Audio Engineering Society

Convention Paper

Presented at the 115th Convention
2003 October 10–13 New York, New York

Intelligent Program Loudness Measurement and Control: What Satisfies Listeners?

Jeffrey C. Riedmiller, Steve Lyman, and Charles Robinson

Dolby Laboratories Inc. San Francisco, CA. 94103, USA

ABSTRACT

The broadcast, satellite and cable television industries have been plagued for years by the inability of personnel to accurately interpret and thus consistently control program loudness utilizing traditional measurement devices and methods. As a result, most listeners feel compelled to make adjustments to their television volume controls (in the home). A recent survey of channel-to-channel and/or program-to-program level discrepancies and subjective listening tests confirms that current the practice is unacceptable to listeners.

In this paper we describe loudness measurement techniques that improve accuracy, usability, and consistency relative to previous techniques. Accuracy in this application is determined by correlation to listener opinion, with the particular goal of minimizing annoyance resulting from level mismatch. Usability is improved by minimizing the interaction required by the user. Consistency is achieved by minimizing the amount of meter interpretation required. The keys to this method are: providing a single numeric indication of loudness for a given program or segment; and isolating and measuring the portions of the program that are primarily speech, and using speech loudness as the basis for overall program level thereby improving listener satisfaction.

1. Introduction

The need for loudness measurement for program level matching is an active and growing concern for people in the audio content production and distribution industries. With the increase in content distribution mechanisms (cable, satellite, and terrestrial television), formats (analog, digital, data compressed), as well as the increase in dynamic range of many of the delivery channels, level mismatch is clearly getting worse. Therefore, it is clear that a reliable and consistent method for measuring program loudness is needed.

2. Current Practices and Problems

To begin, within the past few years we have entered an age whereby many television viewers are presented with programming in both the analog and digital domains (in some cases unbeknownst to them). And this situation has, in many cases, generated audio level discrepancies that far exceed what most of the television viewing population would consider acceptable. This is probably most apparent, although not exclusively, with today's cable television programming since many local cable systems now offer digital programming in addition to all of their existing analog programming.[1] The evidence of this is clearly visible in figures 1 and 2.

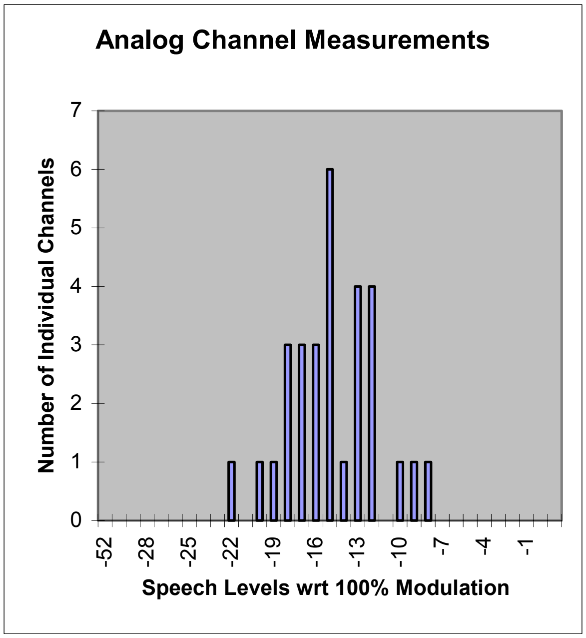

Figure 1 – Analog Cable Levels (30 channels)[2]

Figure 1 represents measurements taken from 30 individual services available on analog cable. Figure 2 represents measurements taken from 13 services available on digital cable, for a total of 43 services combined. Both figures (i.e. Histograms) indicate the average (mean) dialogue levels (Leq(A)) with respect to the maximum permissible level. Whereby, the maximum permissible level is defined as 100% modulation (i.e. 25kHz peak deviation) as per FCC rules. As a side note, the digital services were measured from the channel 3-4 remodulated output of the digital set top box. The 2-channel AC-3 decoder (within the set top) was operating in RF mode and the indicated dialnorm value for each of the digital measurements was –27. This method of assessing the digital service levels, in the home, was based on information provided to us by several cable MSOs (multiple system operators) whereby most of their digital subscriber installations utilized the channel 3-4 remodulated output of the set top.

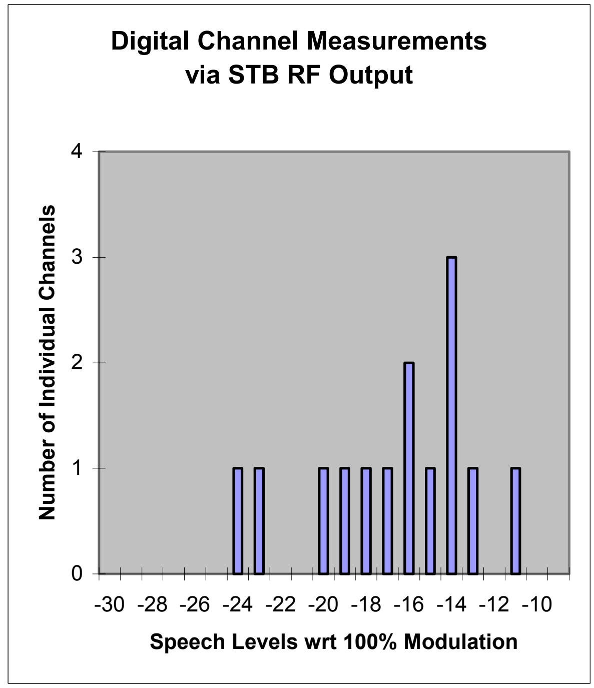

Figure 2 – Digital Cable Levels (13 Channels)[2]

The results indicate that there are significant level differences within and across both tiers of service (i.e. Digital and Analog). When combined, the average A-weighted dialogue levels vary over a 16dB range. This range is well outside the comfort zone (defined in Section 3.3) of listeners tested during our research.

2.1. A Little Known Fact Regarding Digital Set Top Boxes

Many broadcasters, including cable programmers and cable television personnel do not realize that the digital set box has been designed with a several assumptions regarding analog broadcast levels, digital decoder operating mode, the validity of the digital audio metadata and the level relationships within the STB itself. For our discussion, we will focus on three of them. 1. Digital set top boxes assume that while tuned to an analog service (either "off-air" or via "cable") the average dialogue level is ~ -17dB Leq(A) below 100% modulation. 2. While tuned to a digital service (either "off-air or via "cable") the transmitted dialnorm value (in the AC-3 bitstream) is assumed to be correct for that program. 3. While utilizing the channel 3-4 remodulated RF

---

[1] *Within North American Cable Systems*

[2] *Measurements were taken during the spring of 2003 in the greater San Francisco, CA area.*

[4] *The dialogue normalization value in an AC-3 bitstream must be set by the program originator to indicate the dialogue level of the program relative to 0dBFS. The valid range for this value is from –1dBFS to –31dBFS. Also note that the dialnorm value is utilized by the decoder (within the set top box) to normalize the programming to a consistent level.*

output of the set top box, the AC-3 decoder must default to the RF operating mode.[4]

All of the assumptions listed above are referenced in the Open Cable Host Device Core Functional Requirements (i.e. for Open Cable Digital Set Top Boxes) and in a bulletin issued by the Electronic Industries Association (EIA) and the Consumer Electronics Association (CEA) entitled, EIA/CEA-CEB-11 NTSC/ATSC Loudness Matching [3]. This document provides guidance to digital set top box manufacturers on how to maintain uniform audio loudness between existing NTSC programming and digital television services while simultaneously preserving the dynamic range capability of the digital services. The bulletin also addresses the capabilities of consumer broadcast products to match loudness from the listener's perspective, internal gain structure, and output specifications.

Therefore, in order for digital programming to match analog programming (at the channel 3-4 RF output on the set top box) all of the conditions listed above must be met. Now, if we refer back to figure 1 we can see that only three out of the thirty analog services exactly meet our first assumption where the average A-weighted Leq speech level is ~ -17dB below 100%. Referring to figure 2 we measured only one service out of 13 to be at –17dB below 100%. Even though the AC-3 decoder was operating in RF mode as per assumption number three. Another interesting note regarding the digital measurements was that the dialnorm value for each of these services was set to –27. Interestingly, this value is the default value found in most broadcast encoding systems and indicates to us that these programmers are not taking full advantage of the dialogue normalization feature included in the AC-3 system. Therefore, based on the data in figure 2, only one broadcaster had the dialnorm value set correctly at –27 (i.e. when decoded in RF mode and presented to the RF remodulator the speech level is equivalent to –17dB below 100% modulation).

Obviously, we know that services that also measured within some value around –17 are acceptable to listeners as well. We will define these limits of acceptability in section 3.2. Furthermore, during discussions with cable television technical personnel they often complained that some digital services are too quiet. Our response to this question was "How do you know that that your analog channel modulation levels are not too high and that the dialnorm values for the digital services are correct?" As you can also see from the measurement data we collected even if a digital service was provisioned correctly (i.e. dialnorm set correctly, and the STB decoder is operating in the correct mode, etc) and the digital speech level emerged from the RF output of the STB at –17dB below 100% that 21 out of the 30 analog channels would be louder and 6 channels would be quieter. Once this was explained to them they realized the importance of developing methods and acquiring the tools to accurately assess the speech levels within their programming.

In summary, it is clear that the methods currently in use across the broadcast industry to minimize the loudness discrepancies have not made a significant impact on improving the situation. Many but not all of the programs and channels we measured were clearly utilizing some form of dynamic range processing. However, dynamics processing is not sufficient in itself. Dynamics processing can provide self consistent levels, however, listener complaints can still occur when a channel is compared to another channel that may or may not utilize dynamics processing. It is then in the hands of the personnel making the level assessments and/or adjustments, to utilize measurement methods that meet the level requirements of downstream equipment (i.e. such as a set top box) thereby improving listener experience.

3. Listener Satisfaction

There is presumably some loudness range, call it the Comfort Zone, within which a listener will accept loudness changes within and between programs. Assuming further that the non-speech elements of the programs have been appropriately "balanced" around the speech elements, listeners will not be annoyed by the natural changes in loudness that occur during programs if the speech elements fall within their individual Comfort Zone.

To the best of our knowledge, the magnitude of this Comfort Zone has never been determined. A series of subjective experiments was undertaken to determine the range of loudness levels that define the Comfort Zone.

The basic method was to present listeners with a Reference program segment whose loudness they adjusted to their own comfortable listening level. Thy were then asked to adjust the loudness of a test segment until it was louder than the reference item, but still acceptable, an then to adjust the same item until it was softer than the reference item, but still acceptable. These thresholds were taken as defining the listener's Comfort Zone. Listeners were also asked to adjust the loudness of the test item to several other thresholds, both to investigate these points and to allow for the necessary randomization within in the experiment.

Determination of the size or range of the Comfort Zone required that the relative loudness of the test and reference items be known. This required another series of experiments in which listeners were asked to equalize the loudness of various program samples to that of a Reference program sample. Initial experiments revealed that it was much easier for listeners to consistently match the loudness of a variety of test items to some reference items than to others. This spawned another experiment to determine the most appropriate Reference item.

3.1. The Choice of a Reference Item

Six candidate reference items were selected. Three standardized weighted noise signals (ITU-T G.227 (10/68) "Conventional Telephone Signal", "Long-Term Average Speech Spectrum" (LTASS), and ITU-T P.50 (09/99) "Artificial Voices") that have been used as references in other experiments were selected, along with samples of male and female announcers and an advertisement, in case the noise signals proved unsatisfactory

Six separate experiments were conducted, in which each of the sixteen subjects were asked to match the loudness of a selection of nine program samples and the five other candidates to one of the candidate reference items. The reference item that allowed the listeners to make the most consistent evaluations of the loudness of the test items (judged by the largest number of loudness estimates with standard deviation below 1.5 dB) was taken to be the most suitable reference. The Female Interviewer candidate item met this target for eight of the fourteen items; the next best candidate only achieved this for five items. Interestingly, none of the noise signal candidates performed as well as the "program" type candidates. The histograms in figures 3 & 4 clearly show that the spread of equal loudness estimates is much wider for the experiment, which used the P.50 Noise signal as a reference than when the Female Interviewer program sample was used as the reference signal.

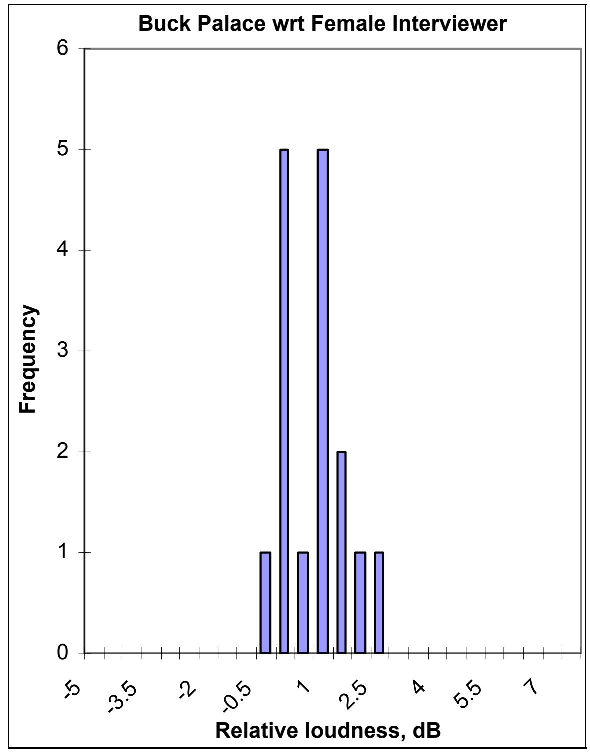

Figure 3

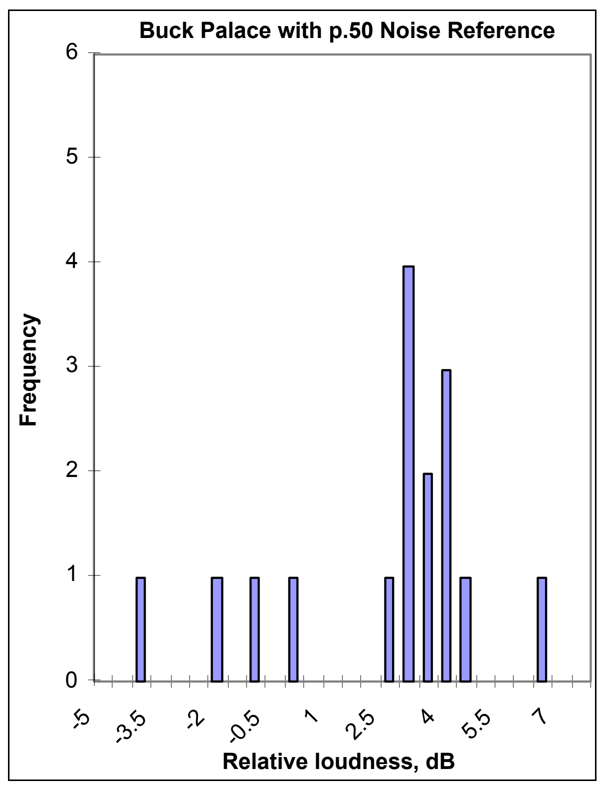

Figure 4

3.2. Loudness Matching Experiment

Once a suitable Reference item had been selected, we could proceed to the second set of subjective experiments in which the relative loudness of a wide variety of different program samples was determined. The experimental procedure was basically the same as was used in the reference selection experiments. Listeners were presented with a randomly ordered sequence of test items, at randomly assigned loudness levels, and asked to use the up and down arrows on a keyboard to make the loudness of the test item the same as the loudness of the reference item. All items were monophonic and played back through a single loudspeaker placed in front of the listener in a quiet room. They could switch seamlessly between reference and test items by either clicking on a screen button or pressing the corresponding key on the keyboard. The same computer that played back the sound samples was used to log the gain changes each listener made, thus avoiding data logging errors. The loudness of each of these test items was expressed as the change in amplitude (in dB) required to make its loudness equal to the loudness of the Reference item.
The loudness matching experiment produced reliable data. The 95% confidence interval for the amplitude change required to equalize the loudness of a typical test item averaged 0.6 dB and was never greater than 0.9 dB. A smaller "sanity check" experiment was done to check the results, and it verified that the data was indeed accurate.

The gain changes resulting from the experiment were then applied to the test items to provide a new library of program material samples of equal subjective loudness.

3.3. Finding the Comfort Zone

A useful objective loudness measurement should be able to put most programs into most listeners' Comfort Zone most of the time. Obviously, it is impossible to do this all the time for all listeners; both the impression of loudness and the comfort zone are individual opinions. It is not always desirable to make all programs equally loud either. A rock concert or some parts of an action film would seem silly if they were not louder than a current affairs discussion. The point of this experiment was to develop criteria for the requirements of a useful loudness meter, and to give some guidance to broadcasters about what the Comfort Zone of their listeners might be as well as to let them know what their listeners might find objectionable.

The experimental scenario was to place listeners in a typical listening or viewing situation during which the program would switch from one type of program to another. Subjects were presented with several paired, monophonic Reference and Test program samples, reproduced by a single loudspeaker in front of them, in a typical listening room.

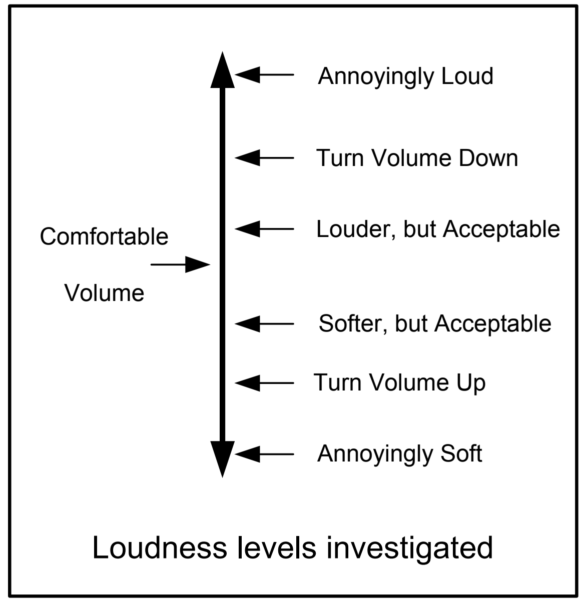

Figure 5

They were instructed to adjust the master playback level until the Reference item was reproduced at what they considered to be a "comfortable volume". The experimenter then asked them to set the Test volume control ("volume" is familiar to most listeners) to one of the six points shown in Figure 5. The questions were asked in random order, and the order of presentation of the pairs of reference and test items was randomized between subjects.
The pairs of program samples were taken from the library of equal subjective loudness samples developed from the loudness matching experiment, so the gain offset the subjects applied to the test item in response to the questions from the experimenter were a direct measurement of the subjects' comfort zone and other critical loudness levels.

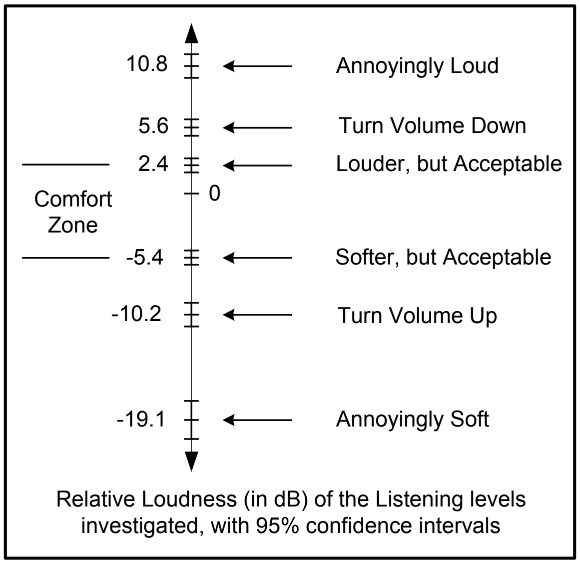

Figure 6

Figure 6 shows the results of the Comfort Zone tests. The interesting aspect is that the results make it quite clear why television broadcasters (and others) have been plagued by complaints of "loud ads" for so long. An increase of two to three dB in subjective loudness is enough to move a program out of the typical listeners Comfort Zone, and toward the point at which they would like to turn the volume down. There is much more latitude available on the softer side of the "comfortable volume" point (shown here as "0"). One point should be mentioned, however. The ambient noise level in the listening room used for the tests was quite low; similar to a suburban living room on a tranquil evening. Since the "Annoyingly Soft" point can reasonably be expected to fall somewhere above the ambient noise level in the listening environment, the figure of –19.1 dB may depend on the ambient noise level. The other points are far enough above the ambient that they should not be affected.

4. Measurement Techniques to Improve Listener Satisfaction

In this section we describe a number of techniques for improving listener satisfaction as described in the previous section.

4.1. Measure Dialogue

In Section 1 we used dialogue levels as the basis for our measurements. During our research into subjective vs. objective loudness estimations it became apparent that listeners could agree with each other more consistently when evaluating content that primarily contains speech. On the other hand, when the listeners evaluated other types of program content such as music and/or effects there was a very large disagreement among them at times. This is shown in figure 7. Figure 7 compares the results of 21 listeners evaluating the level of an audio program containing speech and a program only containing the sound of footsteps (i.e. a sound effect from a drama) compared to a reference. The results show that 19 out of 21 listeners agreed with each other to within 1dB when evaluating a speech item. However, when the same 21 listeners evaluated the footsteps item they disagreed with each other by up to 12dB! Where one listener indicated that the footsteps item was 3dB too quiet and yet another listener indicated that the footsteps item was 9dB too loud. Based on this evidence, we conclude that adjusting the speech portions of programs to a consistent loudness will lead to greater listener satisfaction.

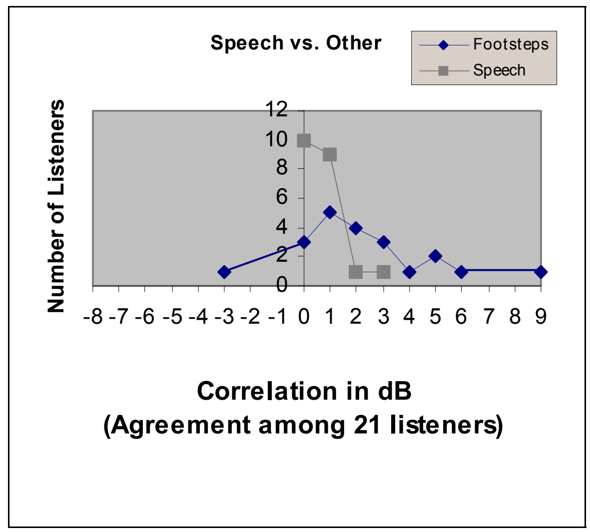

Figure 7– Agreement among listeners when evaluating speech items vs. other signal types (music, effects)

Other studies support the idea that normalizing the speech portions of programs to a consistent loudness would increase listener satisfaction. An unpublished work by Benjamin [1] concluded that television listeners, in a "living room" setting, preferred the dialogue to be set at a mean sound pressure level (SPL) of 60.5 dB (A weighted, slow). Listeners in a home theatre setting preferred a mean dialog SPL setting of approximately 69 dB (A weighted, slow). Pearsons et al [2] found that the SPL of conversational speech, in many different situations, fell in the range from 55 to 66 dB (A weighted, slow). We can thus conclude that listeners most likely adjust their television volume to mimic the levels normally encountered in everyday conversational speech.

Broadcast and postproduction personnel often overlook the importance of speech loudness while program level assessments are being made. Furthermore, many of the measurement methods and devices in use today do not specifically consider dialogue portions of the signal. For example, if one were to take several samples of speech and adjust them so that their peak levels are all about the same on a Peak Program Meter (PPM), they will not differ greatly in perceived loudness. If, on the other hand several other types of programming (e.g. programming that includes a mixture of speech, music & effects, etc) are also adjusted to have the same peak levels and then presented to several listeners, some of these items will sound quieter and some quite a bit louder than the speech itself, depending on the frequency content and on their individual peak-to-rms ratios.

Note when program levels are based on dialogue it is necessary to leave headroom for other types of (i.e. non-dialogue) program content. When broadcasters and programmers make an effort set levels correctly, it is normal to make speech peak to some predetermined meter reading but also to adjust other sound elements (e.g. music & effects, etc) to an artistic level with respect to the speech level without overloading their systems. Therefore, setting the level of speech so that its peaks just reach the system limit will lead to a de facto loudness normalization of speech with the penalty of little or no headroom above the speech peaks. Thus, broadcasters can only provide wide dynamic range and headroom above the normalized speech peaks by ensuring that the speech peaks are several dB below the broadcast system limit.

4.2. Use an Objective Loudness Measure

At the heart of any loudness measurement or loudness matching tool is a method for estimating the absolute or relative loudness of a segment of audio. For years the broadcast industry has quantified the level of their programming utilizing Peak Program Meters (PPM) and Volume Unit (VU) indicators. It is important to note that both are used to read signal voltages, and therefore make no attempt to measure subjective loudness. Thus several different voices, adjusted in level so that they all deflect meters to the same mark, may sound somewhat different in loudness to the listener. PPM and VU meters are also frequently used to measure and/or align to a pre-determined “house” reference level, and thus only have an arbitrary relationship to the dialog or speech loudness within a given program. For example, if a VU meter and a PPM meter are calibrated to display a reference tone equally, and speech that averages 0VU is applied to both, the PPM meter will indicate levels considerably above its reference level and possibly above the maximum permitted level. On the other hand, speech that averages at the PPM reference will most likely indicate many dB below 0VU (our original reference). This confirms the very important idea that the reference level is not the same as the speech or dialog level of a program.

To briefly summarize standard VU characteristics, the VU meter only approximates momentary level changes in program material. Its rapid response does not correspond to the overall impression of the program loudness that we perceive. Further, both the PPM and VU meters also have a flat frequency response over the entire audio spectrum and therefore does not address the fact that the human auditory system is non-linear with respect to frequency, which can result in very large meter deflections that do not highly correlate with a change in perceived loudness. Even to the experienced operator the VU meter is often very difficult to interpret due to its dynamic characteristics and its small useable dynamic range. The useable dynamic range is approximately 13dB whereby the top 6dB of this range is dedicated to 50 percent of the meters overall scale. Thus, with uncompressed material the indicator tends to fluctuate more than the perceived loudness change therefore making this type of device difficult to assess the subjective loudness of broadcast programming among multiple operators. Experienced operators generally use their ears to balance the different elements of a program and to set the overall loudness, while watching the PPM or VU meter to make sure that signal peaks are not overloading the equipment.

As implied by the name, a PPM responds very quickly to changes in the signal level. It was designed to help identify potentially problematic peaks or transients which may exceed distortion limits in a device being used. Because the instantaneous fluctuations in a signal may be so brief as to pass unnoticed, the PPM was designed to react quickly to peak onsets but fall more gradually from these peak levels. The human ear is not particularly sensitive to instantaneous peaks in signal level and while such peaks of short duration may be present in a signal the perceived loudness of the overall signal may not be significantly effected. This makes a peak meter less effective in indicating how loud the signal being passed might sound to the human ear.

Both these meter types were originally developed for measuring signal level in order to best match the signal to audio production, reproduction, or transmission equipment. This is a very different problem than matching audio levels to a desired perceived loudness. None the less, these and other "signal level" meters are often used to estimate subjective loudness or to level programs to reference loudness.

Ad hoc procedures have been developed for interpreting the movements of a meter needle/pointer (or fluctuations of LED's) that do in fact provide a reasonable approximation of subjective loudness. Such procedures generally involve visual averaging of the meter indication and can be quite sophisticated, such as "estimate the average value of the meter considering only the middle two quartiles of the meter values" or "estimate the value at which the meter most frequently falls to within 2dB, disregarding silence". It is difficult to get reliable and repeatable results in this way. Different users may use different procedures. Different people (or even a single person on a different occasion) using the same procedure will likely obtain different results. Procedures can and sometimes are standardized within a studio; this can minimize user to user fluctuation, but can result in systematic differences between facilities. The difficulties associated with meters of this type are magnified by the fact that the meter reading procedure used is often effected - or deliberately altered - based on content type (compressed vs. full dynamic range, speech vs. music).

In the end, however, meters that are open to or require interpretation while useful, still result in a subjective result. In order to estimate perceived loudness, we have substituted one subjective process for another: visual interpretation in place of aural interpretation, and a skilled meter operator in place of a skilled listener. It is clear that an accurate, purely objective measure would be highly desirable. Such a tool or algorithm would provide an explicit loudness value, would be repeatable, and would not require a skilled user. Furthermore, it is possible that such a tool could operate at much greater than real time for rapid analysis of file-based audio. We have analyzed a number of new and existing loudness algorithms and compared the results to subjective listening results. As shown in figure 8, we have found that it is possible to meet or greatly exceed the accuracy of a VU meter with even relatively simple objective algorithms.

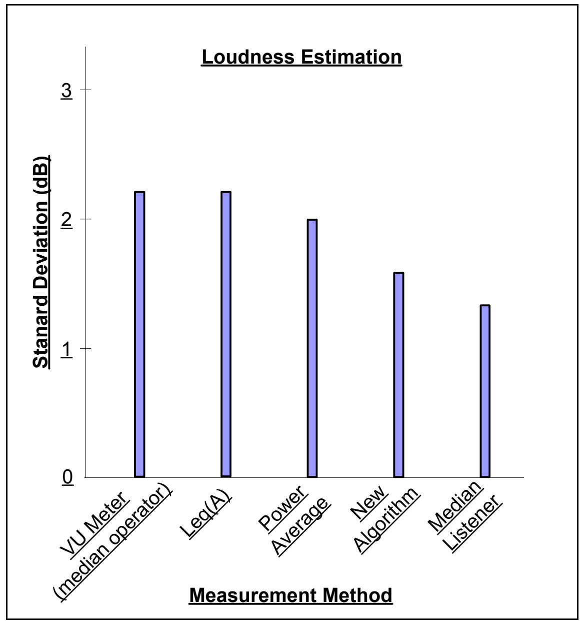

Figure 8

4.3. Provide Multiple Time-Scales

During our investigation it became apparent that in order to satisfy the requirements of estimating loudness in several applications multiple time scales are necessary.

The long-term or “infinite” time scale is necessary for determining the loudness of an entire program that results in a single numeric loudness value. For example, a single loudness value can be utilized for normalization purposes at the program ingest point. A single loudness value can be calculated in real-time while the content is being transferred, or in an offline process on existing files. This value can then be used to normalize the program to a desired level or stored and carried along with the file as a part of metadata where downstream equipment could take advantage of it upon playback. If the single loudness value only considered the loudness of speech over an entire program this value could also be utilized to provision the dialogue normalization value (dialnorm) in an AC-3 emission encoder as well.

There are several applications in which a short-term loudness indication is useful. Level shifts at program boundaries can be more quickly identified with a short-term method. For example, this feature can be used identify and address the loud commercial problem. In addition, measurements performed in short-term mode allow the operator to see short-term variations in loudness within a program. This can be particularly useful in live event situations. In any case, skilled audio operators may prefer to use the short-term measurement in some cases, as they find the information on near-term dynamics to be very useful when mixing or producing a program.

Short-term mode is also very useful for measuring and logging the "dynamic" loudness history of a given program during the QC, postproduction process, or particular television service/channel in a cable head-end facility.

4.4. Provide a Loudness History Log

Periodically logging the short-term speech measurement values against time can be used to identify and correct endemic errors in broadcast. This simple, yet very effective means of analysis can clearly show whether a channel has the proper level, is self consistent and would only require a simple gain change, or has a large loudness variation during or between programs. For example, figures 9 - 16 show the short-term speech loudness history for 8 analog cable channels. The x-axis corresponds to the time of day the measurement was taken, whereas the y-axis corresponds to the Leq(A) speech level (in dB) relative to 100% modulation (i.e. 25kHz peak deviation). To further clarify, 0(dB) on the y-axis is equivalent to 100% modulation. And each data point on the chart represents an individual speech measurement integrated over the previous 60 seconds taken at the time indicated on the x-axis. Also, the dark line at -17dB indicates where the speech loudness should be. Recall from section 2.1 that -17dB was the analog speech value assumed during the design of the digital set top internal gain structure. It is important to note that all of the data we captured utilized an automated system where each channel was measured for 60 seconds. And the channels were grouped into blocks of six which gave us ten 60 second speech measurements per hour per channel. (i.e. one measurement every six minutes) The measurement method utilized was Leq(A) as per IEC 60804.

An analysis of figure 9 indicates that a significant portion of the speech data is approximately at -17dB which is our target level. However, at times could vary significantly above and below our target level. You will also notice two data points at -52dB, this result indicates that speech was not detected during that particular measurement period and our measurement system returns the value of -52dB under these conditions. The data in figure 10 indicates that the speech levels are very consistent throughout this measurement period. To a cable headend technician this data indicates that only a simple gain adjustment (at the analog modulator) of ~ -5dB is all that is required to set the speech at -17dB.

In contrast, figure 14 indicates that this particular channels speech levels varied significantly over time. To the unsuspecting cable headend technician not knowing this information, a simple gain change will not be valid for long. Thus, having a system capable of logging loudness history in this manner can be a very important tool to address loudness complaints. And more importantly determine whether the problem is in house or at the programmer's origination facility.

The data in figures 17 & 18 show the speech loudness history for a single Network DTV channel taken from ~ 5:30pm until ~ 5:30am on February 19th and 20th, 2003. In both of these figures a speech measurement was generated every 30 seconds (note the finer detail). It is clearly evident that the program that aired from 8pm to 9pm was approximately 5 to 6dB louder (on average) when compared to the programming that preceded and followed it. Figure 18 shows that at approximately 2:45am the channel loudness dropped by ~ 7dB to ~ -33dBFS and remained at that level for the remainder of the measurement period.

5. CONCLUSION

The need for better loudness management in content creation and broadcast is well recognized. In this paper we have quantified the accuracy needed to maximize listener satisfaction. We have also described several techniques and tools that can be employed to help meet this objective.

We believe that the best results can be achieved by measuring and leveling the dialogue portions of programming, using an objective loudness measure, utilizing a combination of long and short term loudness averaging. Lastly, the ability to periodically log this type of data is advantageous in identifying and correcting existing loudness discrepancies.

6. REFERENCES

[1] Benjamin, E., "Comparison of Objective Measures of Loudness Using Audio Program Material", presented at the AES113th convention, Los Angeles, California, 2002 October 5-8.

[2] Pearsons, Bennet, Fidell; "Speech Levels in Various Noise Environments", Report No. EPA-600/1-77-025 (1977)

[3] EIA/CEA-CEB-11 NTSC/ATSC Loudness Matching.

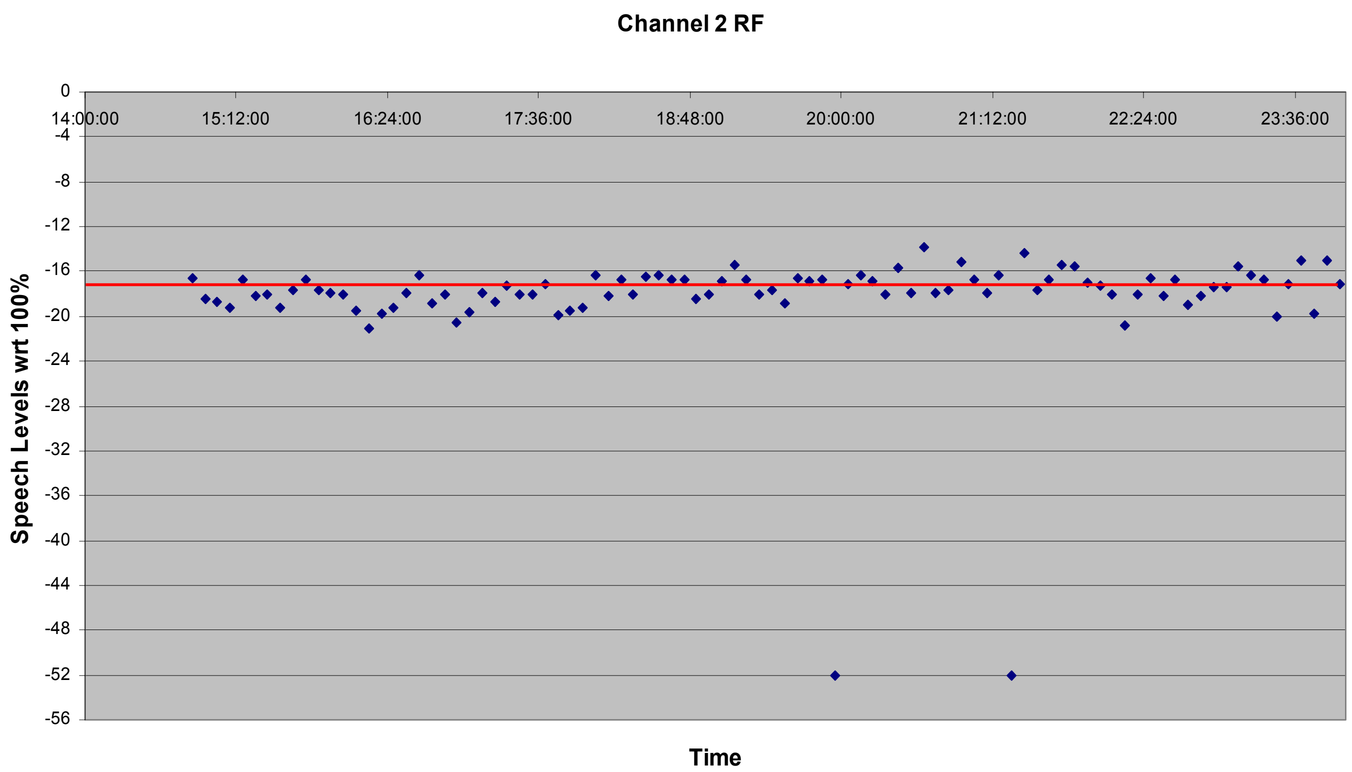

Figure 9

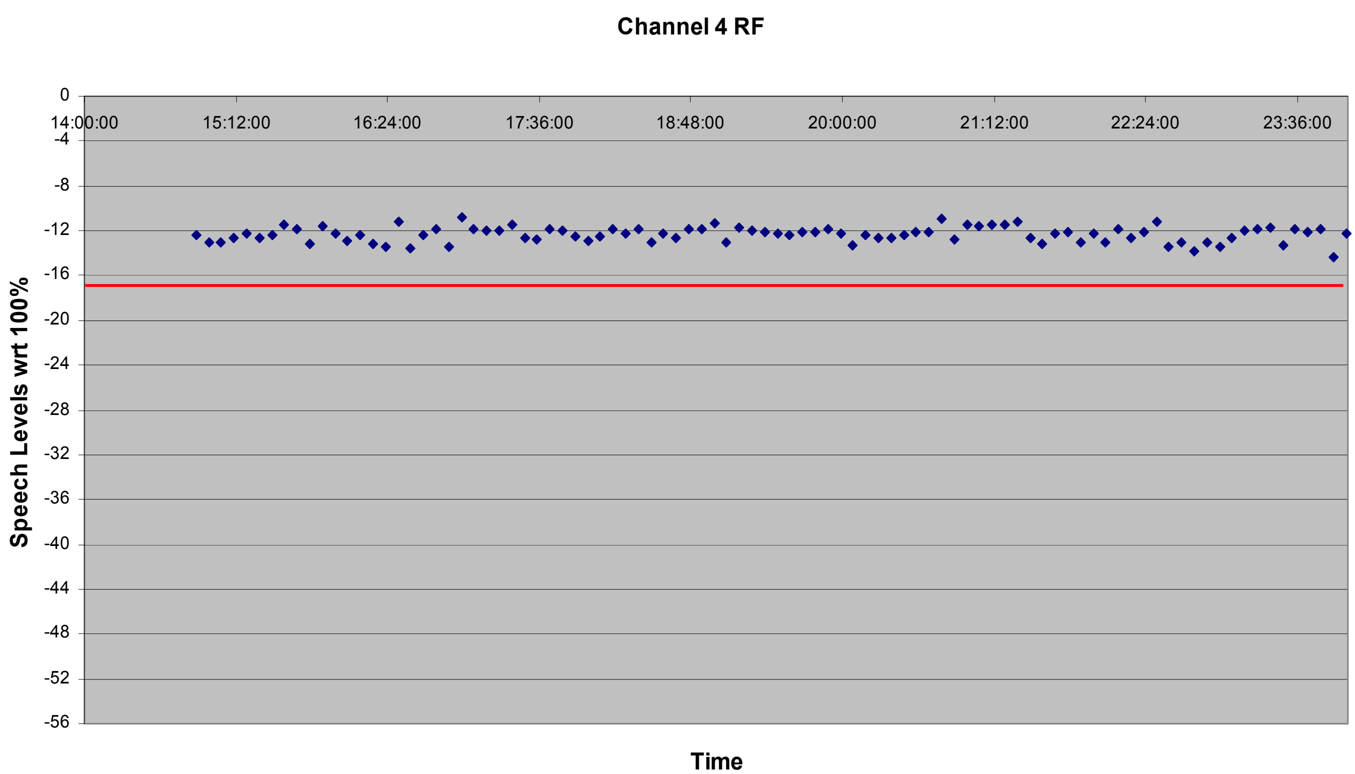

Figure 10

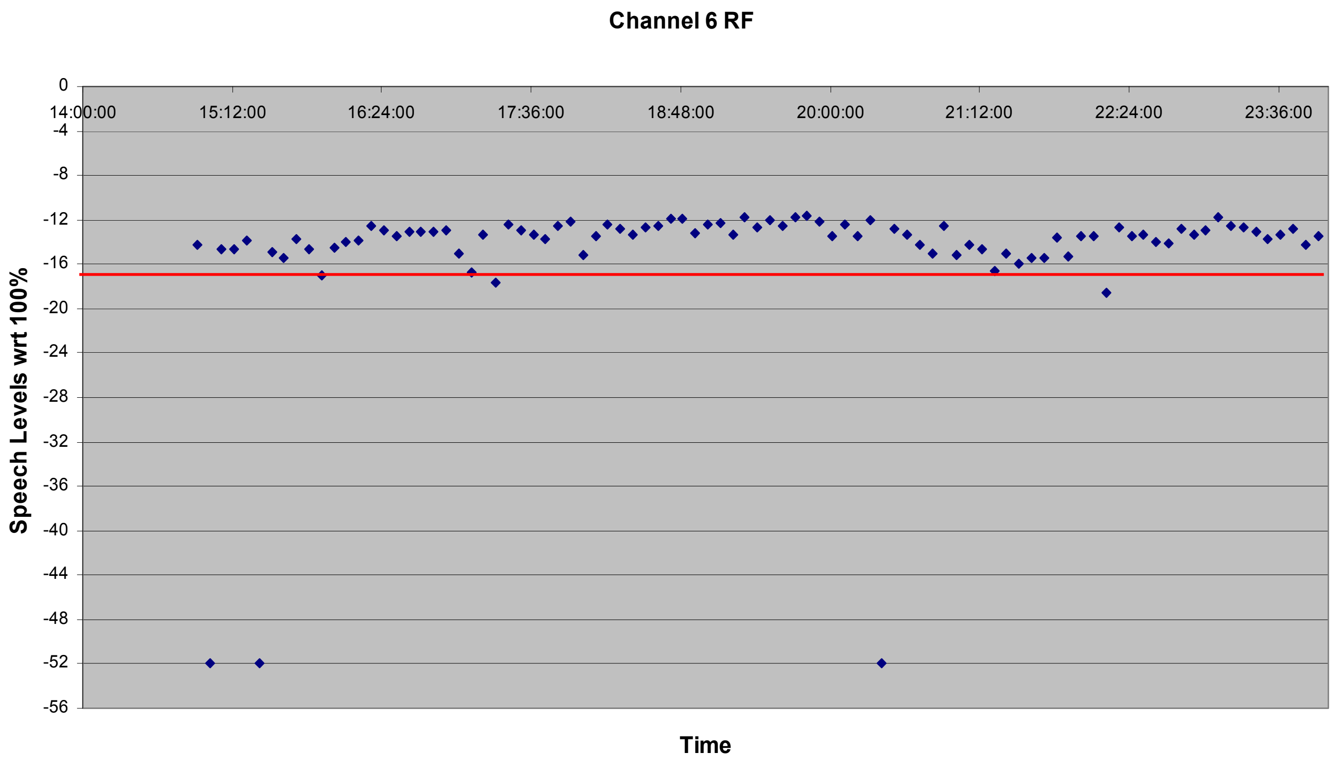

Figure 11

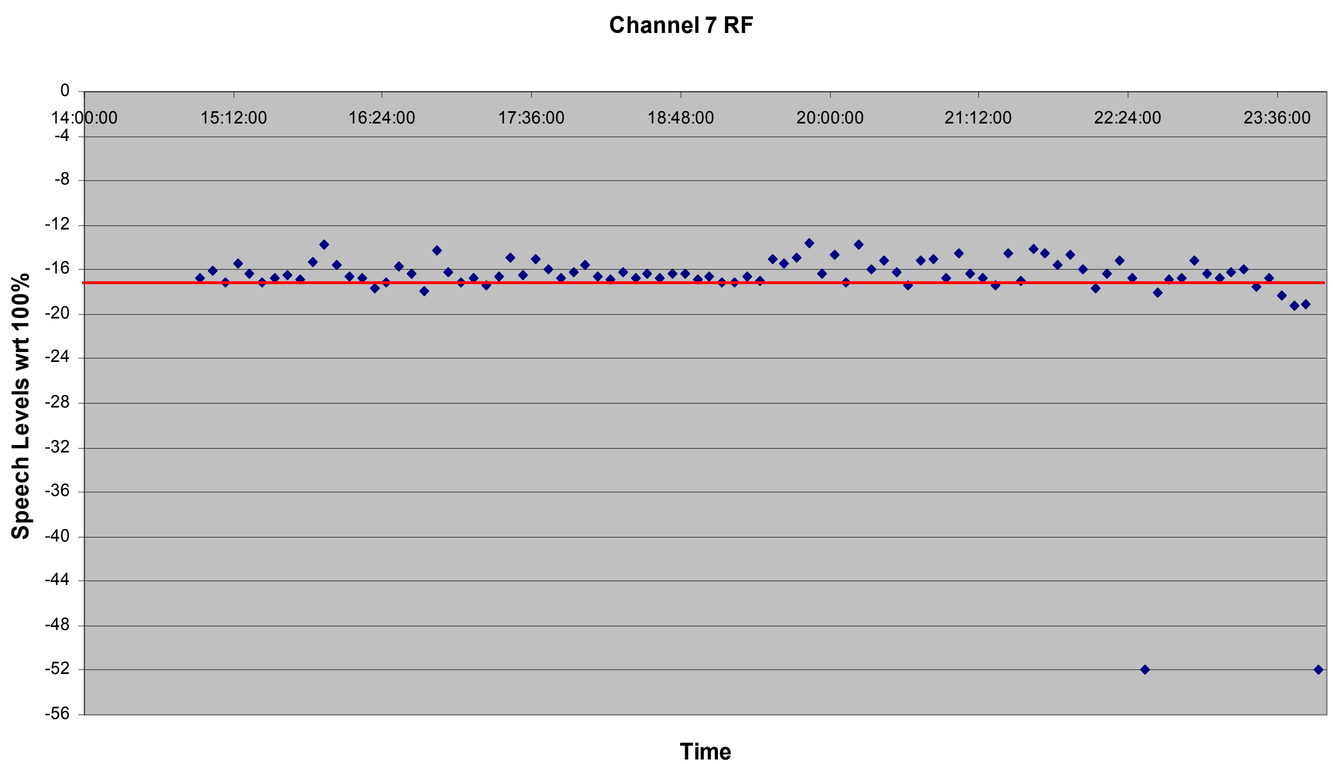

Figure 12

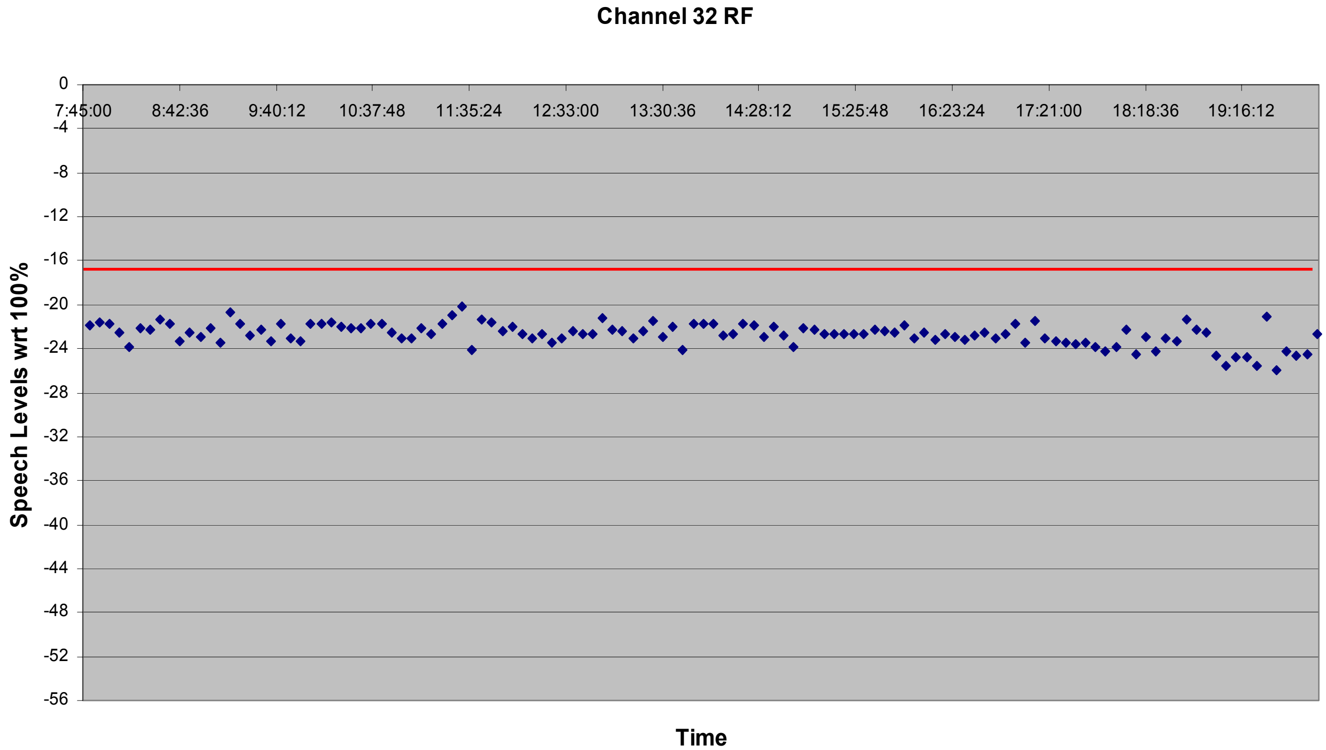

Figure 13

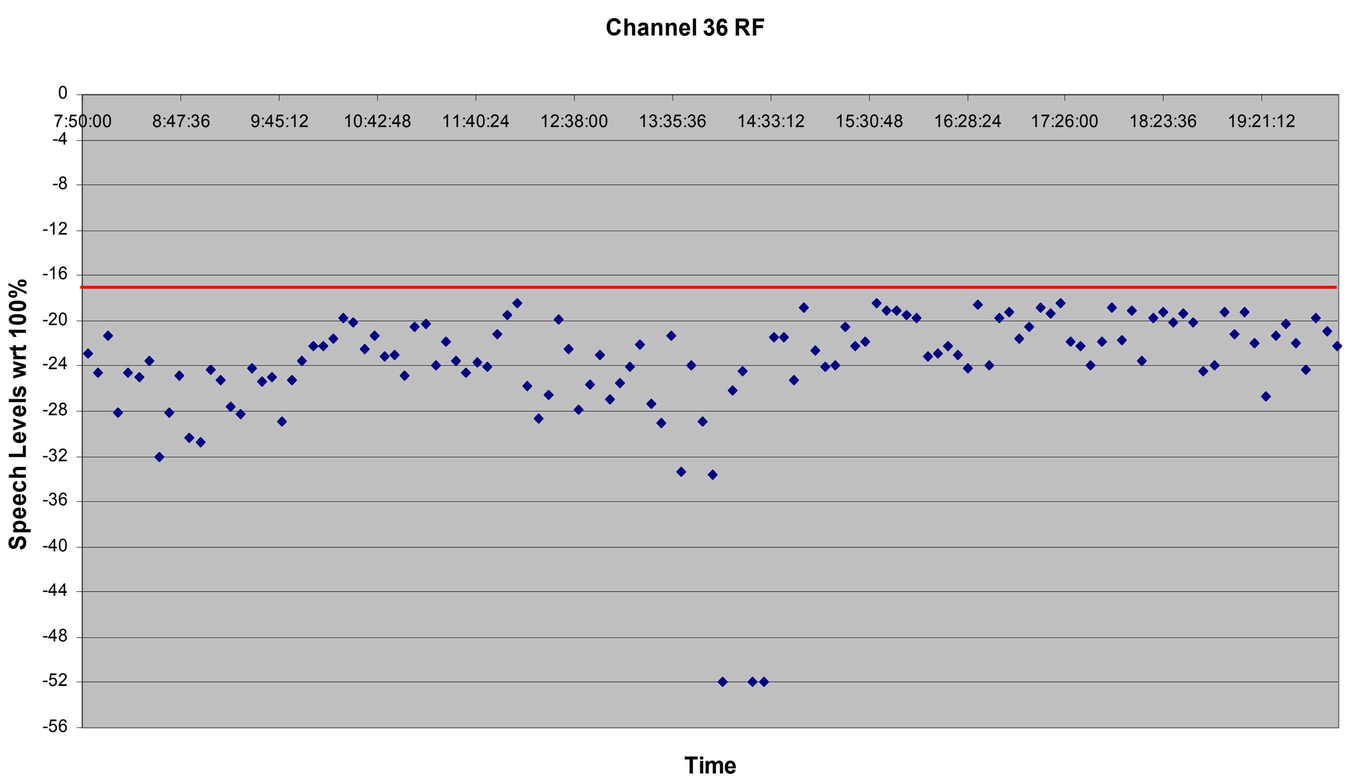

Figure 14

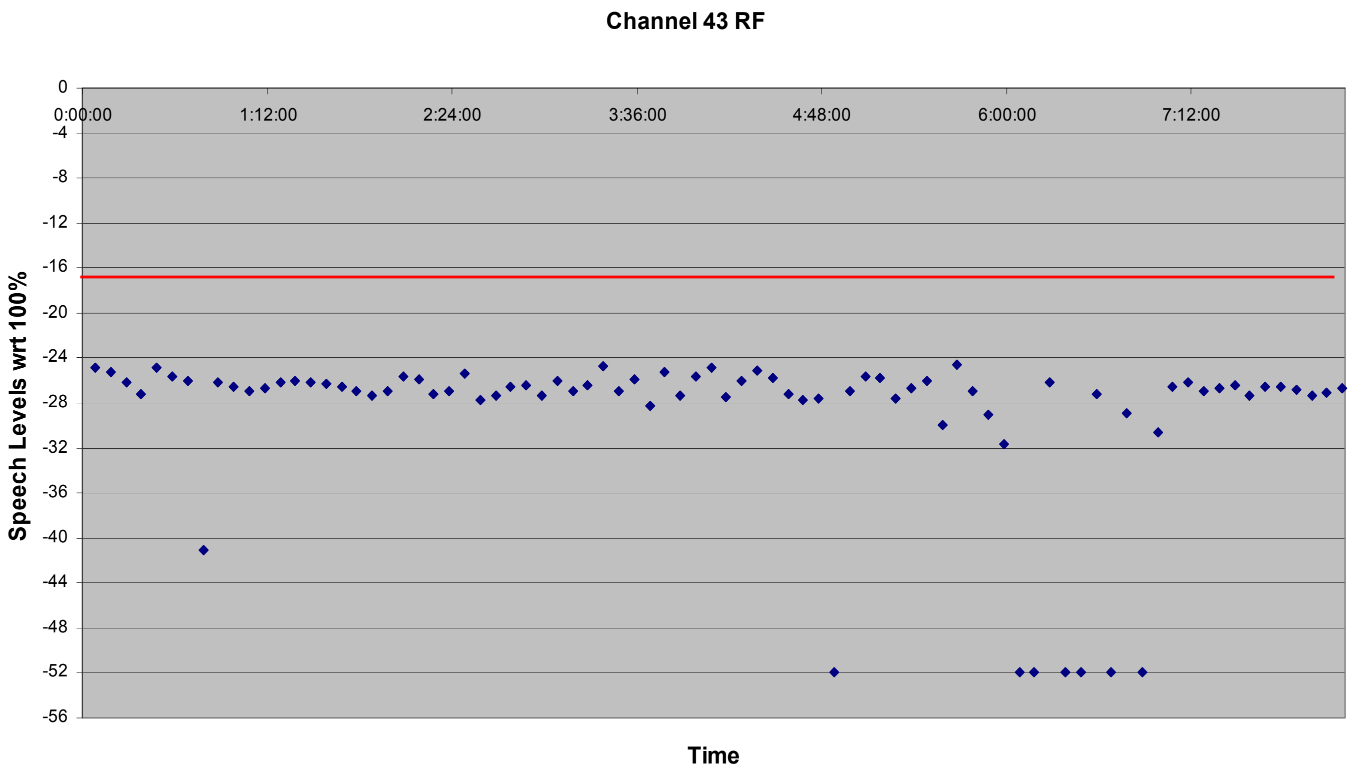

Figure 15

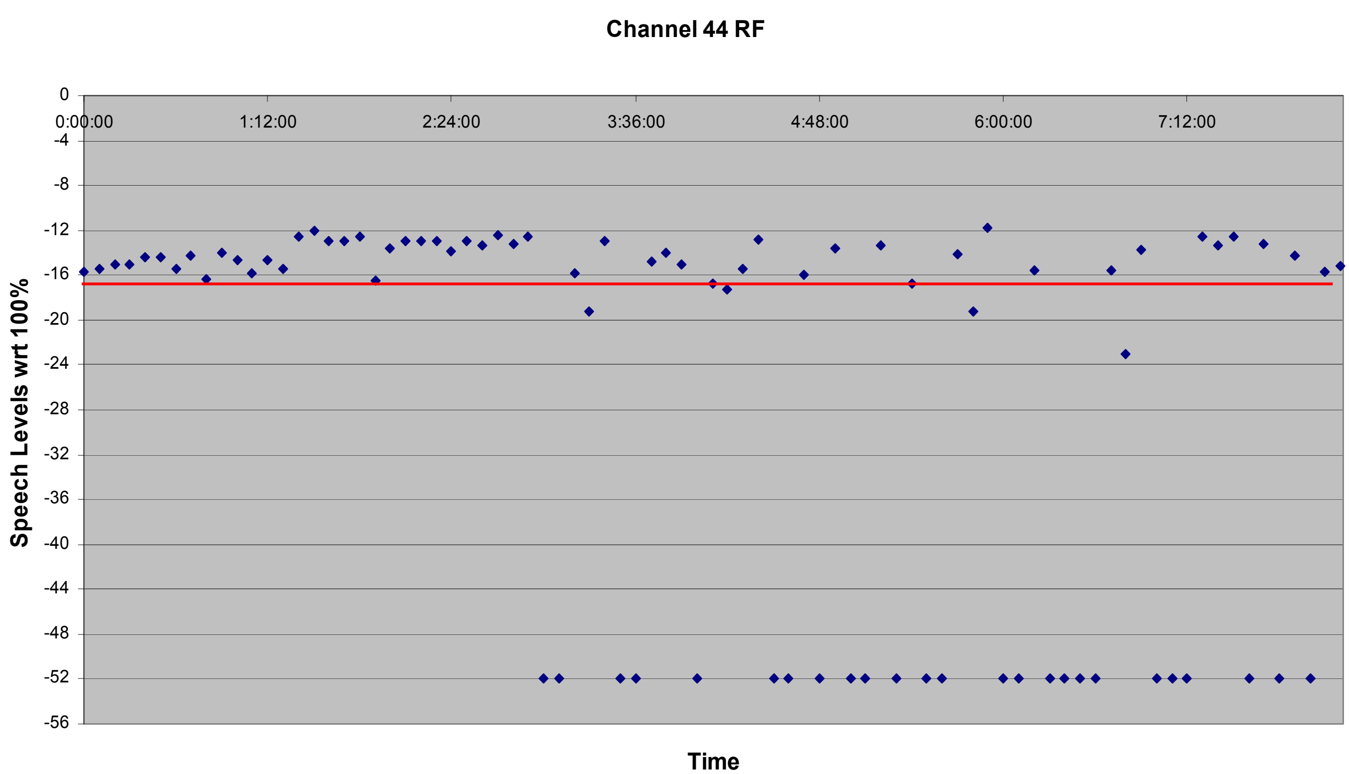

Figure 16

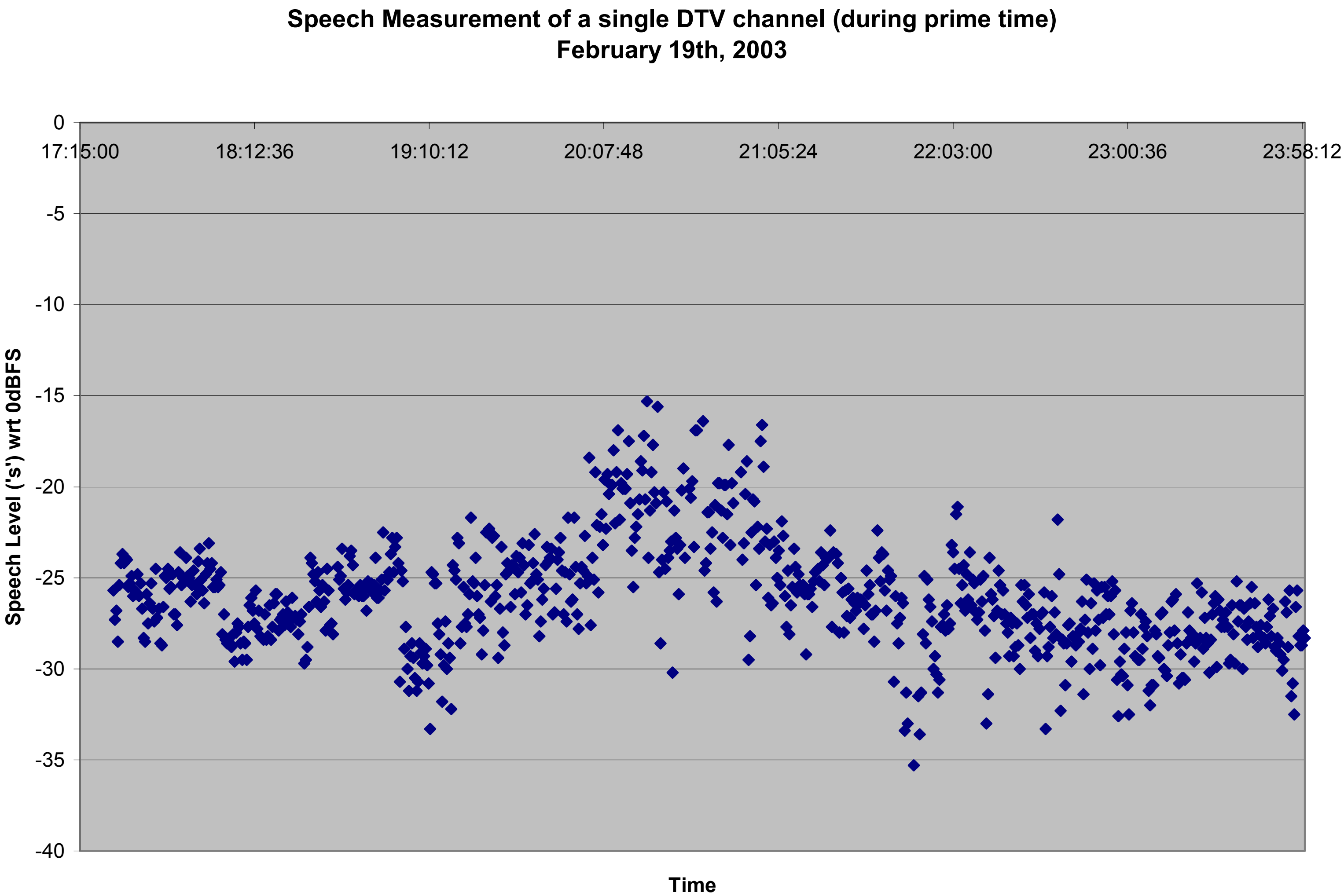
Speech Measurement of a single DTV channel (during prime time)
February 19th, 2003
0
-5
-10
-15
-20
-25
-30
-35
-40
17:15:00
18:12:36
19:10:12
20:07:48
21:05:24
22:03:00
23:00:36
23:58:12
Speech Level ('s') wrt 0dBFS
Time

Figure 17

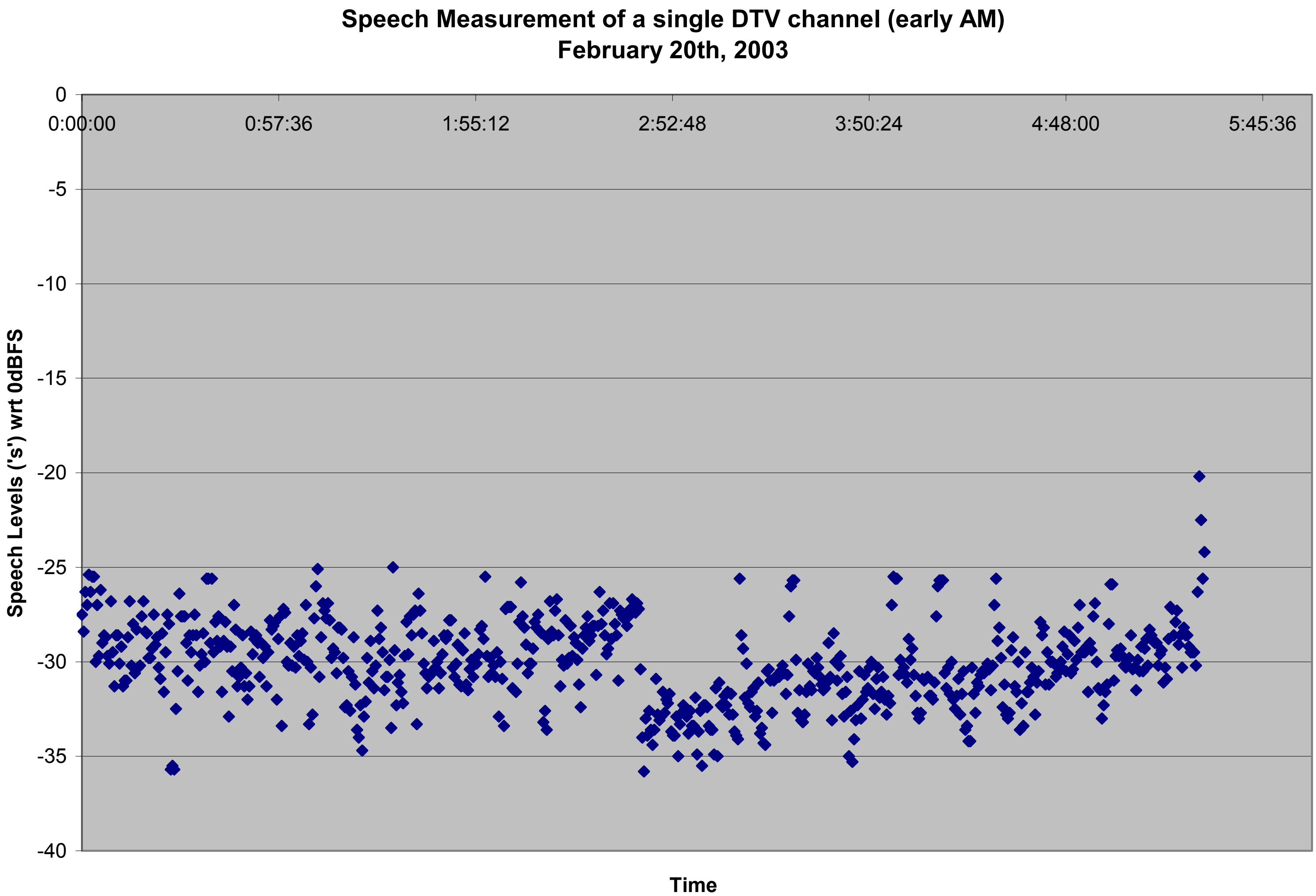
Speech Measurement of a single DTV channel (early AM)
February 20th, 2003
Speech Levels ('s') wrt 0dBFS
Time
0
-5
-10
-15
-20
-25
-30
-35
-40
0:00:00
0:57:36
1:55:12
2:52:48
3:50:24
4:48:00
5:45:36

Figure 18