Audio for Mobile TV, iPad and iPod

Thomas Lund
TC Electronic A/S
Risskov, Denmark

Abstract - *For five years, the author has systematically studied audio capabilities of Pod and Mobile TV devices from Apple, Nokia, Samsung and Sony Ericsson. This paper is the first public report from parts of the test investigating what a Mobile user is able to hear, and what she can't. Taking test results and perceptual criteria into account, guidelines are given for optimum station handling of programs for Mobile devices. Furthermore, the paper presents a transparent and codec-agnostic audio path from HDTV to multiple personal platforms, attaining the goal without a need for "sausage processing". The techniques described aim at high audio quality, based entirely on open standards and a low station workload.*

INTRODUCTION

Among audio lovers, digital has acquired a bad name for its massive use of lossy data reduction, and for its prolific loudness wars that have caused numerous casualties in music, broadcast and film.

Most of our music heritage from the past 20 years is in bad shape, and sounds even worse when played on a fine reproduction system [1-4]. Unfortunately, the music can be considered lost because neither session recordings, nor a non-squashed, linear master is obtainable. Let's just not hope the period had a Beatles, a Kathleen Ferrier, a Dylan, or a Pink Floyd to offer; but that we won't know without extra hindsight years from now.

Another field where hyper-compression is senselessly applied is in broadcast commercials. They are, however, short-lived, and there will be no regrets not being able to hear them in the future.

More worrying, feature films have also gone into a peak level managed death spiral. Despite well standardized reproduction systems, thanks to SMPTE, Dolby and THX, playback gain is now being systematically decreased in film mixing as well as in theaters, thereby voiding the immense benefits of calibrated listening. A recent survey of cinema playback level in Denmark places the average at "4.75" on the arbitrary Dolby scale. Reducing the gain by 8-10 dB on the average is a sad sign of the times.

At the root of the problem lies peak level measurement. Rookies as well as skilled operators are misled by an instrument that should *only* be used to avoid clipping. Even worse, the cheapest way to implement meters in digital audio is based on sample peak detection, so that's what engineers and editors are still looking at in their ProTools, MediaComposer, Logic, Final Cut, Premiere etc. Note how sample peak meters are notorious for not even showing reliably if a signal will cause overload [1-3].

In all areas, audio production must break lose from peak level measurement and from peak level normalization. Also, studios should cover sample peak meters with gaffer tape.

This paper is about getting well sounding audio safely distributed to mobile platforms, and not about producing for lowest common denominator requirements. What makes predictable transmission possible at all is a new broadcast standard where loudness and not peak level takes center stage. The worldwide cornerstone is ITU-R BS.1770-3 which works across genres, across platforms, and regardless if linear audio or a wide range of lossy data reduction codecs is employed in parts of the signal-chain.

THE STANDARD

Since ITU-R BS.1770 was revised in 2011, it has included an all-important measurement gate, allowing for reliable discrimination between foreground sound and background sound. Normalizing broadcast audio based on its foreground (mezzoforte) loudness level gives significant benefits:

- Optimal measure for use across genres.
- One Target Level => transparent from production onwards.
- One Target Level => simple, works well without metadata.
- Application friendly measure: Automatic start and stop.
- Open standard. No patents to muddy the waters.
- More headroom than a speech based measure.

Details about the five first topics may be found in [5, 6]. The standard's latest revision at the time of writing is BS. 1770-3, which is now the world reference [7].

Peak to Loudness Ratio, PLR

A transparent definition of Program Loudness isn't the only virtue of BS.1770. The standard also specifies a technique to stay clear of overload, namely to observe "true-peak" level, a superior technique compared to sample peak.

Programs and music tracks may consequently be evaluated by their Peak to Loudness Ratio, or actually true-peak to loudness, abbreviated "PLR". This is a measure of how demanding on headroom a program will be for the downstream signal-path.

Modern pop/rock music and commercials generally have the lowest PLR values, which is a sign of extensive use of compression and limiting in the production process. A recent study compares PLR over time of the most popular music tracks in US, UK and Germany, see *Fig 1*. It reveals a high point with the introduction of CD in the mid 80'ies, and a significant decline ever since [8]. Some tracks today have a PLR of less than 8 dB. At the opposite side of the scale, feature films and classical music have the highest PLR values, sometimes over 20 dB.

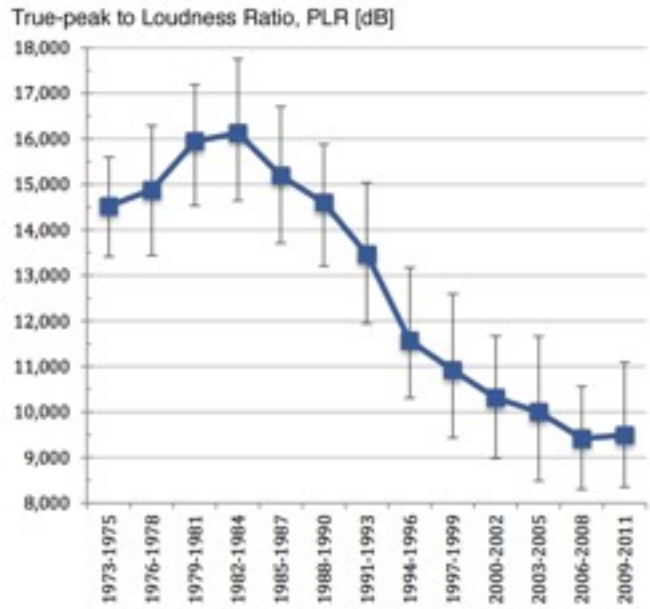

FIG 1. PLR FOR THE 7488 MOST POPULAR MUSIC TRACKS IN US, UK AND GERMANY, 1973 THROUGH 2011

Headroom

For a certain signal-path, the ratio between the maximum peak level it handles and the normal RMS operating level is called headroom. Any part of a signal-path constitutes a possible headroom bottleneck, and the entire chain is limited by its weakest link.

The normalization method used at a station has a serious effect on the amount of headroom available, and therefore on how well high PLR programs may be passed without clipping or processing. The phenomenon is illustrated in *Fig 2*, which is a typical example [6, 9].

The bar marked "1" shows a program with a Loudness Range of 20 LU normalized using speech anchoring, as in ATSC guidelines [10]. A side-effect of this cinema approach is a pronounced loss of headroom in broadcast, with clipping or limiting as a result. The same program is shown on bar no. 2, but this time normalized using the old BS.1770 measure without gating. Much of the headroom is still eaten by Loudness Range, i.e. parts of the program louder than its normalization point. Finally, the program has also been normalized using BS.1770-3, where less of the Loudness Range has vanished. Bar no. 3 shows how the same program escapes processing just by utilizing a more intelligent normalization scheme [5, 6, 7, 9].

"Headroom" is therefore used about a platform's ratio between max True-peak level and Target level based on BS. 1770-3. The latter refers to the default Program Loudness used, i.e. -24 LUFS in most of the world, where headroom in DTV thus is 22 dB (-2 dBTP relative to -24 LUFS).

Note: This paper uses the ISO compliant unit "LUFS" rather than "LKFS". However, the two are identical, so -24 LUFS is exactly the same as -24 LKFS. Some people have taken the "U" vs. "K" as an indication of measurement gating or not, but that is a misunderstanding. There is *only one* BS.1770 standard, currently the -3, and it employs a relative gate at -10 LU.

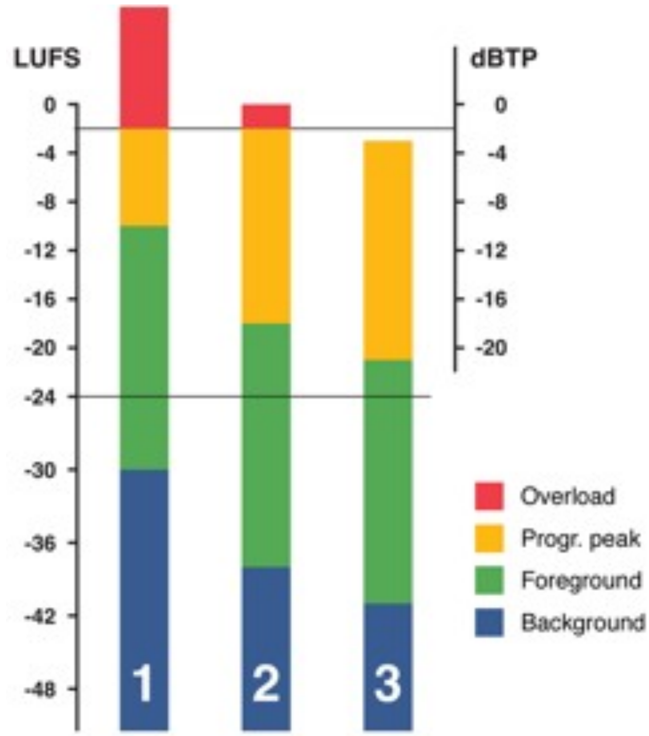

FIG 2. HEADROOM VS. NORMALIZATION METHOD
1: SPEECH, 2: BS.1770-1, 3: BS.1770-3

Because cross-genre balancing and headroom is better with BS.1770-3 normalization than with any alternative, it was a disappointment how ITU chose to not dispose of ambiguous and patent restricted speech normalization in the recently revised ITU-R BS.1864 concerning international program exchange. Instead, the two methods, BS.1770-3 norm and Speech norm, are both recognized. A crying shame for transparency and therefore for audio at large.

PERCEPTION AND TRANSIENTS

Sensation can be divided into reception and perception. Sound is mechanical energy transmitted through a medium, typically air, which has to fall within a certain frequency interval for the ear to recognize a stimulus. Our eyes also only detect a small frequency range of the electro magnetic waves that come our way, and the sharp spot (called macula) covers just roughly the size of the moon on the sky.

Nevertheless, the bandwidth of our senses is much higher than the bandwidth of our consciousness. Therefore we constantly prioritize between hearing, seeing, touch, taste, smell and various somatic assessments to bring down the total to a mere 30 to 40 bits per sec, see *Fig 6* [11].

Physiologically speaking, senses arrive at the brainstem, which is the center of prioritization and of cross-modal correlation. The most temporal acute processing we have is performed in this region, namely L/R ear comparison for localization. It's important to realize how sensation takes time, and changing priority from one sense to another takes even longer. 400-500 ms to be precise, see *Fig 3*. The illustration shows Libet's groundbreaking findings that were doubted for years [12, 13]. Note how startle associated with some senses compensates for latency.

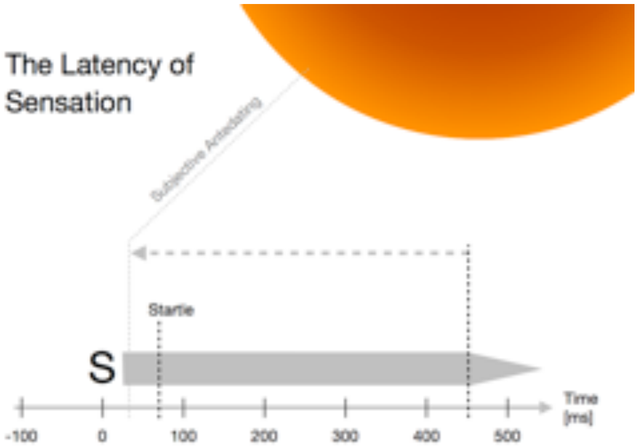

FIG 3. SENSATION TAKES TIME AND USES ANTEDATING
S: STIMULUS. ORANGE CIRCLE: CONSCIOUSNESS

Fig 3 also justifies why the Momentary Loudness measure of ITU-R [19] and EBU [5] at 400 ms is relevant (see *Fig 5*), and why it's not meaningful to define a shorter time interval for measuring *loudness*.

Fig 4 shows the perceptual bottleneck of the brainstem. The primary auditory path is fast without many synapses, so we react to acoustic startle quickly (70-100 ms). Hearing is the undisputed king of temporal sensing, and by having startle associated, we're able to react quickly to threats when newborn, or if we want to win a 100 m sprint.

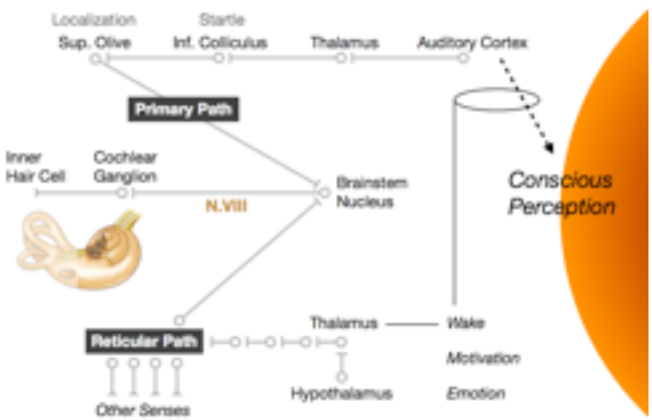

FIG 4. AUDITORY PRIMARY AND RETICULAR PATHS
N.VIII: AUDITORY CRANIAL NERVE

However the reticular path with its final "switchboard", the Thalamus, decides which sense we're actually conscious of, and that takes far longer (*Fig 3*). This is a reason why the combination of cell phone and driving is a lethal cocktail.

We obviously hear sounds shorter than 400 ms, but such sounds should not be confused with the general sense of loudness all humans share across borders. Phonemes, the building blocks of language, are much shorter, but unlike loudness they take training to learn [14]. Intelligibility, clarity and loudness are not the same.

Transients play an important role for the two former in speech and in music. Transients are even more important when a speaker is under quiet conditions and the listener under noisy. In such cases the Lombard reflex doesn't kick in to make the speech clearer and less transient dependent [15, 16]. In those cases, restricting PLR through transient limiting can be bad for speech clarity and for intelligibility.

Similarly, peak limiting of music tends to offset the balance between direct and reverberant sound in favor of the latter, which again may be bad for clarity. For music, however, the story has a twist. Like phonemes in language, we need to learn music transients before we can appreciate them at all. Either we have to be familiar with real, acoustic instruments, or we must have learned their sound aided by a decent reproduction system.

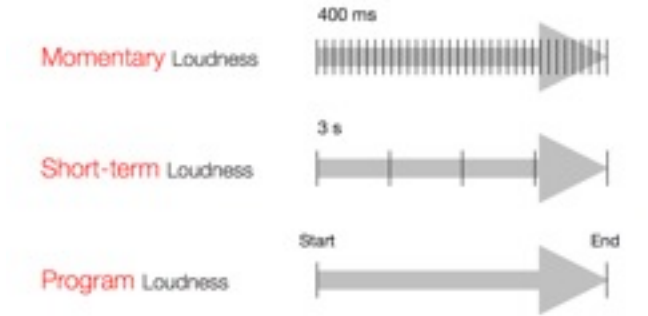

FIG 5. LOUDNESS ON THREE TIME-SCALES
PER ITU-R BS.1770, BS.1771 AND EBU R128

The declining PLR in pop music is not good for younger generations. Even if a kid buys fine speakers, she doesn't stand a chance of finding out how transients sound because they are missing from the source. Consequently, many of them do not hear transients. This has been verified on young employees at TC. With good speakers, most couldn't tell a difference between a music track, and the same track with a PLR of 6, 9 or even 12 dB lower. It's like a language they've never learned.

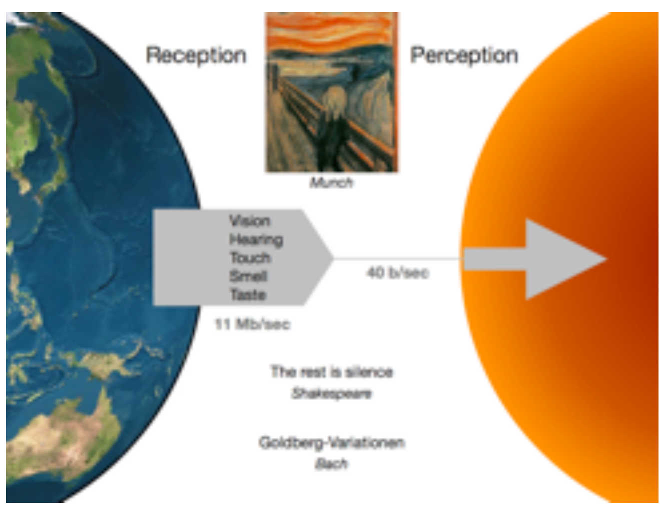

FIG 6. BANDWIDTH OF CONSCIOUSNESS
ART ALLOWS US TO BREAK FREE OF A 40 BIT/SEC REALITY

Just because kids haven't learned to appreciate transients, that's not a reason for chopping them off. Shakespeare isn't made a cartoon because some can't read. Art is our chance of experiencing more than 40 bit/sec when stimuli *develop* in our mind, see *Fig 6*. As professionals, we have an obligation to ensure how art based on audio retains this potential.

Loudness Range, LRA

Unlike PLR, anyone is able to hear Loudness Range, which is a statistical measure of loudness variation inside a track or a program [9, 17].

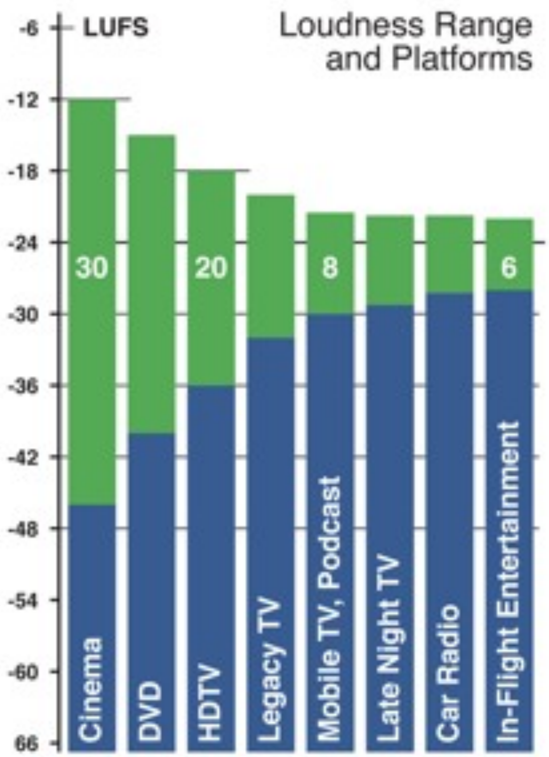

FIG 7. SUGGESTED MAX LOUDNESS RANGE AND TARGET PLATFORMS
LRA FOR MOBILE TV SHOULD NOT BE MUCH HIGHER THAN 8 LU

Loudness Range isn't symmetrical around the Target. For instance, typical broadcast content with an LRA of 8 LU, normalized to -24 LUFS, will likely have most of the Loudness Range come from sources softer than -24 LUFS, as seen in *Fig 7*. Essential parts of the program could have a short-term loudness level around -30 LUFS, which - as seen in *Tab 2* - is expected to generate an SPL of around 72 dB (78.6-6 dB with pink noise) on an iPhone 5 with gain turned up full.

Fig 7 provides a Loudness Range to aim for when mixing regular programs with a specific platform in mind. The measure has proven useful in production by helping to settle expectations early, which is its main purpose in the EBU guidelines [5]. Even classical music or drama for HDTV should not have an LRA in excess of 20 LU, but there is no specific requirement.

Contrary to LRA, a high PLR is generally not a problem for a pod listener. A program with a PLR of, for instance, 15 dB is fine, as long as its Loudness Range isn't much higher than 7 or 8 LU. Try and listen to Donald Fagen's "New Frontier". It has a high PLR (18 dB), but with an LRA of 6 LU it needs no processing to sound great on iPhone - or on a superb set of speakers [4].

APPLE ITUNES STUDY

At the AES convention 2009 in New York, the author took part in a panel on the music loudness wars, contributing with a study on the normalization function embedded in iTunes, so-called Sound Check [18]. Data showed how Sound Check overall is a benign feature, able to gain offset tracks in a playlist based on loudness. Old and new tracks can live side by side without adjustment of the gain, though normalization isn't based on BS.1770, but on an Apple algorithm.

By and large, Sound Check was found not to be far off, most tracks sitting +- 2LU from where BS.1770-3 would have put them. This is an immense improvement compared to deactivating Sound Check, where inter-track loudness can deviate +- 10 LU or more.

The *median Target* level using Sound Check was found to be -16.2 LUFS on a BS.1770-3 scale; very reasonable taking *Fig 1* into account. With the company's attention to audio detail, this is presumably the level Apple device's gain structure is optimized for.

It was also shown how positive normalization in general is disabled in case peak level would go above 0 dBFS, so tracks with a PLR of more than 16 dB get normalized to play quieter than -16.2 LUFS.

One of the reasons why this paper was written in the first place also relates to iTunes: When traveling, the author often listens to BBC Radio 4 podcasts, in particular "In Our Time" where history, culture and science is discussed in a stimulating way. One particular program about Benjamin Franklin I couldn't turn up loud enough. Parts of it drowned in background noise on the flight.

Measuring it later, the combination of a relatively high LRA and a soft Program Loudness, see *Fig 8*, was part of the problem. However, Sound Check had been able to boost the podcast by 7 dB (from -23.3 to -16.2 LUFS) if its PLR had just not been so high. The program was simply stuck at low level and not suitable for flight.

FIG 8. LOUDNESS PLOT OF BBC PODCAST THAT TRIGGERED THIS PAPER.
PART OF THE PROGRAM IS TOO QUIET FOR AN IPOD

MOBILE TV AND IPOD TESTS

From an audio point of view, mobile devices operate under less than ideal conditions: Physical size limits the amount of voltage and current (i.e. power) they are able to feed a pair of headphones, the listening environment is often noisy, and the package is highly price sensitive.

We tested iPods, iPads and Smartphones on a number of audio parameters. The equipment used was Otto, a head and torso with fine, built-in condenser microphones, nearfield monitors, main monitors, calibrated SPL meter and loudness software. The setup is shown in *Fig 9*.

FIG 9. OTTO IN THE STUDIO LISTENING TO APPLE EARPODS

Otto was calibrated using pink noise at -23.0 LUFS to the nearfield monitors placed 240 cm from the torso. This resulted in an SPL of 77.0 dB slow C per channel, 80.0 dB for both, at Otto's position. The binaural microphones fed an analyzer using integrated Leq with power summing of the channels like BS.1770. The 80 dB SPL point was used as reference when subsequently testing mobile devices and headphones, and calibration was repeated twice per day.

FIG 10. YOUTUBE VIDEO WITH PINK NOISE AT -16.0 LUFS

Mobile devices were tested with a number of signals ranging from speech and music over pink noise and tones. For each test signal, its True-peak and Program Loudness value could be read from the DUT's display while it was playing it, see *Fig 10*. Test videos were uploaded as YouTube clips with AAC data reduced audio. For Apple units, the same tests were repeated using QuickTime files with linear audio. For the SPL study reported here, there was no significant difference between lossy YouTube and linear QuickTime.

The main objective was to predict the SPL for a given Program Loudness one could get from a personal platform when listening in headphones. In the real world, on several occasions, the author had run out of playback gain on his iPod when listening to podcasts.

TEST RESULTS

Each mobile device was tested with its replay gain at max, thereby generating a measurable signal into Otto's ears via a pair of headphones. This provided SPL data when a certain Program Loudness level was reproduced through a certain pair of headphones. DUTs were tested using AKG K240S headphones (55 ohm) as a common reference, and also with the set of phones that originally came with that particular unit, in the table marked "Standard".

The AKGs were chosen for their sound, and as industry standard, semi open types with an impedance that wouldn't make outputs current-limit.

Personal Platform	Phones	PN -24	PN -16
Apple iPod Nano G2	AKG K240S	75.9	83.9
Apple iPod Nano G3	AKG K240S	76.1	84.0
Apple iPad	AKG K240S	77.4	85.3
Apple iPhone 5	AKG K240S	71.8	79.6
Nokia Lumia 920	AKG K240S	66.3	74.3
Sony Erics Xperia	AKG K240S	70.2	78.1
Samsung Gal S3	AKG K240S	69.2	77.2
Samsung Gal IIS	AKG K240S	69.9	77.8

TAB 1. DB SPL WHEN REPRODUCING PINK NOISE AT -24 LUFS OR -16 LUFS THROUGH REFERENCE HEADPHONES

Personal Platform	Phones	PN -24	PN -16
Apple iPod Nano G2	Apple Old	86.0	94.0
Apple iPod Nano G2	Apple New	86.8	94.8
Apple iPod Nano G3	Apple Old	86.4	94.4
Apple iPod Nano G3	Apple New	82.8	90.7
Apple iPad	Apple Old	86.0	93.9
Apple iPad	Apple New	87.0	95.0
Apple iPhone 5	Apple New	78.6	86.6
Nokia Lumia 920	Standard	86.2	94.2
Samsung Gal S3	Standard	81.8	89.8
Samsung Gal IIS	Standard	82.5	90.5

TAB. 2. SAME AS TAB. 1, BUT DEVICE-BUNDLED HEADPHONES. FOR APPLE, NEW OR OLD HEADPHONE TYPE IS INDICATED

SUBJECTIVE VERIFICATION AND NOTES

To check speech and music based programs under real world conditions, the author also listened to devices using Standard and AKG K240S reference phones. In order to evaluate normalization level, programs were accompanied by a meter display as shown in *Fig 10.*

Based on a variety of programs, only one set of the in-ear phones tested provided a decent spectral response and good imaging, namely Apple's new standard "Earpods". The in-ear types generally block surrounding noise even less than semi-open AKGs.

With new Apple Earpods, listening to pop/rock music in a car or a train, it took an SPL of 80-83 dB at the minimum for a good experience. When listening to clear speech under the same conditions, words can be expected to get lost below approximately 78 dB SPL. In the following, below 78 dB SPL will be considered surely not enough.

Though not the scope of this study, hearing loss as an effect of Pod over-dosing should be taken seriously: Otto also listened to much hyped "Beats by dr. dre" headphones. Driven by an Apple device, SPL exceeded 100 dB at -16 LUFS!

INTERPRETATION OF RESULTS

Tab 1 and *Tab 2* gives an estimate of the SPL a mobile user would experience when fully turning up the gain ("volume"), playing audio at -24 LUFS or -16 LUFS.

With the reference headphones and either target level, playback SPL falls within a range of 11 dB. In general, Apple devices are better at driving the AKG reference headphones than other vendors tested here (*Tab 1*). With programs at -16 LUFS, the SPL is high enough to hear music and clear speech under most conditions, while devices other than Apple were too soft.

Using the vendors' headphones, devices were generally loud enough when audio played at -24 LUFS. Apart from Apple headphones, however, the bundled ones sound so bad that users would likely buy a different pair, thereby adding to the variation.

Older Apple iPods and iPads play roughly 1 dB louder with the new headphone type than the old one. This is in contrast to new iPhones and iPods that play softer when the new Apple Earbuds are used *(Tab 2)*. The reason for the difference could be technical, but it's likely rather a sign of Apple getting SPL tighter under control.

MOBILE TV AND PODCAST GUIDELINES

A station should think carefully about immediate and future requirements when deciding on the best overall strategy for the handling of Mobile and Podcast. The procedure needs to be automatic, transparent, well sounding and flexible.

Automatic

It's a waste of valuable time to prepare content for more than one platform, namely HD. Transcoding to Mobile TV with a Target level between -18 LUFS and -14 LUFS must happen automatically. *Fig 11* uses -16 LUFS as the goal.

Transparent

From production onwards, it should be easy to check how a program will sound on any given platform, so there should be nothing ambiguous about the transcoding; for instance if a program is normalized to Speech Level or to its Program Loudness, or if metadata values are right or wrong.

Well sounding

To optimize audio quality, Target level should not be raised more than needed. -16 LUFS for Mobile TV is a reasonable choice. Some programs need restriction in LRA and PLR, but "sausage processing" must be avoided so distinction between foreground and background sounds isn't washed out, even when delivered to Mobile TV.

Flexible

Nobody knows the requirement of tomorrow's listener for sure, but we live in a dynamic world. Be careful not to get locked to a certain data reduction codec, because Ogg Vorbis or lossless coding might be your best choice tomorrow.

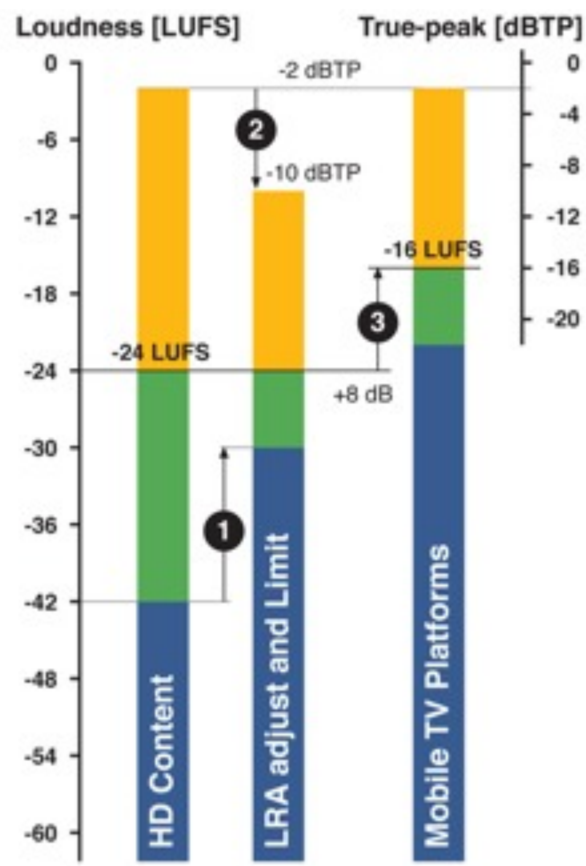

FIG 11. AUTOMATIC TRICKLE-DOWN FROM HD TO MOBILE AND POD THREE EASY STEPS WHEN PROGRAMS SIT AT -24 LUFS

Adding it all up, ideal and easy cross-platform encoding is shown in *Fig 11:* Use BS.1770-3 to normalize content to -24 LUFS. Follow the numbers on the illustration to -

1) Bring up low level sounds that would drown on Mobile.
2) Limit peaks to -10 dBTP.
3) Add a static gain offset of, for instance, +8 dB.

CONCLUSION

A case has been made for not pursuing a lowest common denominator approach to audio in Pod and Mobile TV. For consumers with flat panel TVs and matchbox sized loudspeakers, personal platforms and headphones is the closest they get to a decent audio experience these days. However, two things could prevent even that from remaining possible: Hyper-compression at the source, and/or more lossy data reduction on the platforms consumers listen to.

In this paper, a study on the SPL from various iPods and Mobile TVs has been presented, showing a spread among vendors of around 11 dB. Different headphones taken into account, the variation is over 20 dB. SPL in Mobile TV has also been linked to the best method of normalizing audio in broadcast, the Program Loudness measurement of ITU-R BS.1770-3, so level is well defined at transmission at least.

Using standard headphones, devices from Apple were not only found to be the best sounding, they also offered the highest gain. Headphone gain is a crucial element of sound quality in order for programs and music tracks not having to be hyper-compressed before transmission. If all Mobile TVs and Pods were from Apple, the need for a higher Target level than -24 LUFS in Mobile broadcast wouldn't be strong.

With other less capable systems to take into account, however, it has been justified how a Program Loudness of -16 LUFS is an informed choice still allowing for high quality playback without much dynamics processing. An automatic, simple and transparent method of taking audio from HD specs at -24 LUFS to Mobile specs at -16 LUFS was therefore described, based entirely on open standards.

Program Loudness and Loudness Range are important parameters when preparing programs for Mobile TV while peak level is less important. However, a too low platform headroom has an adverse effect on sound quality. True-peak level for Mobile platforms doesn't have to be restricted to -2 dBTP. Lossy codecs need a conservative peak limit only if the measurement is sample peak, or if level sits around full scale all of the time. That's not the case when Target is set to -16 LUFS. Consequently, the best sound with the most headroom for transients comes from Mobile transmission allowing True-peak level all the way to 0 dBTP, or at least to -0.5 dBTP. Part of the low level peak myth stems from AC3's lack of headroom for downmix, but that shouldn't hurt the sound of Mobile TV. Transient headroom also helps clarity and speech intelligibility.

Preventing adequate output gain on personal platforms is detrimental because it forces source audio to be squashed to make a program be heard. The European tech standards committee, CENELEC, unfortunately hasn't been a help by putting restrictions on the allowable amount of gain. That is not an intelligent concept for reducing SPL in Pods.

The responsible way forward, reinvigorating audio as a possible carrier of art, must use alternatives to CENELEC type of "solutions"; and phase out the dependency on lossy data reduction. Regarding the former, manufacturers reading this are encouraged to consider an integrated normalization and gain control [20] for the next generation of Mobile TVs, iPods, iPads and other consumer devices.

REFERENCES

[1] Nielsen, S.H. & Lund, T., "Level Control in Digital Mastering", *Paper of 107th AES Conv.*, New York, NY, 1999.

[2] Nielsen, S.H. & Lund, T., "Overload in Signal Conversion", *Paper #11 of 23rd International AES Conference*, Copenhagen, Denmark, May 2003.

[3] Lund, T., "Stop Counting Samples", *Paper of 121st AES Convention*, San Francisco, CA, 2006.

[4] Serinus, J.V., "Winning the Loudness Wars", *Stereophile Magazine*, New York, NY, Nov 2012.

[5] EBU, "EBU Technical Recommendation R128 - Loudness normalisation and permitted maximum level of audio signals", *European Broadcasting Union*, 2010.

[6] Lund, T., "The CALM Act and Cross-platform Broadcast", *Paper NAB BE Conf.*, Las Vegas, NV, 2012.

[7] ITU-R, "Rec. ITU-R BS.1770-3, Algorithms to measure audio programme loudness and true-peak audio level", *International Telecommunications Union*, 2012.

[8] Ortner, R., "Je lauter desto bumm! - The Evolution of Loud", *Donau Universität*, Krems, Austria, 2012.

[9] Skovenborg, E. & Lund, T., "Loudness Descriptors to Characterize Wide Loudness Range Material", *Paper of 127th AES Convention*, New York, NY, 2009.

[10] ATSC, "Techniques for Establishing and Maintaining Audio Loudness for Digital Television", *Advanced Television Systems Committee.* Doc A/85, 2011.

[11] Küpfmüller, K., "Nachrichtenverarbeitung im Menschen", *Springer*. Berlin, Germany, 1962.

[12] Libet, B., "Subjective Referral of the Timing for a Conscious Sensory Experience", *Brain*, 102, Oxford, England, 1977.

[13] Libet, B., "The Timing of Mental Events: Libet's Experimental Findings and Their Implications", *Consciousness and Cognition,* 11, 2002.

[14] Tremblay, K. et al., "Central auditory system plasticity: Generalization to novel stimuli following listening training", *Journal of Acoustical Society of America*, 102(6), 1997.

[15] Summers, W.V. et al., "Effects of noise on speech production: Acoustic and perceptual analyses", *Journal of Acoustical Society of America*, 84(3), 1988.

[16] Lau, P., "The Lombard Effect as a Communicative Phenomenon", *UC Berkeley Report*, CA, 2008.

[17] Skovenborg, E., "Loudness Range (LRA) – Design and Evaluation", *Paper of 132nd AES Convention*. Budapest, Hungary, 2012.

[18] Lund, T. et al., "Loudness Wars", *Data from "tribunals".* AES Conventions 127, 129, 131, 133. 2009-2012.

[19] ITU-R, "Rec. ITU-R BS.1771-1, Requirements for loudness and true-peak indicating meters", *International Telecommunications Union*, 2012.

[20] Camerer, F. et al., "Loudness Normalization: The Future of File-Based Playback", *Paper for the audio industry.* Vienna, Austria, 2012. http://www.music-loudness.com