately in conversation yet struggle in a reverberant courtroom. Courts are also subject to accessibility obligations, including the provision of assistive listening systems under the Americans with Disabilities Act.

Limitations of Ceiling-Based Reinforcement

The conventional courtroom system mixes all microphones and distributes the result through an array of ceiling loudspeakers. This approach achieves uniform loudness but carries three perceptual liabilities.

First, every voice appears to come from above, and from the same place. The spatial cue that would distinguish judge from witness from counsel is eliminated, so listeners must separate talkers on voice quality alone. ***Second***, the auditory location contradicts the visual location of the talker, a mismatch that the ventriloquist effect can only partially reconcile. ***Third***, ceiling loudspeakers excite the reverberant field of the room, lowering the ratio of direct to reflected sound and blurring the temporal cues on which intelligibility depends.

Hybrid proceedings magnify these problems. A remote expert's voice, reproduced overhead, is detached from the face on the display, and the same ceiling loudspeakers that serve the room can return sound to the room microphones, complicating echo cancellation for the remote party.

An Alternative Design: Audio-Visual Anchoring

The proposed design places reinforcement where the talkers are, so that each voice arrives first from its visible source.

Bench anchor. A free-standing, floor-mounted column loudspeaker placed beside the bench, rather than built into the casework, reproduces the judge's voice and serves as the primary acoustic reference for the room. Because it is co-located with the judge and the witness the voice is localized correctly and the judge's authority is reinforced perceptually as well as visually.

Display anchor. A loudspeaker mounted adjacent to the large display used for videoconferencing reproduces remote participants, including defendants appearing for remote arraignment and remote expert witnesses. The voice is heard to come from the face on screen, preserving the audio-visual

correspondence that supports both intelligibility and the jury's assessment of demeanor. Audio and video should be kept synchronized within accepted broadcast tolerances, since lip-sync error itself increases listening effort.

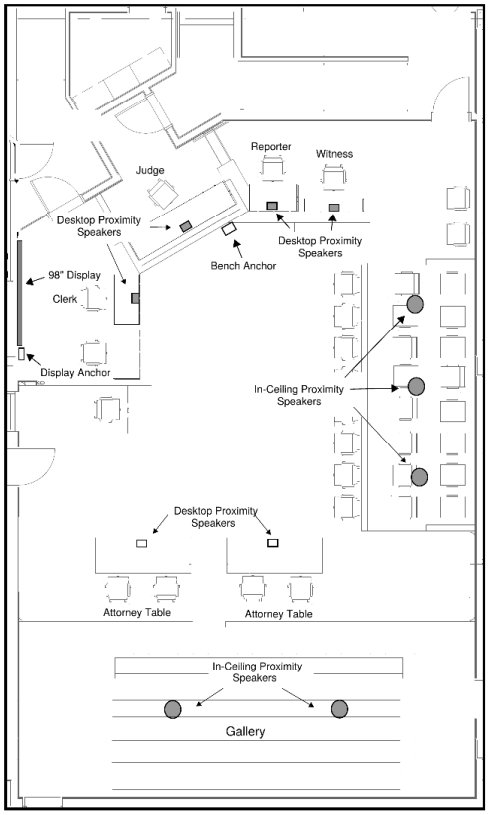


Figure 1. Plan view of the proposed courtroom sound system, showing the bench anchor, the display anchor adjacent to the videoconference screen, desktop proximity loudspeakers at the judge, clerk, court reporter, witness, and counsel positions, and time-aligned overhead reinforcement above the jury box and gallery.

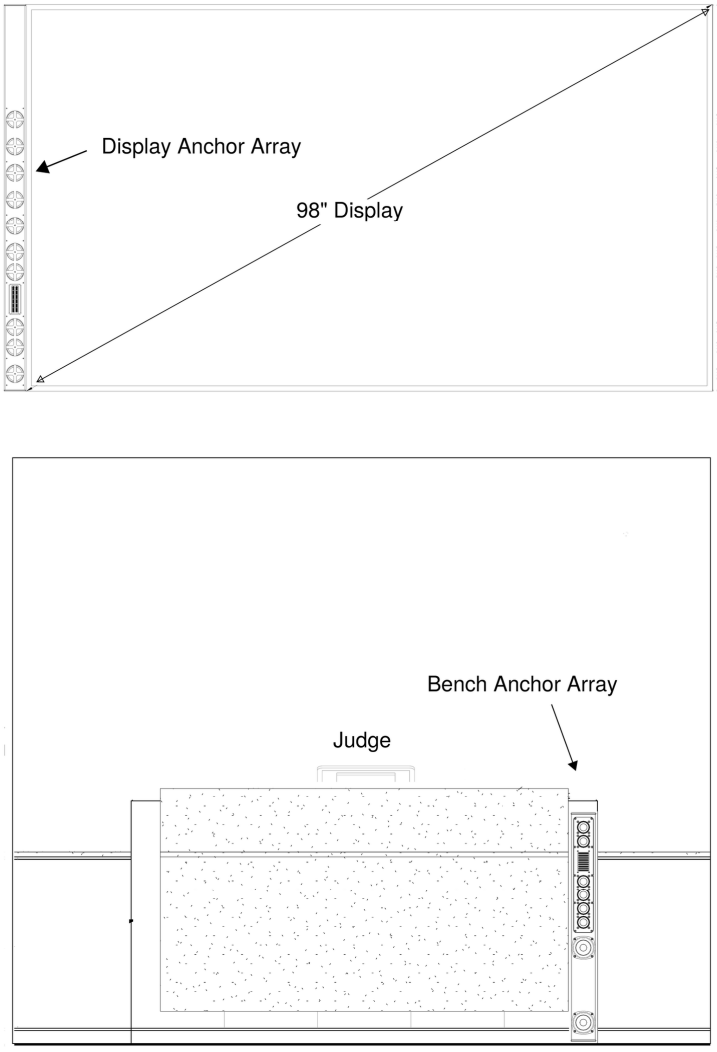
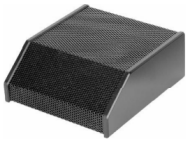


Figure 2. Elevations of the anchor loudspeakers. Top: the display anchor array, a vertical column loudspeaker mounted along the edge of the 98-inch videoconference display. Bottom: the bench anchor array, a free-standing, floor-mounted vertical column loudspeaker placed beside the judge's bench.

Desktop proximity loudspeakers at the bench, clerk, reporter, witness, and counsel positions. Small near-field loudspeakers at each position allow the judge, clerk, court reporter, witness, and attorneys to hear one another and remote participants at modest level. For the court reporter in particular, a clear near-field feed supports an accurate record. Each station receives a mix that excludes its own microphone, which limits feedback and permits lower overall gain. A digital signal processor synchronizes sound coming from the nearfield speakers to exactly match the flight-time of sound arriving from the anchor speakers. Near-field reproduction yields a high direct-to-reverberant ratio without raising the general sound level of the room.



Nano-monitor



Micro-lift

Figure 3. Examples of desktop proximity loudspeakers. Left: the Nano-monitor, a low-profile bar loudspeaker. Right: the Micro-lift, a compact angled loudspeaker that directs sound up toward the seated listener.

Overhead reinforcement for the jury and gallery. Because jurors are distant from most talkers, a dedicated overhead system above the jury box supplies sufficient level. Its signals are delayed so that sound from the relevant anchor arrives first; by the precedence effect, jurors then localize each voice to the talker even though much of the energy comes from above. A second set of in-ceiling proximity speakers above the gallery serves members of the public, the press, and parties awaiting their cases. These are time-aligned in the same way, so that voices remain anchored to the well rather than to the ceiling.

Supporting elements. The design should be paired with acoustic treatment to control reverberation and lower background noise from building systems, provision for an assistive listening system such as a hearing loop, and a clean feed to the court reporter and recording system. Intelligibility should be verified by measurement, for example using the Speech Transmission Index (IEC 60268-16), before and after installation.

Implications for the Bench

The proposed design asks courts to evaluate sound systems not by loudness or coverage alone but by whether listeners can readily attribute each voice to its source. Judges reviewing renovations or new construction may reasonably ask whether the system localizes voices to talkers, whether remote participants are heard from the display on which they appear, whether intelligibility has been measured, and whether jurors with hearing loss have been considered.

Conclusion

Research on selective auditory attention demonstrates that the brain's capacity to follow a single voice depends on spatial and visual cues that conventional ceiling-based systems discard. An anchored design, with reinforcement at the bench, at the remote-participant display, at counsel tables, and time-aligned speakers above the jury and gallery, restores those cues for both in-person and remote proceedings. The result is a courtroom in which listening demands less effort, so more attention remains for the evidence itself.

References

- Alais, D., & Burr, D. (2004). The ventriloquist effect results from near-optimal bimodal integration. *Current Biology*, 14(3), 257–262.
- Bregman, A. S. (1990). *Auditory Scene Analysis: The Perceptual Organization of Sound*. MIT Press.

- Broadbent, D. E. (1958). *Perception and Communication*. Pergamon Press.
- Bronkhorst, A. W. (2000). The cocktail party phenomenon: A review of research on speech intelligibility in multiple-talker conditions. *Acustica*, 86, 117–128.
- Cherry, E. C. (1953). Some experiments on the recognition of speech, with one and with two ears. *Journal of the Acoustical Society of America*, 25(5), 975–979.
- IEC 60268-16. Sound system equipment, Part 16: Objective rating of speech intelligibility by speech transmission index.
- Litovsky, R. Y., Colburn, H. S., Yost, W. A., & Guzman, S. J. (1999). The precedence effect. *Journal of the Acoustical Society of America*, 106(4), 1633–1654.
- Mesgarani, N., & Chang, E. F. (2012). Selective cortical representation of attended speaker in multi-talker speech perception. *Nature*, 485, 233–236.
- Pichora-Fuller, M. K., et al. (2016). Hearing impairment and cognitive energy: The Framework for Understanding Effortful Listening (FUEL). *Ear and Hearing*, 37, 5S–27S.
- Shinn-Cunningham, B. G. (2008). Object-based auditory and visual attention. *Trends in Cognitive Sciences*, 12(5), 182–186.
- Treisman, A. M. (1964). Selective attention in man. *British Medical Bulletin*, 20(1), 12–16.
- Wallach, H., Newman, E. B., & Rosenzweig, M. R. (1949). The precedence effect in sound localization. *American Journal of Psychology*, 62(3), 315–336.

 © 2026 SDA Consulting. All rights reserved.