



What Did They Just Say?

How Machine Learning Gives TV Audiences
the Audio Experience They Want

Table of Contents

Executive Summary	03
Introduction: The Growing Challenge of Dialogue Intelligibility	04
Why Dialogue Gets Lost in Translation: Understanding the Problem	05
The Business Case for Improved Dialogue Intelligibility	07
Solutions Overview: From Conventional Fixes to Advanced AI and ML Solutions	09
Introducing DTS Clear Dialogue	10
Strategic Considerations for OEMs: Choosing the Right Partner for Dialogue Intelligibility Solutions	13
Solving Dialogue Intelligibility with DTS Clear Dialogue	15

Executive Summary

As TVs evolve to offer larger screens and sharper visuals, one key feature has taken a hit: sound quality. Despite the stunning picture, hearing and understanding dialogue has become more challenging. Dialogue intelligibility – or the ability to clearly hear and understand spoken words while watching TV or movies – is an ongoing frustration for global audiences. From a strategic and technical perspective, solving the dialogue intelligibility problem is both a challenge and an opportunity for TV manufacturers.

If consumers are unhappy with a TV's sound quality, manufacturers face declining brand loyalty, lower engagement with smart TV platforms, and missed opportunities to differentiate products in a crowded marketplace. However, new advancements in machine learning (ML) provide innovative solutions for understanding dialogue. Manufacturers that adopt these emerging solutions, like DTS Clear Dialogue™, can deliver superior audio clarity, meet consumer expectations, and gain an important competitive advantage in the crowded home entertainment space.

**In this e-book,
TV manufacturers
will learn:**

- ▶ Why traditional solutions like subtitles and volume adjustments fall short of consumer expectations and fail to solve dialogue intelligibility issues.
- ▶ How ML-powered technologies are transforming dialogue clarity through precision, personalization and accessibility.
- ▶ The business case for TV manufacturers to invest in innovative sound solutions to differentiate products and drive customer satisfaction.
- ▶ What to look for when selecting a strategic solution for dialogue intelligibility.

Introduction: The Growing Challenge of Dialogue Intelligibility

Imagine this: A viewer is watching the season finale of their favorite TV series. The stakes are high, and the lead character is about to deliver the big reveal. But just as dialogue is spoken, it's drowned out by a swelling musical score. They lean forward, turn up the volume, and wonder: What did they just say?

This is the last thing manufacturers want consumers to experience when watching content on their devices. However, a recent Dynata survey revealed that 84% of viewers have struggled to understand dialogue while watching TV.¹ The problem is even more pronounced for older viewers, with a study from the Fraunhofer Institute finding that the number rises to 90% for people over 60.² As dialogue clarity diminishes, viewer satisfaction with their devices drops. This isn't just consumer annoyance; it's a business risk for manufacturers. Distracted and dissatisfied viewers are less engaged, less likely to interact with platforms, and less inclined to see value in their smart TVs. Over time, this can impact brand loyalty and monetization opportunities for TV manufacturers.



84%

of viewers have
struggled to
understand
dialogue while
watching TV

The good news is that dialogue intelligibility isn't an unsolvable problem. While traditional solutions have fallen short on effectively addressing this problem, advancements in ML for audio post-processing offer new implementations that eliminate frustration for consumers and create differentiation for TV manufacturers. Addressing dialogue intelligibility can ultimately increase product quality, improve customer satisfaction, and drive revenue growth.

Let's explore the root causes of dialogue intelligibility and why traditional solutions have failed to measure up.



Why Dialogue Gets Lost in Translation: Understanding the Problem

Dialogue intelligibility refers to how easily audiences can hear and comprehend spoken words in TV shows, movies, and other entertainment content. It's not necessarily a loudness issue or a volume problem.

Instead, it focuses on how clearly speech – or what's usually referred to as "dialogue" in the entertainment industry – can be understood without requiring extra effort from the listener. Ideally, viewers aren't increasing the volume or watching TV with subtitles. Good dialogue intelligibility lets viewers stay engaged with the story, understanding critical details without struggling with issues like subpar audio, excessive background noise, hearing limitations, or environmental distractions.

In today's home entertainment environment, no single factor makes it hard to hear dialogue. Instead, there are a myriad of causes that are important to understand:

Industrial Design and Speaker Constraints:

Modern ultra-thin TVs lack the physical space required for high-quality built-in speakers. Manufacturers often compensate for this design tradeoff with digital signal processing (DSP) algorithms, but these may unintentionally make the dialogue more difficult to understand.

Artistic Intent and Sound Mixing:

Directors and content creators often prioritize immersive soundscapes to accompany stunning pictures. Content mixed in ideal studio environments with professional studio equipment systems may not translate effectively to two-channel home environments. Downmixing immersive tracks into stereo may also introduce unintended masking effects, as dialogue competes with background sounds and music.

Shrinking Production Budgets and Accelerated Timelines:

Content production today operates within tighter budgets and faster delivery requirements than ever before. These constraints leave less time to optimize audio tracks for various real-world listening environments. As a result, audio may not sound as originally intended in everyday home settings.

Environmental Noise:

Real-world listening environments present significant challenges. Household noise – appliances, HVAC systems, animals, and street traffic, to name a few – masks dialogue,

especially in open-plan spaces with hard surfaces that amplify reflections. Noise constraints, such as sleeping children or close neighbors, can further limit volume levels and make dialogue even harder to discern.

Language, Accents, and Rapid Speech:

Language barriers, diverse accents, and rapid-fire dialogue create additional hurdles for many viewers all over the world, as they have access to content created anywhere. Original audio may be in a language or with a dialect that is unfamiliar to the viewer.

Aging Populations and Hearing Impairments:

Hearing loss is a growing issue, with 90% of viewers over 60 reporting difficulty understanding TV dialogue "often" or "very often." According to the World Health Organization (WHO), by 2050, roughly 2.5 billion people worldwide will have some degree of hearing loss, creating an urgent need for individualized audio solutions that address varying stages of impairment.³ Current personalization features fail to serve these audiences adequately.

Fragmented Content Sources: Content is consumed across streaming platforms, cable, and physical media, each with its own compression and delivery standards. Switching between different content sources can create inconsistent audio quality, forcing audiences to adjust settings frequently or rely on stopgap solutions like subtitles.



The Business Case for Improved Dialogue Intelligibility



94%

of U.S. viewers
rank sound
quality as critical

Audio clarity has become a core expectation for today's TV audiences. A survey from Opinium research and Xperi Inc. found that 94% of U.S. viewers rank sound quality as critical to their viewing experience, outpacing screen size and smart capabilities.⁴ It's a concern that spans demographics. Baby boomers facing age-related hearing challenges want clearer dialogue to maintain their connection to stories. Younger viewers want seamless audio that doesn't interrupt multitasking.

Yet, consumers are relying on imperfect solutions to cope. For example, 77% of viewers reported using subtitles to follow dialogue in a recent Dynata survey.⁵ However, subtitles disrupt the immersive experience, shifting attention away from the visuals and adding cognitive effort to what should be an effortless activity. Subtitles are a workaround for an issue that demands a more sustainable solution.



In the same survey from Dynata that was previously mentioned, it was uncovered that nearly 38% of viewers want dedicated remote buttons or user-friendly features to adjust dialogue clarity without the frustration of navigating complicated menus.⁶ Highlighting there's a growing demand for simple solutions to this common pain point.

For manufacturers, addressing dialogue intelligibility is a strategic opportunity to lead in a competitive market. Mid-market manufacturers can achieve several audio innovation goals, including:



77%

of viewers reported
using subtitles to
follow dialogue



38%

of viewers
want dedicated
remote buttons or
user-friendly features

- ▶ **Attract Premium Consumers:** As high-end buyers and technology adopters focus on AI-driven audio enhancements, manufacturers can position their devices as premium offerings with exceptional sound.
- ▶ **Differentiate Mid-Market Offering:** Buyers want clear audio at all price points. Exploring solutions such as ML post-processing options can allow manufacturers to roll out new features in a highly price-competitive TV market.
- ▶ **Expand Accessibility:** Accessibility features can differentiate products and expand market reach to a variety of populations, including aging TV audiences, non-native speakers and neurodiverse individuals.
- ▶ **Boosting Engagement and Monetization:** Clearer dialogue encourages viewers to engage more with TV platforms by watching more content. Stronger engagement drives viewership, which in turn increases opportunities to drive post-sale revenue through advertising experiences.

For manufacturers, embracing innovation in audio clarity is a chance to lead the next frontier of home entertainment.

Solutions Overview: From Conventional Fixes to Advanced ML Solutions

Dialogue intelligibility challenges have led to the development of diverse solutions, which have addressed the issue with varying levels of success. Traditional approaches like dynamic range compression (DRC) and speech enhancement filters offer some relief but fall short. However, advancements in ML-powered solutions for post-processing audio are shifting the conversation from why traditional methods haven't worked well to what's possible today.

Conventional Approaches: Why They Often Miss the Mark

Various approaches have been used to make dialogue easier to understand. DRC is commonly used to manage loudness by lowering the volume of loud sounds and amplifying softer ones. While this technique helps balance audio dynamics, poorly tuned algorithms can create jarring "pumping" effects or even suppress dialogue.

Many TVs include dialogue enhancement features that boost frequencies where speech is most dominant, typically between 125Hz and 5,000Hz. While these filters can improve speech clarity, they often affect non-speech elements in the mix, like background music and effects.

Individual content services like Amazon Prime Video have introduced features like Dialogue Boost, offering alternative mixes optimized for dialogue clarity. However, these solutions are limited in scope and often restricted to specific languages (e.g., English) so their impact is limited.

Tailoring soundtracks for individualized preferences is often intensive, as it requires separating dialogue, music, and effects (DME) into discrete tracks. Significant barriers exist, including increased storage demands and the impracticality of creating separate mixes for every viewing scenario.

Machine learning's rapid evolution has created another path. Unlike traditional DSP approaches, it's possible to train ML models to identify acoustic feature patterns. The model can then identify individual parts such as dialogue, music, and sound effects. From there, the solution can learn to "unmix" the audio for greater flexibility and personalization.



Introducing DTS Clear Dialogue

DTS has decades of experience partnering with TV manufacturers to develop innovative solutions to audio challenges. The knowledge gained from these experiences was used to develop DTS Clear Dialogue, a proprietary machine-learning solution designed to solve dialogue intelligibility challenges in a new way.

DTS Clear Dialogue utilizes an optimized deep-learning network trained on diverse datasets to help the model generalize to different scenarios. By addressing the complex interplay of speech, background music, and effects within soundscapes, it's possible to enhance the clarity of spoken languages.

The solution offers several unique benefits for TV manufacturers and their viewing audiences:



Advanced Audio Separation:

DTS Clear Dialogue isolates speech from other audio elements, such as music and sound effects. This precision enables independent processing of dialogue, maintaining clarity even in complex soundscapes. By isolating dialogue and amplifying it without distorting the overall audio mix, DTS Clear Dialogue ensures a frustration-free viewing experience.



Real-Time Personalization:

Audio preferences and needs vary widely among users. DTS Clear Dialogue offers dynamic, user-controlled adjustments. Viewers can personalize their audio experience through intuitive features like voice commands, on-screen controls to adjust dialogue levels, or dedicated remote buttons that enhance dialogue on request. Flexible interface choices let manufacturers select the best custom delivery for their brand and target demographics.



Content and Language Inclusivity:

DTS Clear Dialogue is designed to work across languages, content types and sources, from streaming Hollywood blockbusters to live sports. This content-agnostic approach ensures consistent performance for diverse global audiences and content. A content-aware solution ensures consistent audio quality for the viewing experience, regardless of what's being watched or where it is being watched from.



Greater Engagement and Monetization Opportunities:

Designed with inclusivity and accessibility in mind, viewers no longer need to juggle remotes, subtitles or manual volume adjustments. The focus on accessibility and creating a more enjoyable viewing experience encourages viewers to spend more time on their TVs and platforms.



Future-Proof Technologies:

As consumer expectations for high-quality audio experiences grow, integrating DTS Clear Dialogue helps manufacturers remain at the forefront of innovation. In the short term, a flexible solution to clear dialogue is a competitive edge. Early adopters will stay on the cutting edge as the solution matures and sets a continually rising bar for audio performance.

DTS Clear Dialogue isn't just a feature. It's a comprehensive solution that solves fundamental pain points for consumers and presents a strategic opportunity for TV manufacturers. By investing in this technology, manufacturers can deliver superior audio clarity, build stronger brand loyalty, and differentiate their products in an increasingly crowded market.

Strategic Considerations for OEMs: Choosing the Right Partner for Dialogue Intelligibility Solutions

Taking action to solve dialogue intelligibility challenges can set a brand apart. However, selecting the proper dialogue intelligibility solution is critical for TV manufacturers that want to offer an outstanding viewing experience while keeping costs and technical complexities down. Vetting both the solution and the provider as a potential partner can help reinforce strong investments.

Several core questions can be considered to assess whether a solution is the right fit. Ideally, the right solution will fix key challenges, adapt to diverse user needs, and align with market demands for innovative and accessible audio technology. Assess potential partners and technologies based on the following criteria and questions:

1. Content Compatibility:

Does the solution work seamlessly across streaming platforms, broadcast services, and physical media? Can it handle varied content formats, including immersive audio tracks and downmixed stereo configurations?

2. Language Inclusivity:

Does it process multiple languages and accents without requiring additional tuning or content-specific adjustments? Is it truly content-agnostic, ensuring consistent performance across diverse global markets?

3. Ease of Use:

Are interfaces intuitive and user-friendly, offering features like dedicated remote buttons, sliders, or voice commands? Can the solution be tailored to user needs, from accessibility to advanced personalization?

4. Performance in Real-World Listening Environments:

Can the solution adapt dynamically to challenging listening environments, such as noisy households or level-constrained spaces? Does it address the diverse needs of users, from aging populations with hearing impairments to tech-savvy younger audiences?

5. Partnership:

Does the brand behind the solution have a track record of experience in audio processing? Do they understand the TV manufacturing landscape and have solutions that fit in effortlessly? Are they prepared to help manufacturers navigate broader questions, or taking the steps needed to disclose AI use in connection with emerging regulations?

Adopting ML-powered audio solutions is both a technical upgrade and a strategic move. Solutions like DTS Clear Dialogue differentiate products in a competitive market, enhance customer satisfaction, and address one of the most frequent consumer complaints: difficult to understand dialogue. More critically, ML-powered solutions offer scalable, hardware-compatible solutions. These technologies provide a user-first experience that delivers on both accessibility and personalization. Taking the key steps now to invest in these solutions can help future-proof devices while meeting fast-evolving consumer demands.

Solving Dialogue Intelligibility with DTS Clear Dialogue

As now proven, dialogue intelligibility impacts nearly all TV viewers and can be a complex challenge to solve for. Manufacturers who are dissatisfied with the solutions on the market have a new option to consider.

DTS Clear Dialogue stands out as the strategic choice for manufacturers with its ability to deliver personalized, content-agnostic audio clarity that integrates seamlessly into devices for global markets. Manufacturers have an option that is scalable, efficient and a globally applicable solution that can enhance revenue and engagement on TV platforms without adding significant costs.

Now is the time to reach high-performance audio standards, keeping viewers focused on the dialogue and not fussing with controls. Manufacturers that invest in solving key user challenges like difficult-to-understand dialogue will capture engagement, revenue, and customer loyalty while staying on the cutting edge of competitive differentiation in a crowded market.

For manufacturers ready to explore how ML audio processing solutions like DTS Clear Dialogue can transform their devices, [contact us today](#) to begin the conversation.



0.4208



DTS Clear Dialogue

Improving the television experience by
making dialogue more intelligible

FIND OUT HOW

