Improving Dialogue Intelligibility in Media

A Report of the AES Technical Committee on Broadcast and Online Delivery

Co-chairs: Kimio Hamasaki, Matthieu Parmentier, Jim Starzynski, David Bialik

AES SHOW 2025
LONG BEACH

Working Group on Dialogue Intelligibility:
Chairman: Jim Starzynski
Editors: Jim Starzynski, Robert Orban, Robert Bleidt
Contributing Authors: Robert Bleidt, Jon Greasley, Hannes Muesch with Scott Norcross, Robert Orban, Jim Starzynski
Derek Barrentine, Hannah Baumgartner, David Bialik, Robert Bleidt, Manuel Briand, Richard Friedel, Jon Greasley, Rainier Huber, Scott Isabelle, Mark Johnson, Bob Katz, Scott Kramer, John Kean, Scott Levine, Steve Morris, Hannes Muesch, Scott Norcross, Robert Orban, Matthieu Parmentier, Sean Richardson, Jim Starzynski, Martin Walsh

AES Technical Council
Audio Engineering Society

Improving Dialogue Intelligibility in Media

A Report of the AES Technical Committee on Broadcast and Online Delivery

AES SHOW 2025
LONG BEACH

Technical Council
Audio Engineering Society

Broadcast and Online Delivery Working Group on Dialogue Intelligibility

Chairman: Jim Starzynski

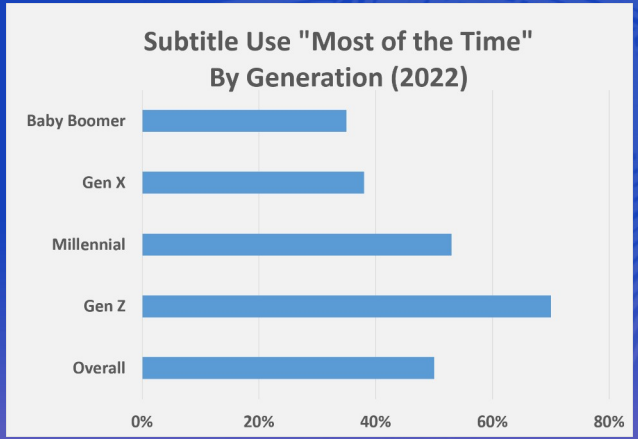
Editors: Jim Starzynski, Robert Orban, Robert Bleidt

Contributing Authors: Robert Bleidt, Jon Greasley, Hannes Muesch with Scott Norcross, Robert Orban, Jim Starzynski

Derek Barrentine, Hannah Baumgartner, David Bialik, Robert Bleidt, Manuel Briand, Richard Friedel, Jon Greasley, Rainier Huber, Scott Isabelle, Mark Johnson, Bob Katz, Scott Kramer, John Kean, Scott Levine, Steve Morris, Hannes Muesch, Scott Norcross, Robert Orban, Matthieu Parmentier, Sean Richardson, Jim Starzynski, Martin Walsh

Technical Council
Audio Engineering Society

How the Paper Came Together



Presenters:

- **Robert Bleidt (Streamcrest.com)**
- **Richard Friedel (formerly of Fox TV)**
- **Scott Kramer (greenfieldsound.com, formerly of Netflix)**
- **Scott Norcross (Dolby.com)**

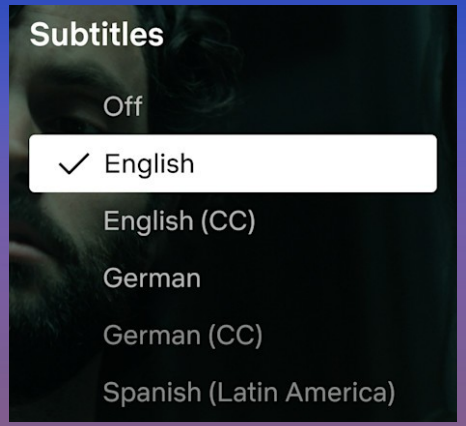
Reasons for Subtitle Use:

- **“Muddled Audio” or “Poor Audio Quality**
- **Can’t Hear Some Words**
- **Noisy Environment**
- **Accents**
- **Avoid Disturbing Others**

57% watch “in public”

Data from Preply survey 2022 and Associated Press survey 2025

The Dialogue Intelligibility Issue



Viewer Strain and Disengagement

Common Use of Subtitles: Use of subtitles also detracts from the visual quality and story engagement.

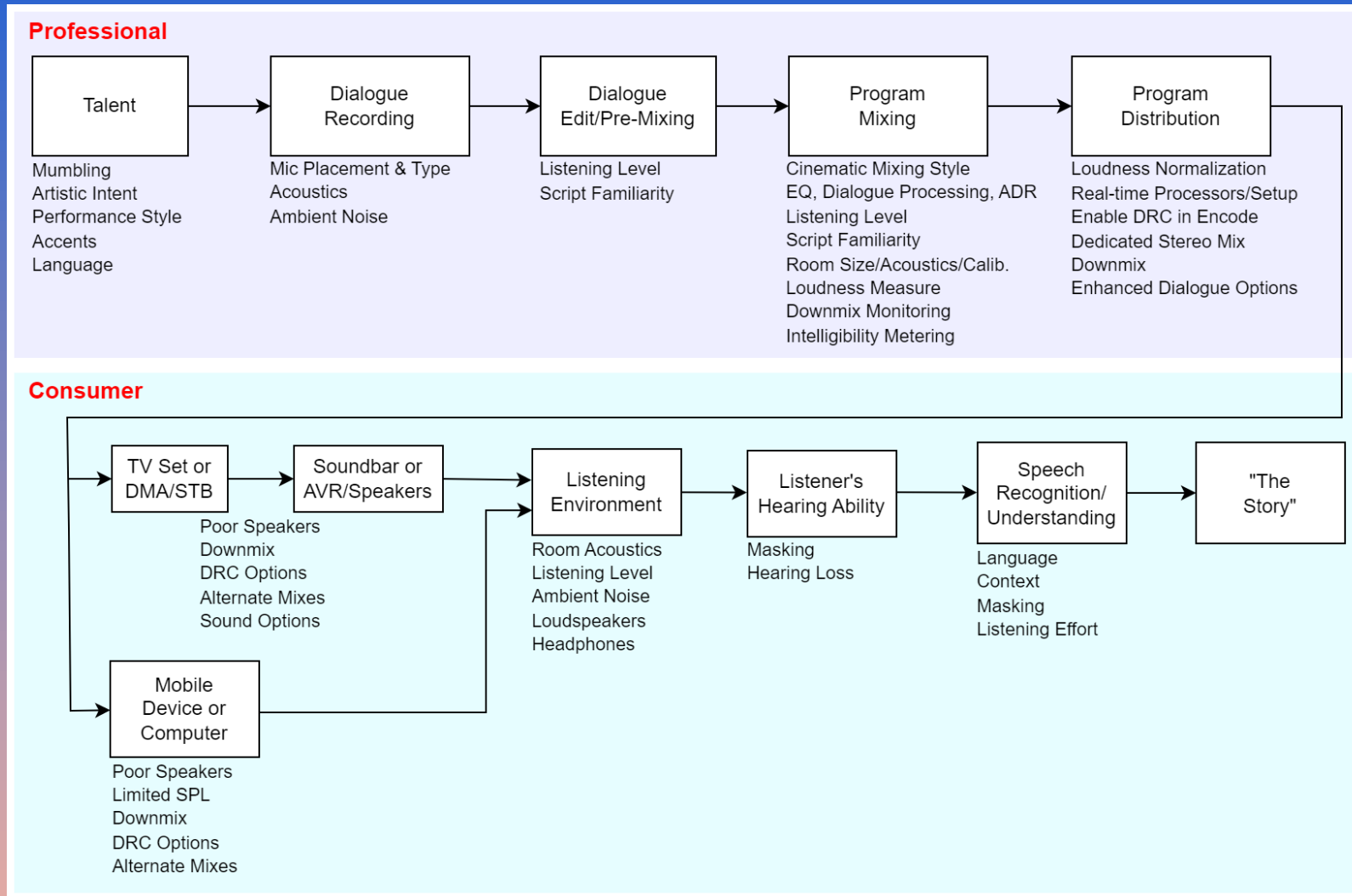
Evolution of Soundtracks: Expanded dynamic range in video soundtracks can compromise clarity.

Complex and Nuanced Issue

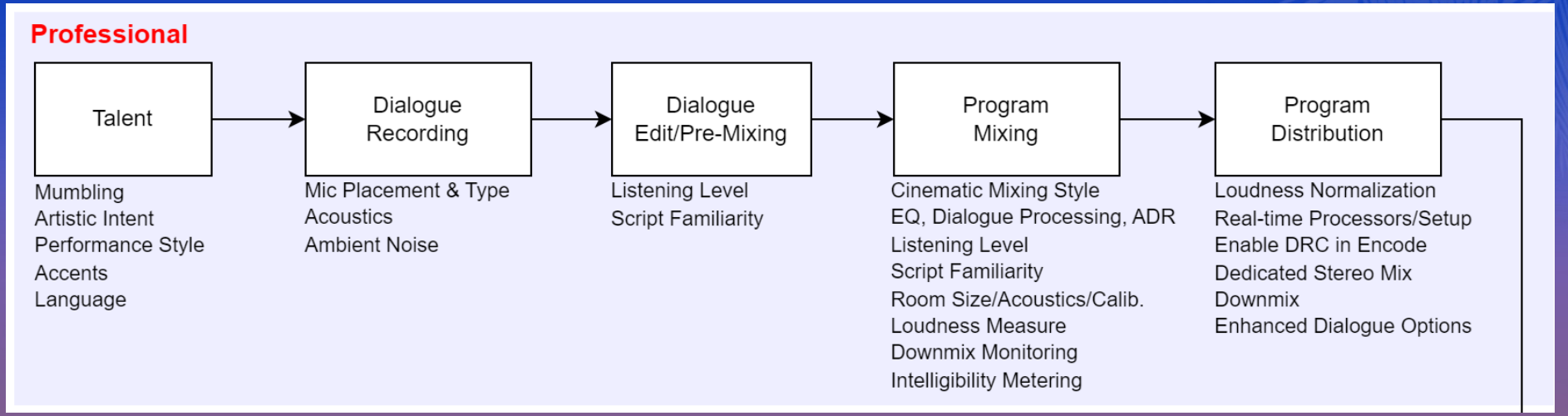
Achieving dialogue intelligibility is a shared responsibility across the entire content creation and delivery chain.

Netflix.com

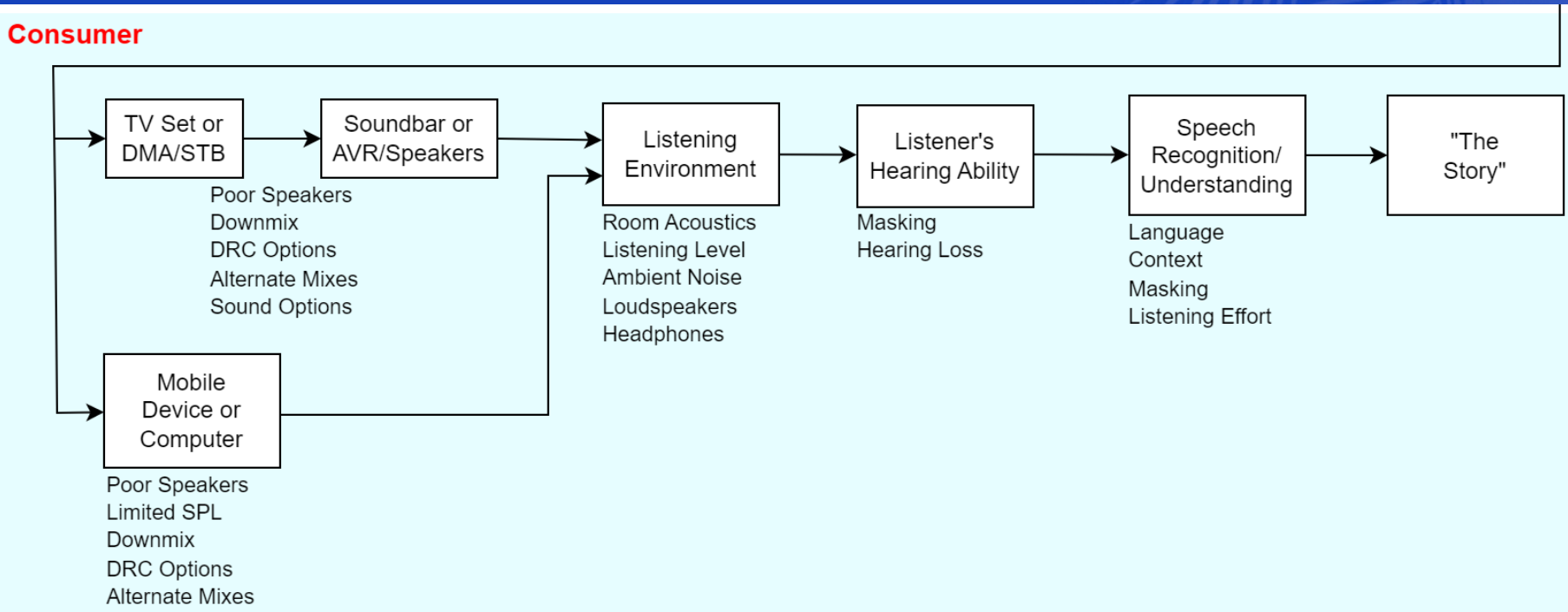
Technical Council
Audio Engineering Society


Professional
Talent
Mumbling
Artistic Intent
Performance Style
Accents
Language
Dialogue Recording
Mic Placement & Type
Acoustics
Ambient Noise
Dialogue Edit/Pre-Mixing
Listening Level
Script Familiarity
Program Mixing
Cinematic Mixing Style
EQ, Dialogue Processing, ADR
Listening Level
Script Familiarity
Room Size/Acoustics/Calib.
Loudness Measure
Downmix Monitoring
Intelligibility Metering
Program Distribution
Loudness Normalization
Real-time Processors/Setup
Enable DRC in Encode
Dedicated Stereo Mix
Downmix
Enhanced Dialogue Options
Consumer
TV Set or DMA/STB
Poor Speakers
Downmix
DRC Options
Alternate Mixes
Sound Options
Soundbar or AVR/Speakers
Mobile Device or Computer
Poor Speakers
Limited SPL
Downmix
DRC Options
Alternate Mixes
Listening Environment
Room Acoustics
Listening Level
Ambient Noise
Loudspeakers
Headphones
Listener's Hearing Ability
Masking
Hearing Loss
Speech Recognition/ Understanding
Language
Context
Masking
Listening Effort
"The Story"

Factors Impacting Intelligibility

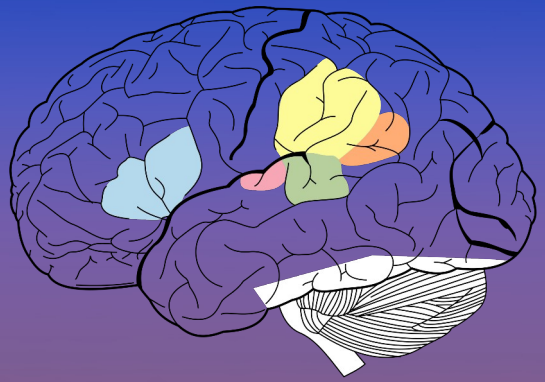


Factors Impacting Intelligibility



Technical Council
Audio Engineering Society

Factors Impacting Intelligibility - Human Hearing



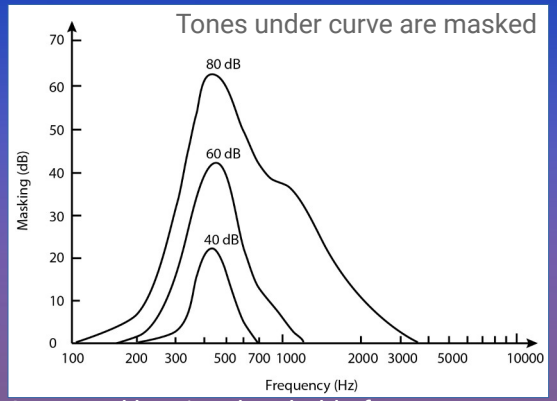
Auditory processing areas of the brain

Intelligibility vs. Listening Effort:
Intelligibility is relative. Significant listening effort can lead to fatigue and irritation. As speech signals become clearer, listening effort decreases.

Bottom-Up and Top-Down Processing:

a. "bottom-up" processing (decoding acoustic cues) and
b. conscious "top-down" processing (using context and expectations).
c. When acoustic cues are degraded, listeners rely more on top-down processing, hurting comprehension.

Factors Impacting Intelligibility - Human Hearing



Increased hearing threshold of pure-tone signal masked by 410 Hz narrow-band noise

Masking:

a. Competing sounds (music, effects, or environmental noise) in the same frequency range
b. Heavy bass can particularly mask higher-frequency dialogue

Reverberation: Reduces intelligibility by smearing the speech envelope.

Listener Characteristics:

a. Hearing loss (even with normal audiograms)
b. Age (children and elderly)
c. Non-native language listening
d. Intelligibility is an individualized experience
e. Audience demographics and listening environment must be considered
f. A single intelligibility measurement is insufficient

Adapted from Egan and Hake, as reprinted in Moore, *Psychology of Hearing*, 5ed., p88

Technical Council
Audio Engineering Society

Factors Impacting Intelligibility - Production & Post

Subdued Acting & Noisy Recordings

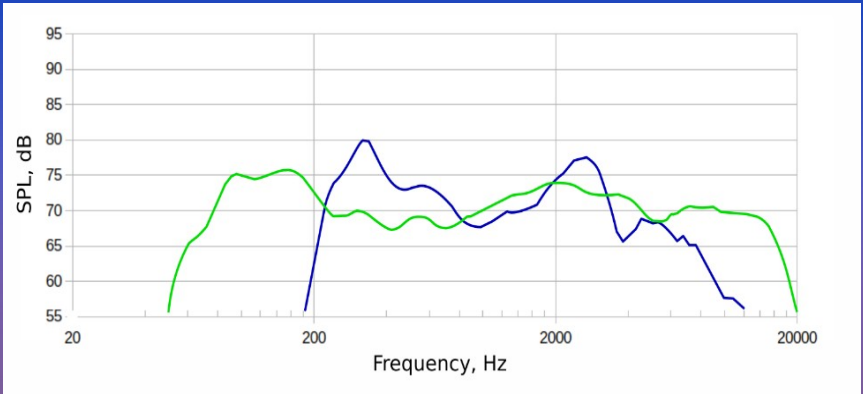
Cinematic Mixes for Home Viewing: Mixes too dynamic for comfortable home listening.

Conflicting Loudness Recommendations: Loudness recommendations vary and are often misinterpreted or misapplied. Example: Full program vs Dialogue based specifications.

Loudness Range (LRA) Targets: LRA can change unpredictably and does not account for genre or dialogue level.

Misuse of Dialogue Processing Tools: Can render quieter parts of words inaudible or cause unintended "ducking" of dialogue.

Factors Impacting Intelligibility – Distribution and Devices/Environment



Frequency response of the best and worst 2023 TVs measured by rtings.com

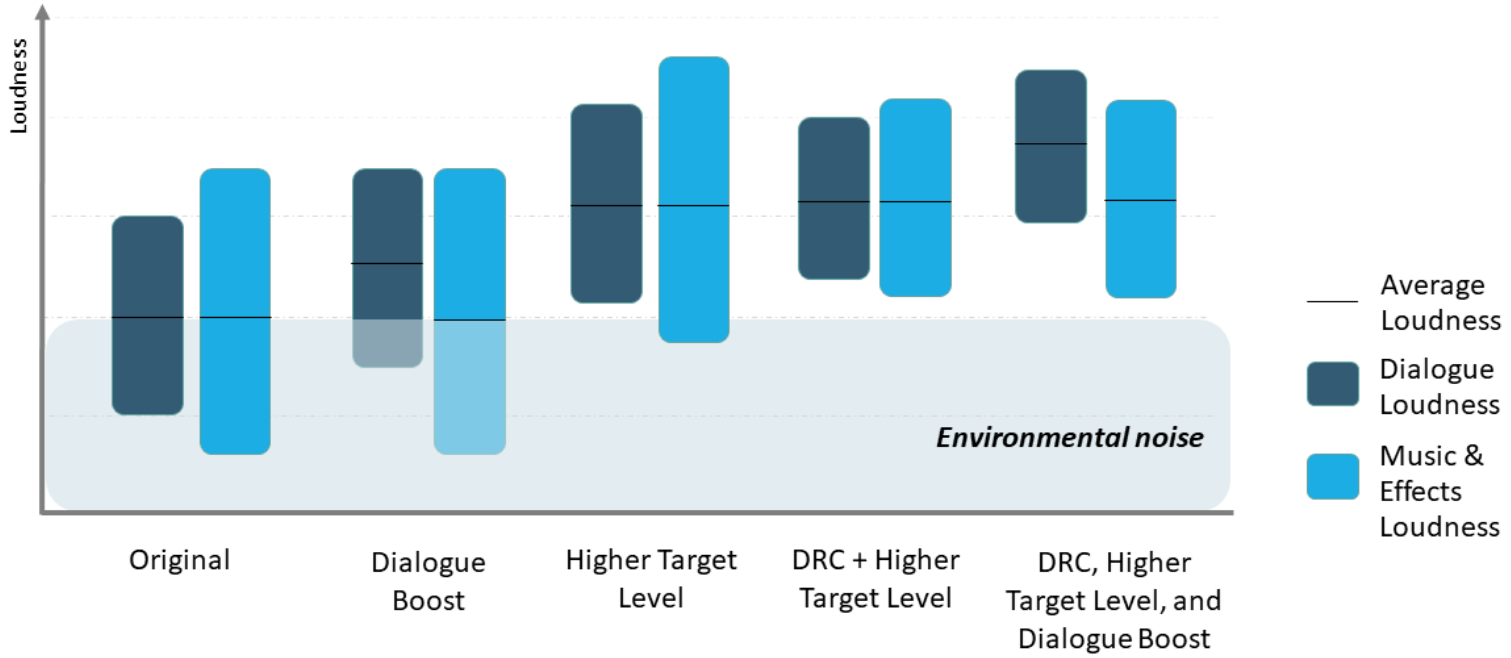
Lower Consumer Playback Levels

Inconsistent Loudness Management: Without proper loudness normalization, consumers are fatigued by constantly adjusting volume controls.

Downmixing Challenges: Leads to content buildup

Technical Council
Audio Engineering Society

Target Level, Dialogue Boost & Dynamic Range Control

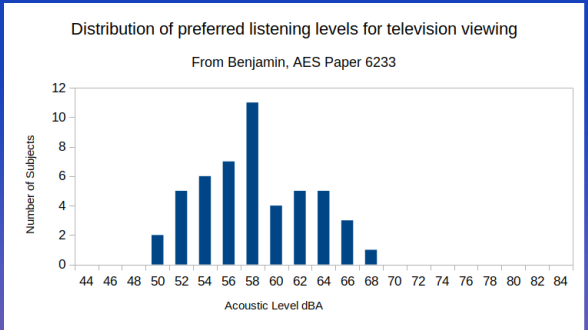


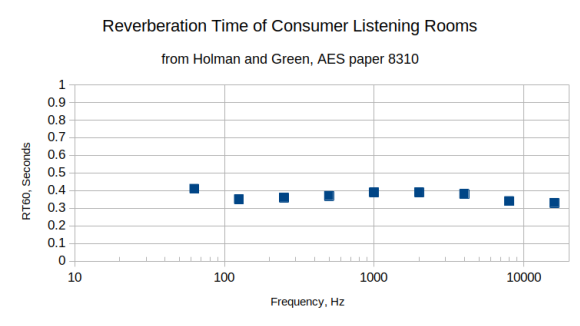
Demonstration

- Original Mix – courtesy Jon Greasley
- Traditional DRC – Film Standard
- Strong modern DRC
- Separated and boosted dialogue
- Boosted dialogue and Strong modern DRC

- Recorded airport ambience added to all examples to simulate noisy conditions
- Nuendo and Nugen Intelligibility Meter

What Can Go Wrong At Home (Even for Professionals)





Studies have shown consumers listen at much lower levels (SPL) than professionals.

Reverberation times of consumer's listening environments are in the range of 400-600 ms.

Consumers routinely have noise sources (dishwasher, clothes dryer, HVAC noise, other occupants) that would not be tolerated in a professional setting.

Consumers typically have limited maximum SPL due to equipment, family, and neighbors.

To check intelligibility, monitor at consumer levels (60 dBA), perhaps with masking noise (35-40 dBA) added.

Distributors should ensure that Dynamic Range Control is enabled for codecs that support it.

What Can Go Wrong At Home – The Downmix

Speakers from “best sounding TV” in 2023 test at rtings.com. Others may be much smaller

According to major streaming distributors, 90% of consumers listen over their TVs in stereo.

Unless a streaming distributor offers a separate stereo mix, consumers hear an automatic downmix. (Always with broadcast/cable/satellite TV)

Automatic stereo downmixes do not include the LFE channel

Monitor the downmix using small loudspeakers with flat response

Technical Council
Audio Engineering Society

What do we advise consumers (& us) for home listening?

Which decoder are you listening to? Always try to select “bitstream passthrough” if you have a soundbar or AVR.

Consider buying a soundbar. Note that inexpensive soundbars will be stereo, even if they support surround or immersive decoding. A center channel speaker can be a big upgrade.

Experiment with TV settings such as DRC (“night mode”). DRC will help avoid “volume control riding”.

Try turning TV post-processing on/off. It may help or hurt.

Use automatic EQ setup if available

Investigate if your content offers other mixes

Realize that options may vary with content and provider.

Technical Council
Audio Engineering Society

Recommendations for Improving Dialogue Intelligibility

Use Dialogue-Based Loudness Anchors

- Effective means of establishing dialogue intelligibility
- A single industry-wide dialogue-based loudness delivery specification would be beneficial

Prioritize Stereo and Downmix Intelligibility

- Most consumers listen in stereo!
- Carefully check the derived stereo downmix; give it equal priority while mixing
- Optimizing stereo intelligibility benefits all formats

Cautious Use of Dynamics Processing

- Over-use can cause listener fatigue and impair intelligibility
- Attention should be paid from initial capture through distribution

Recommendations for Improving Dialogue Intelligibility

Mix for Consumer Listening Conditions

- Monitor mixes at lower listening levels (e.g. 58-65 dB SPL)
- Use speakers the approximate home listening experience
- Perform quality control (QC) emulating consumer conditions, including adding background noise

Encourage consumers to experiment and perhaps upgrade

- Experiment with their device's processing and DRC
- Consumers should add a center channel speaker if not already using one
- Low cost soundbars often only have left and right speakers

Foster a Culture of Intelligibility

- Encourage clients, showrunners, directors & producers to prioritize intelligibility
- Focus on intelligibility while developing cinematic experiences
- Ensure appropriate resources for sound capture and mixing

Technical Council
Audio Engineering Society

Thank you!

Coming Soon:
AES TD1009 – Improving Dialogue Intelligibility in Media

Kind Regards,

AES Technical Committee on Broadcast and Online Delivery
Working Group on Dialogue Intelligibilty

Contact us at broadcast@aes.org

Presentation and link to report available at: rb.gy/lg2js8