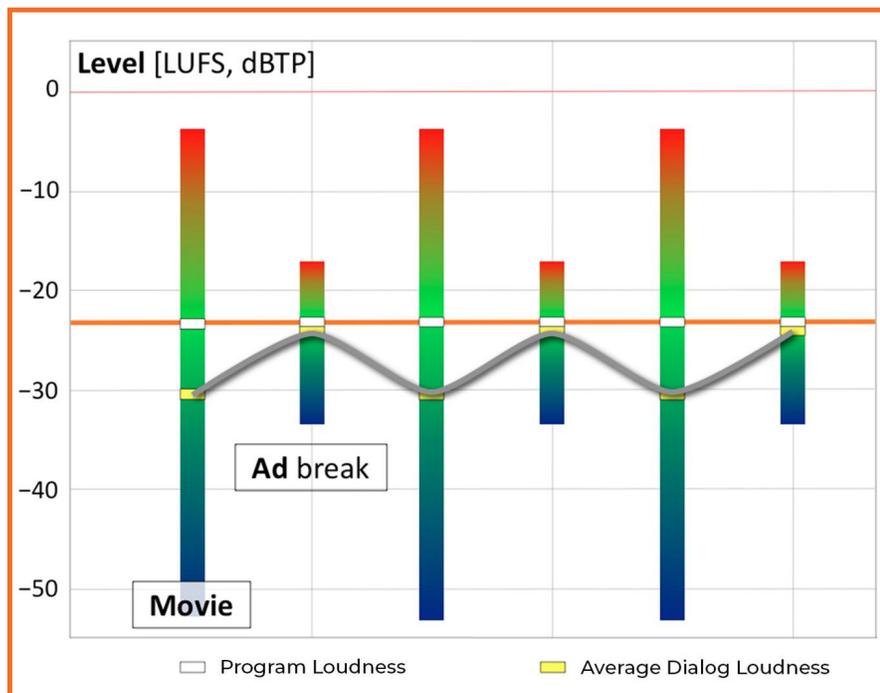


White Paper

The Challenge Of Dialog Across Modern Platforms

Inconsistent dialog level between programs still pose a challenge to today's audiences. AudioTools Server has the tools to fix this and more.



EBU Tech 3343 Guidelines for Production of Programmes in Accordance with EBU R 128. <https://tech.ebu.ch/docs/tech/tech3343.pdf>

The Challenge of Dialog Intelligibility

The EBU R128 recommendations have sought to address variations in audio loudness between different content by recommending that all audio be normalized to a consistent level based on its overall integrated (Program) loudness value of -23LUFS. The drawback of using the overall program loudness alone as the target element is that it doesn't take into account the single most important aspect of most television audio: Dialog.

To understand why this can be a problem, let us consider three different program genres: A documentary, a TV drama, and an action movie or other cinematic content.

The Problem:

While audio specifications, such as EBU R128 and ATSC A/85, have helped to address variations in overall program loudness between different content, they do not take into consideration the relationship between the overall loudness and the dialog level. Newer formats, designed for today's modern streaming services, require better control over Dialog to Program loudness values, as well as the need to switch between Dialog and Program anchor based upon the amount of dialog present.

The Solution:

AudioTools Server's Advanced Loudness Adaptation module can help to address these variations. Its sophisticated loudness engine (powered by our Linear Acoustic APTO™ technology) is able to target and adapt the difference between Program and Dialog loudness, as well as a host of other advanced loudness controls, all while maintaining the creative intent of the original mix. When content suffers from issues with dialog intelligibility, a new technology from Fraunhofer has been incorporated into AudioTools Server that can be an alternative solution: Dialog+.

A documentary program will likely be very dialog-driven, meaning that the difference between the measured overall program loudness and the measured overall dialog loudness will probably be minimal (perhaps even less than 1LU).

The TV drama will likely be more dynamic, with the difference between the overall program and dialog loudness sitting somewhere around 3 to 4LU.

The action movie, on the other hand, could conceivably have a difference of over 10LU or more between the average dialog and program loudness values, as cinematic content is generally mixed with a wider dynamic range than is desired for TV. If, for example, the viewer has set the volume of their TV to a comfortable level for watching the documentary and then switches over to watching a movie, they are going to experience a significant drop in the dialog level (and a corresponding reduction in intelligibility) despite the overall loudness of the two programs being the same. The television audience has voiced its frustration and unhappiness about these loudness shifts by way of ongoing viewer complaints. While recommendations have gone a long way toward improving level shifts, they have yet to solve all the loudness issues the original specifications were designed to achieve.

New Standards Pose New Challenges

With the rise in the popularity of video on demand, many of the major streaming platforms are now taking a different approach to loudness requirements, with “speech-gated” dialog loudness targets becoming increasingly common for audio on their platforms. In addition, recommendations for things like Loudness Range (LRA), “Dialog-LRA” (the dynamic range of the dialog component), and the maximum program loudness concerning the dialog (the “Loudness-to-Dialog Ratio”) are also increasingly being used. The EBU has acknowledged this with the publication of R 128 s4 in November 2023, “Loudness Normalization of Cinematic Content for Broadcast.”

One of the drawbacks of speech-gated loudness normalization is how to handle situations where the content does not contain much dialog, such as a classical music concert. Loudness specifications, such as those required by [Netflix](#)¹ or the [Disney Near Field Standard](#)², get around this by specifying two different target options: -27 LKFS (dialog anchor) as the primary target and -24 LKFS (program anchor) if the measured content contains less than 15% dialog. This use of multiple loudness targets, conditional on the amount of detected dialog, along with recommendations for Dialog-LRA (the maximum Program-to-Dialog difference), poses a challenge to most automated loudness solutions, which cannot easily handle these operations.

This is where AudioTools Server’s Advanced Loudness Adaptation module comes into its own.

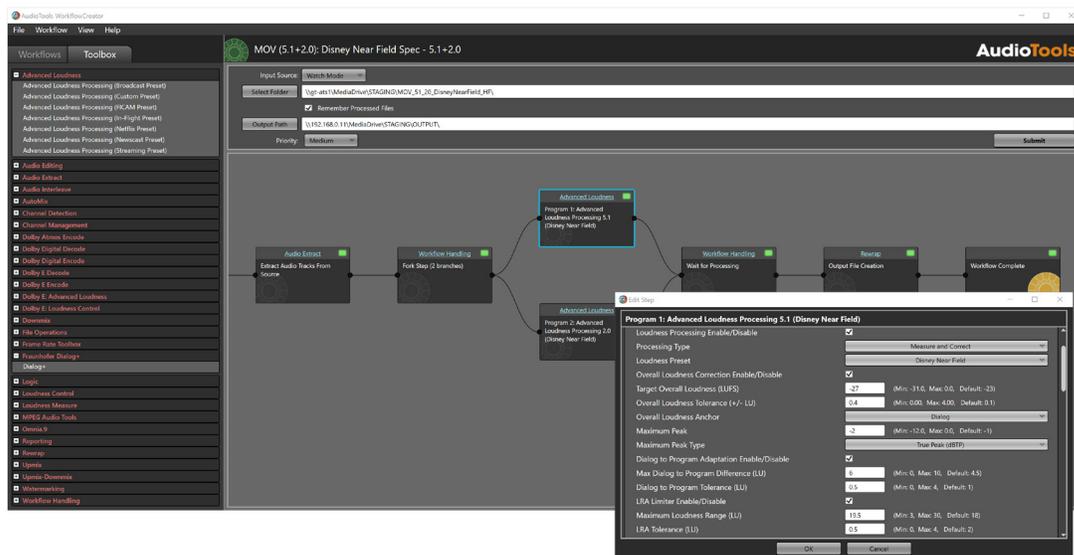
¹ <https://partnerhelp.netflixstudios.com/hc/en-us/articles/360001794307-Netflix-Sound-Mix-Specifications-Best-Practices-v1-4>

² <https://mediatechspecs.disney.com/#/mastering/audio/audio-specifications>

Modern Problems Require Modern Solutions

AudioTools Server’s Advanced Loudness Adaptation

The optional Advanced Loudness Adaptation module for AudioTools Server offers users a much more advanced toolset of loudness controls, including the ability to set independent primary and secondary overall loudness targets, conditional on the amount of detected dialog, a key requirement of the Netflix and Disney loudness specifications. Controls for defining the maximum difference between the measured program and measured dialog (“Loudness-to-Dialog Ratio”), the dynamic range of the dialog component (“Dialog-LRA”), and Maximum/Minimum Short Term Dialog loudness are also provided, each with their own target and tolerance values. More commonly found settings, such as Overall Loudness and Maximum True Peak, are also present, making Advanced Loudness Adaptation the perfect tool for providing great sounding and dynamically controlled audio that can meet all the latest loudness specifications.



Relying on the state-of-the-art loudness processing provided by Linear Acoustic APTO™, the audio signal is carefully controlled to preserve the source content's transients, sonic image, and artistic intent while increasing dialog intelligibility and loudness consistency. Consequently, a dynamic and balanced mix can be tailored to any destination platform, from theatrical to mobile, without resulting in subjectively perceived compressed sound. A sophisticated logic engine determines the order in which the different correction types are processed, ensuring maximum efficiency. With support for multiple correction passes built in, the module can handle even the most challenging audio. Audio deemed already compliant for the intended use case will pass through the workflow untouched.

For users who don't want to define all of the loudness controls themselves, the Advanced Loudness module comes with a comprehensive set of pre-configured presets, including ones for Netflix, FICAM, and Broadcast, making it easy to produce audio that is optimized toward the chosen delivery platform.

The Advanced Loudness Adaptation presets can be selected - and its parameters easily controlled - through applications such as AudioTools Operator or AudioTools WorkflowCreator, with their graphical user interfaces. Using our REST API, they can also be integrated into third-party tools, such as Telestream Vantage. The Advanced Loudness Adaptation module supports configurations from mono up to immersive layouts, such as 5.1.4 and 7.1.2

Fraunhofer IIS' Dialog+ – Current and Legacy Content

While the Advanced Loudness Adaptation module for AudioTools Server has been highly successful at managing modern challenges, and is consequently used by major networks and content holders worldwide, what about content where loudness isn't the only challenge, or where dialog needs to be treated independently from the other mix components? For this, new technology is now available as an optional module for AudioTools Server: Fraunhofer's Dialog+.

The algorithm, developed by Fraunhofer IIS in Germany, separates the dialog component from background music and effects. This enables the automatic creation of a new mix with greatly improved dialog intelligibility that could be offered, for example, as an alternate audio channel for the hard-of-hearing or other custom audiences. Dialog+ parameters offer control over how the dialog and the bed tracks are mixed, including dialog gain or background audio attenuation levels and timing. The background audio attenuation can be static or dynamic. The module comes with factory presets, but users can also create their own library of parameters using tools such as AudioTools Operator or AudioTools WorkflowCreator.

Fraunhofer's Dialog+ – NextGen TV and MPEG-H

Challenging content is not the only use case for Dialog+. Another option is to use the separated speech and background elements in an object-based MPEG-H workflow. The dialog audio is included as an audio object, joining the main background audio to create a program wherein the viewer at home can interactively control the dialog volume using the remote control. That way, Dialog+ can help to upgrade today's legacy recordings for NextGen TV formats.

Fraunhofer Dialog+, as well as MPEG-H Authoring, are available as modules for AudioTools Server.

Advanced Loudness Presets:

BROADCAST

The “Broadcast” preset optimizes the audio to accomplish the most comfortable television listening experience while fulfilling the highest audio quality expectations. The workflow includes a range of processing strengths to suit all tastes and provides fully independent control of all the most common loudness control parameters. It is designed to meet all international loudness recommendations, such as ATSC A/85, EBU R128, EBU R128 s4, FreeTV OP59, ARIB TRB32, Portaria 354, FICAM, and more.

NETFLIX

The Netflix preset is designed to ensure suitability for Netflix distribution, in accordance with the latest Netflix Best Practices. Overall loudness targets are fixed in accordance with the Netflix delivery specification, including automatic switching of the anchor type based on the amount of dialog present in the source audio. Additional options are provided for targeting LRA, Dialog LRA and the difference between Dialog and Non-Dialog content.

STREAMING

The Streaming preset provides a highly flexible and effective solution for translating theatrical or broadcast content into programs suitable for streaming and mobile presentation, where the overall level must be consistent and loud enough in order to overcome environmental background noise. It has been designed to guarantee compliance with “AES TD1008.1.21-9: Recommendations for Loudness of Internet Audio Streaming and On-Demand Distribution”.

IN-FLIGHT

The In-Flight preset is designed to optimally adapt the audio to fit comfortably into the limited dynamic and frequency range of inflight playback devices. In-Flight Adaptation processing yields higher dialog intelligibility, consistent overall level and comfortable loudness ranges. The mix balance and creative intent of the original mix is maintained and unaltered during the adaptation process, even when downscaling formats are required. The In-Flight Adaptation profile is fully compliant with the APEX 0415 standard and can be further customized to any use case request.

FICAM

The FICAM preset addresses compliance with the FICAM CST-RT-017-TV recommendation. This profile is the most powerful workflow for performing automated adaptation of any content intended to be broadcast according to the French standard. Options tailored for both the short form and long form versions of the specification are available.

CUSTOM

The Custom preset provides full user control of the available Advanced Loudness settings, allowing you to shape the adaptation process to fulfill any custom requirement. Using our state-of-the-art Linear Acoustic APTO™ technology the module allows you to control advanced audio settings, such as the relationship between program and dialog loudness, dialog loudness range and short term dialog, as well as providing full independent control of all the common loudness control parameters.

Ready to Save Time and Money By Automating Your Media Workflows?

Contact Telos Alliance

by visiting TelosAlliance.com/Contact-Sales today to learn more about how AudioTools Server can save you time, money, and human resources.