VoIP Acoustic Design

Hints on How to Get Better Acoustic Performance on a VoIP Phone

Project Code:
Project Name: VSMPG

Revision History			
Rev.	**Date**	**Author**	**Description**
2.11	2008-09-18	HH	Added Chapter 2.3.10.
2.10	2008-09-12	HH	Initial version.

Table of Contents

List of Figures

1 Introduction

The intent of this document is to present suggestions on how to increase sound quality of VoIP Speaker Phones built upon the basic design of VLSI Solution's VS1000+1003 VoIP Speaker Phone demonstration units. This is done by analyzing acoustic feedback that can have an adverse effect on sound quality. The document also gives some hints on how to minimize the effect of acoustic feedback.

Chapter 2 presents what acoustic feedback is and describes the various ways of how it can manifest itself. Then it shows a case study of how acoustic feedback can be minimized.

Chapter 3 describes the Pass-Through VoIP Software that is needed for analyzing acoustic feedback and gives instructions on how to load it to a unit.

Finally, contact information is given in Chapter 5.

2 How to Combat Feedback

This chapter aims to answer to the question: What is feedback and what can you do to combat it?

2.1 What Feedback Is

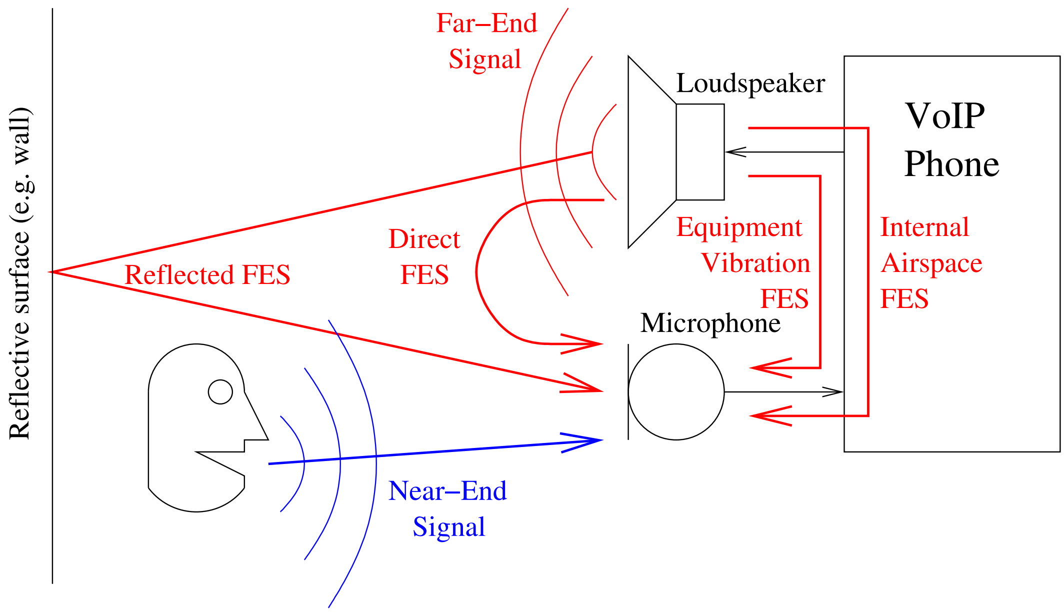

Figure 2.1: Speaker Phone Signals

In the context of speaker phones, feedback is considered to be audio that is by some means transmitted from the of a phone loudspeaker back to its microphone and thus back to the speaker phone and potentionally to the other party.

Speaker phone audio signals, including four mechanisms for feedback are shown in Figure 2.1.

- The Near-End Signal (NES) is that part of the audio signal that the microphone actually *is supposed to* capture and pass on.
- Direct FES (Far-End Signal) is the part of the speaker signal that is transmitted directly from the loudspeaker through the air to the microphone. This can be minimized by designing the device with directional loudspeaker faced away from

the microphone, a directional microphone and by putting the loudspeaker and microphone as far away from each other as possible.

- Equipment Vibration FES is usually caused by the main chassis, which vibrates along with the speaker. This vibration may be transmitted to the microphone and cause very high microphone signal levels.
- Internal Airspace FES is sound that occur inside a unit chassis. There may be structures which boost sound levels and may cause very high microphone signal levels.
- Reflected FES (Far-End Signal) is the part of the loudspeaker signal that gets reflected from different surfaces and get transmitted back to the microphone. This is the signal that adaptive echo cancellation has been designed to combat.

2.2 Feedback Mechanisms in VoIP Speaker Phone Demonstration Unit

Figure 2.2: VoIP Speaker Phone Demonstration Unit

This Chapter will discuss the relevance of the four feedback mechanisms to the VS1000+1003 VoIP Speaker Phone (VVSP) demonstration unit, depicted in Figure 2.2.

The loudspeaker and microphone of the VVSP are omnidirectional, so Direct FES could be a real problem. This has been solved by putting the microphone and loudspeaker as far as possible from each other.

Equipment Vibration FES is not an issue as there is no chassis to conduct vibrations. The same goes for Internal Airspace FES. Thus, the only remaining feedback method is Reflected FES.

All in all, VVSP works very well even without a directional microphone and loudspeaker. Note, however, that this wouldn't probably be as easy if the unit actually had a chassis.

2.3 Case Study: Feedback in an Actual Product

When a product is built, there has to be a chassis containing the electronics, and audio design gets that much harder. This chapter will discuss one example product and will show some suggestions on how to solve the problem.

2.3.1 The Product

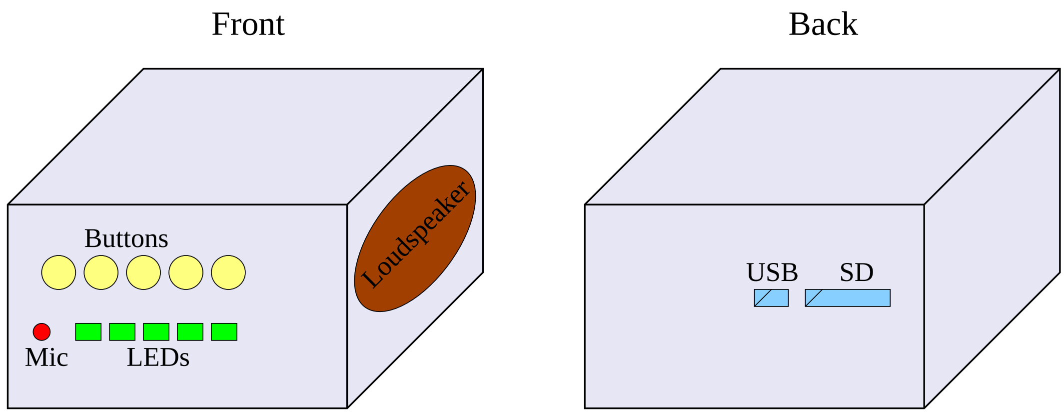

Figure 2.3: Speaker Phone Chassis

The product is a speaker phone, which is in principle as depicted in Figure 2.3. The user interface (buttons and LEDs) and the microphone are on the front side. USB and SD connectors live on the back side.

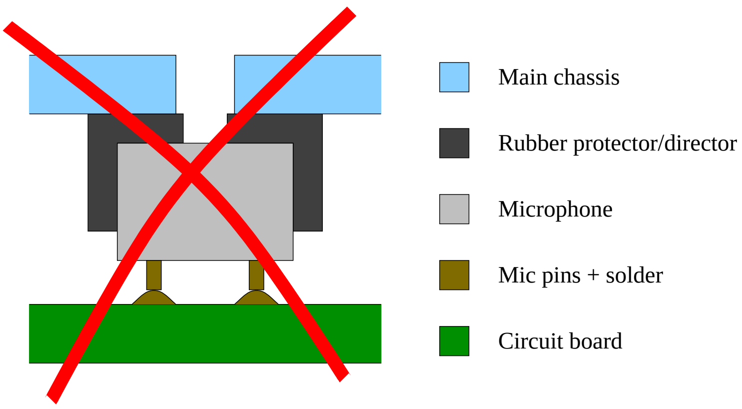

Figure 2.4: Speaker Phone Microphone Encapsulation (Figure 2.13 Is Better)

Figure 2.4 shows how the microphone is been connected. On the other side the microphone has been soldered to a circuit board, and on the other side it has been connected to the main chassis through a rubber protection / director cap. It will later be shown why the microphone shouldn't be connected in this way in a real product. (If you want to see the better setup immediately, see Figure 2.13 at Page 14.)

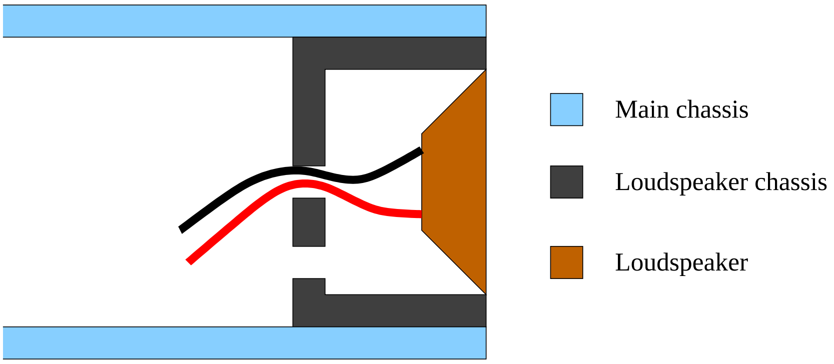

Figure 2.5: Actual Speaker Phone Loudspeaker Encapsulation

Figure 2.5 shows that the loudspeaker has been installed in a separate, plastic chassis of its own. The chassis has relatively thin walls and two holes: one for allowing speaker wires to come through, and another for other purposes.

The reported problem with the product was that echo cancellation didn't work properly. The reasons and solution are discussed in this Chapter.

2.3.2 Measurement Methodology

To determine the amount of speaker sound picked up by the microphone, the following measurement methodology was used:

1. A special pass-through version of the VoIP software was loaded to the unit. This software version does not perform echo cancellation or any other post-processing to the input signal except for a DC removal filter. Thus it will accurately represent what was actually recorded by the microphone. This software version is provided in the same package as this document. See Chapter 3 for details.
2. Wideband white noise sampled at 32 kHz and recorded with an amplitude of -20 dB of maximum amplitude was played to the speakers using full volume on the operating system.
3. Microphone input was recorded for $\approx$10…20 seconds to an 16-bit, 8 kHz mono WAV file.
4. The power and spectrum of the WAV file was measured.

2.3.3 Reference 0.0 dB: VLSI Solution's Demonstration Board

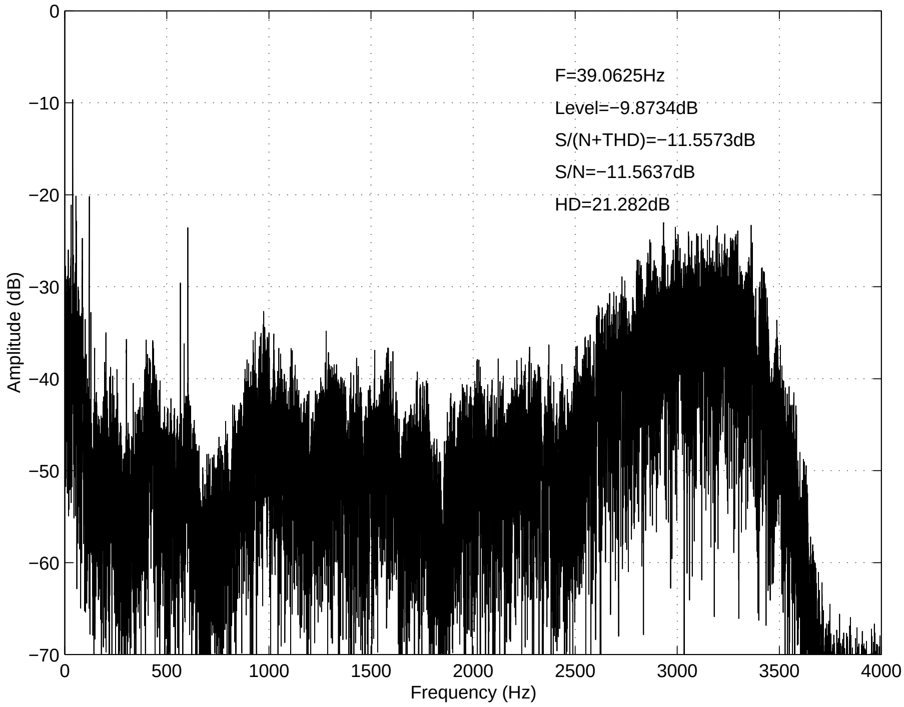

Figure 2.6: 0.0 dB: VLSI Solution's Demonstration Board

As a reference, VLSI Solution's own demonstration board setup (seen in Figure 2.2 at Page 6) was measured. The frequency spectrum of this measurement is presented in Figure 2.6. The feedback floor is very low between 200. . . 2500 Hz, though it rises over 10 dB for frequences between 2700 and 3400 Hz.

Overall, this is performance that ensures that echo cancellation works exactly as intended, although getting the same performance from a device that actually has a chassis may prove problematic.

2.3.4 Meas +30.5 dB: As Is

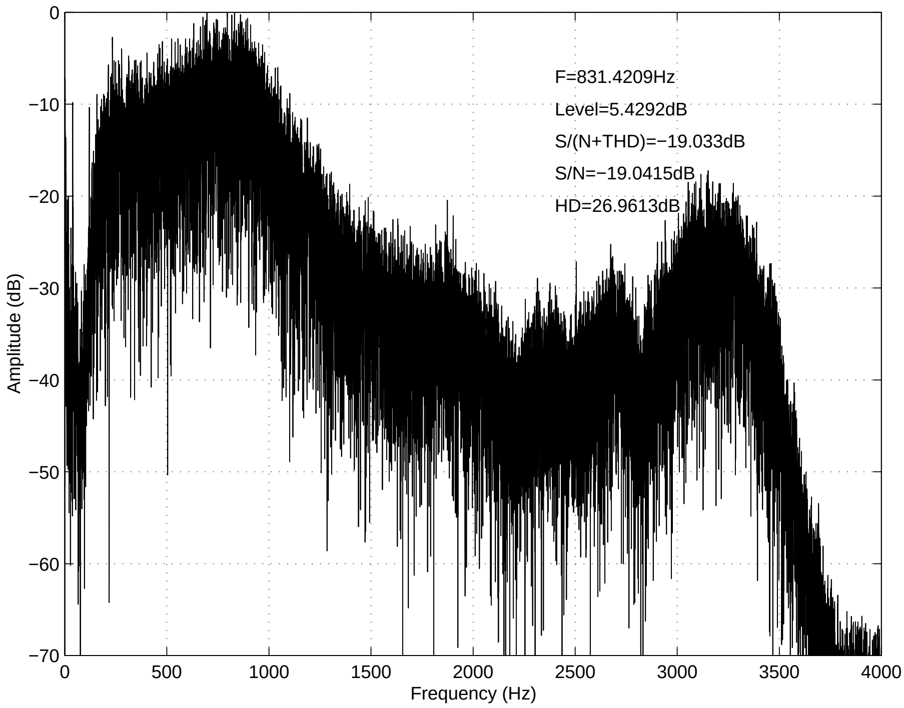

Figure 2.7: +30.5 dB: As Is

First the unit was tested as is, without any modifications except for loading the pass-through software. The spectrum is presented in Figure 2.7. As can be seen, frequencies betweem 200. . . 1400 Hz are greatly boosted.

The overall signal power was a whopping 30.5 dB stronger than the reference so it is not wonder why echo cancellation wouldn't perform properly.

2.3.5 Meas +16.0 dB: No Chassis

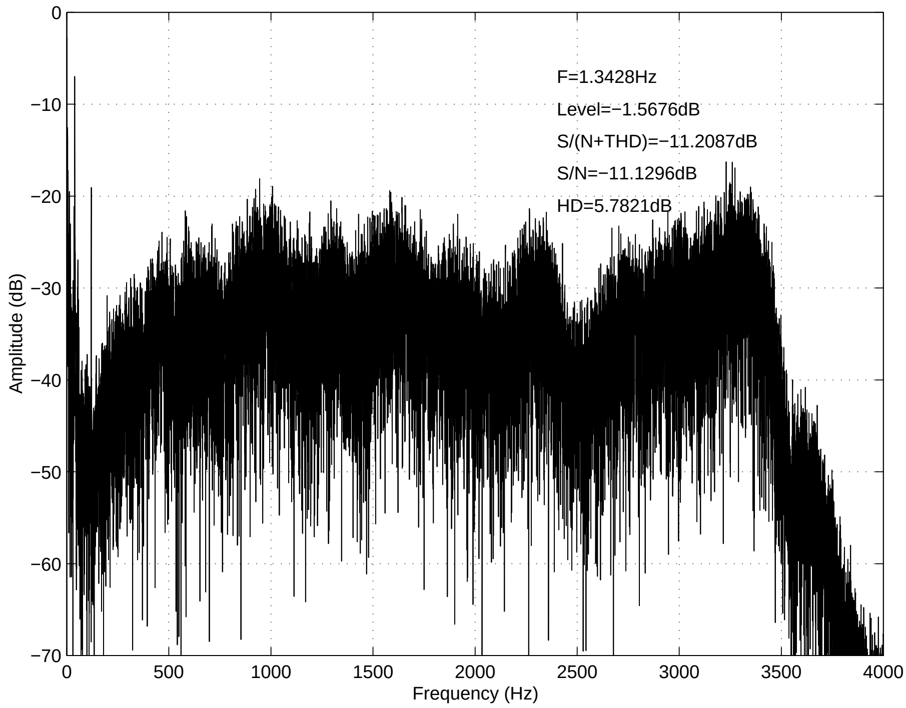

Figure 2.8: +16.0 dB: No Chassis

The main chassis was removed in such a way that the speakers and microphone still were at the same relative positions as before. This resulted in a much enhanced performance as seen in Figure 2.8. All the peaks are gone and the resulting frequency response is relatively straight.

The signal's power has decreased to +16.0 dB. Although it still is much greater than what was with VLSI Solution's Demonstration Unit, the difference is big enough that VoIP functionality works with the official VoIP firmware.

This solution was of course not practical in itself, because a product needs to have a chassis. Nevertheless, now it was clear that the main chassis somehow acted as a sound conductor, because removing it decreased FES so significantly.

The chassis was put together again for the next tests.

2.3.6 Meas +28.5 dB: Speaker Padding

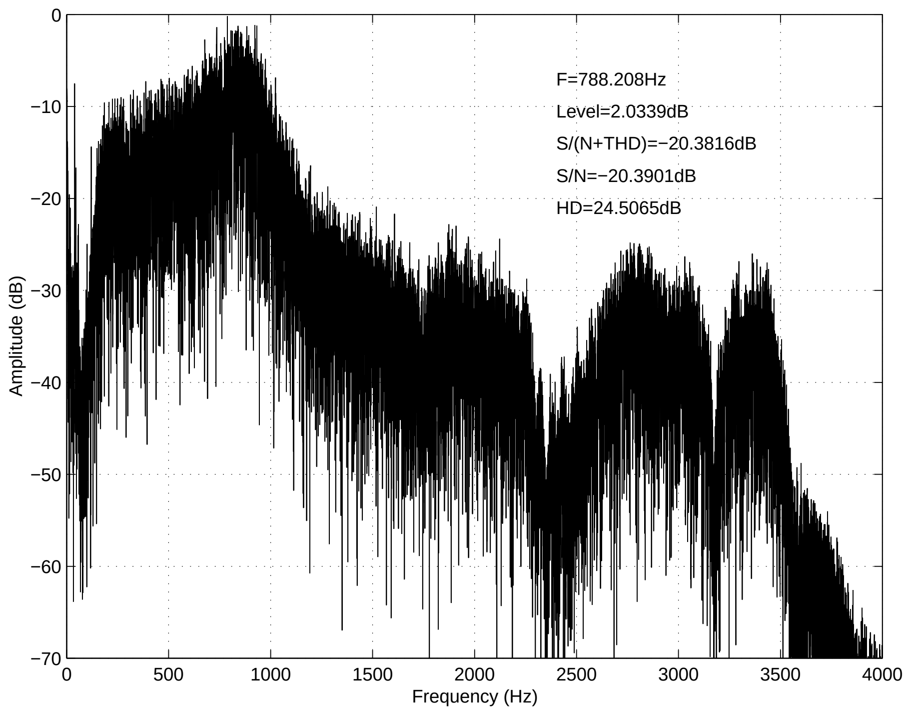

Figure 2.9: +28.5 dB: Speaker Padding

It was felt that the back wall of the speaker element might be vibrating too much, so some mass was added to the back of the inside of the speaker chassis to make it a little bit heavier. Also the speaker chassis holes were filled so that the speaker would be a little bit more like what is shown in Figure 2.10.

This resulted in noticeable attenuation at frequences over 2300 Hz and a signal level of +28.5 dB as shown in Figure 2.9. Unfortunately the main problem, the peak between 200. . . 1200 Hz, was unaffected.

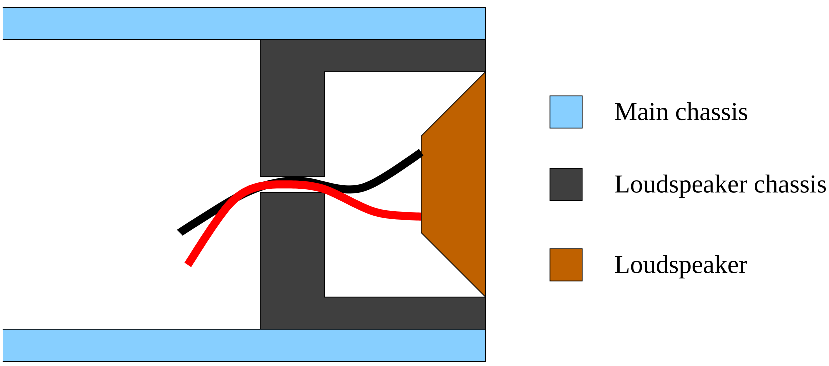

Figure 2.10: Speaker Phone Loudspeaker Encapsulation with Thicker Back Wall

2.3.7 Meas +27.6 dB: Padding around Microphone Circuit Board

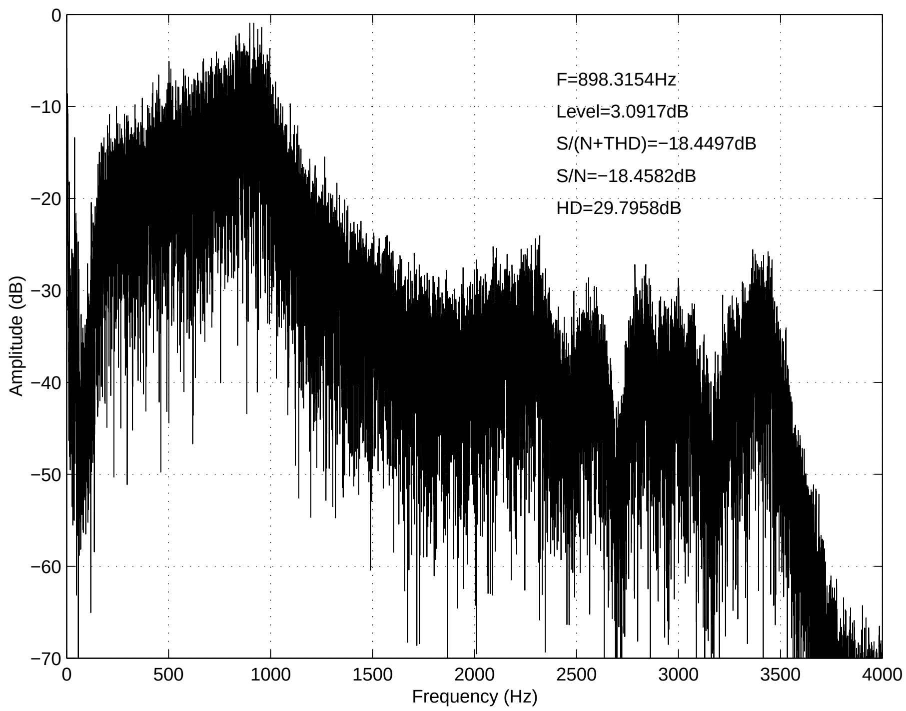

Figure 2.11: +27.6 dB: Blu Tack around Microphone Board

Low-frequency boost could be caused by Equipment Vibration FES through the microphone circuit board. Thus, padding material was put both above and below the microphone board to decrease vibrations.

As can be seen in Figure 2.11, the difference between this and the previous case was insignificant.

2.3.8 Meas +21.8 dB: Mic Removed from Circuit Board

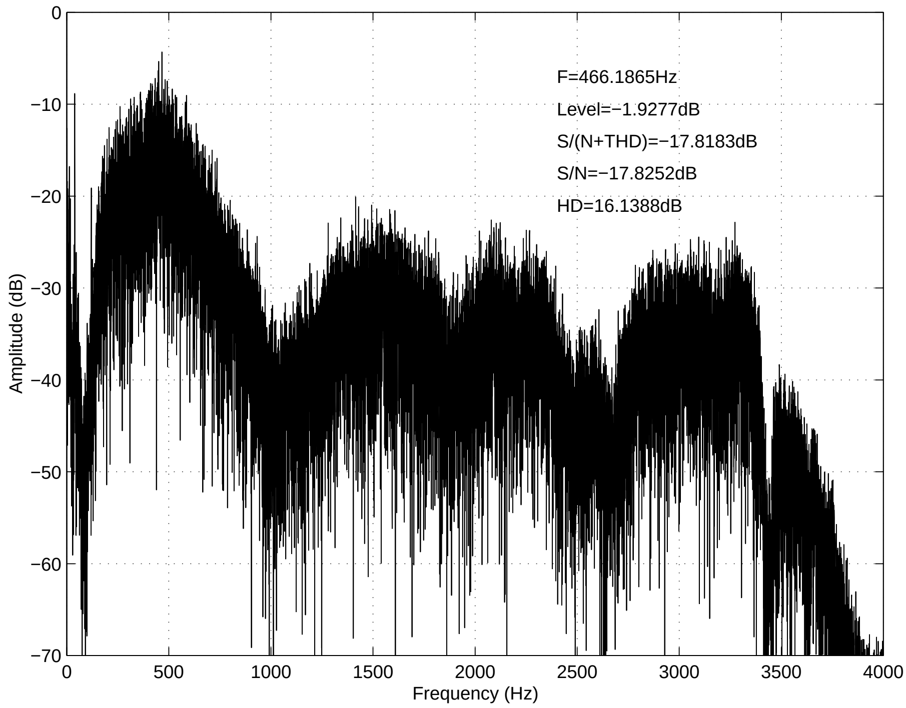

Figure 2.12: +21.8 dB: Free-Hanging Microphone, Closed Chassis

To further combat potential Equipment Vibration FES, the microphone was disconnected from the microphone circuit board. Then it was connected back with thin electrical wires as shown in Figure 2.13. The microphone was secured to the main chassis with adhesive material.

As can be seen in Figure 2.12, preventing Equipment Vibration FES from the circuit board helped: while it didn't affect frequencies below 500 Hz, it greatly reduced frequencies in the most problematic area at 800. . . 1300 Hz. Now signal level was +21.8 dB.

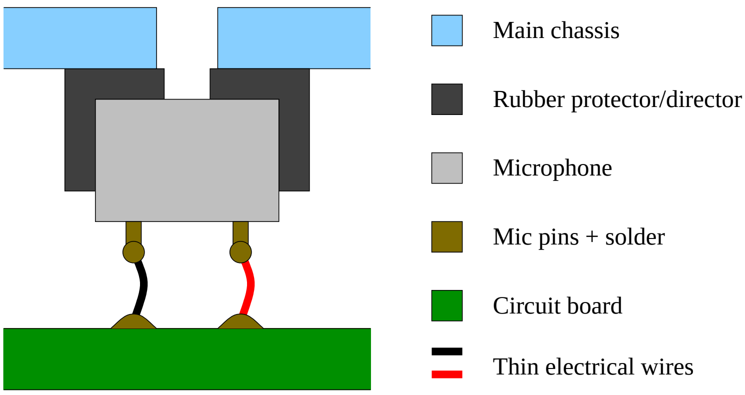

Figure 2.13: Better Speaker Phone Microphone Encapsulation

2.3.9 Meas +16.2 dB: Mic Removed from Circ. Board + Chassis Hole

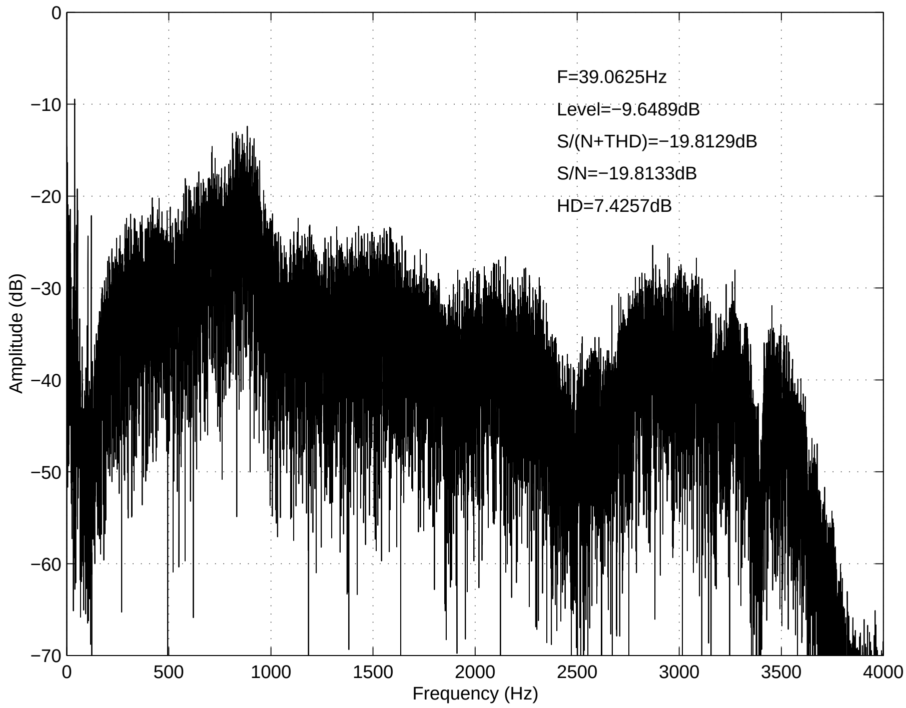

Figure 2.14: +16.2 dB: Free-Hanging Microphone, Open Chassis

Removing the microphone from the main circuit board helped a lot, but there was still much room from improvement.

An extra hole, roughly the size of the SD port hole, was added to the main chassis. As can be seen in Figure 2.14, this moved the high peak frequency range from 300. . . 600 Hz to 700. . . 1000 Hz, but above all it lowered signal level. Now the total signal level was +16.2 dB above the reference level. At frequencies 2800. . . 3400 Hz it is in fact better than the reference unit in Chapter 2.3.3, although at lower frequencies it still is significantly worse. The signal level was now as low as it was in Chapter 2.3.5 where no chassis was used. This was low enough for the VoIP speaker phone to work, although even lower levels would be preferable.

2.3.10 Meas +15.8 dB: Microphone Glued to Chassis

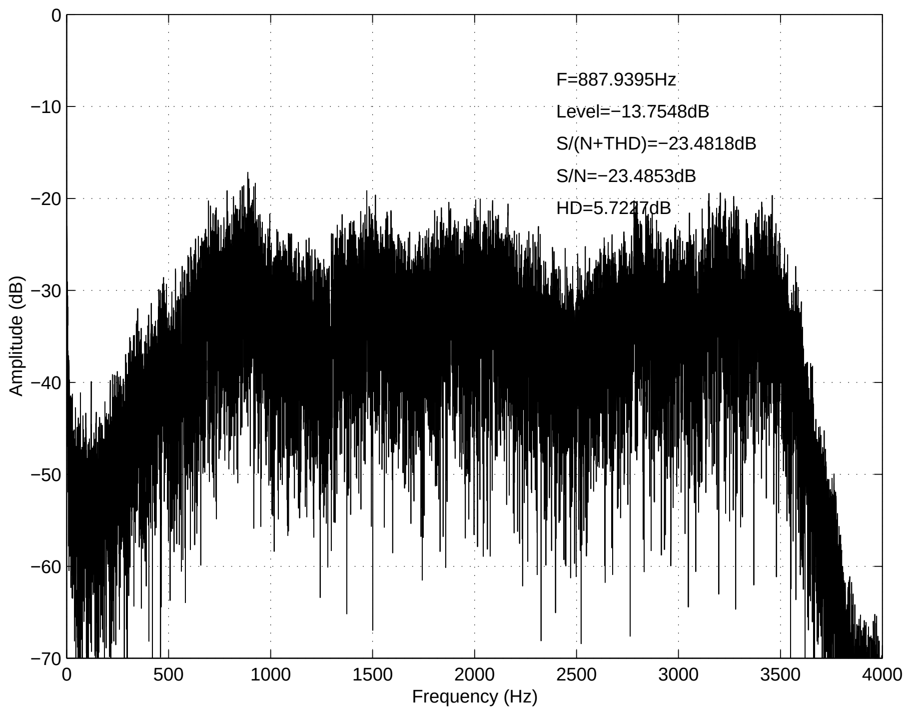

Figure 2.15: +15.8 dB: Microphone Glued to Chassis

Another solution to separate the microphone from the vibrations of the main board and plastics was to solder the microphone to a smaller board and glue them tightly to the main chassis, as shown in Figure 2.16. This unit was then connected to the main board and plastics with a 15 cm thin ribbon cable.

Figure 2.15 shows that this gave the best results: a uniform frequency response and the lowest signal amplitude, and all of this without drilling extra holes to the chassis. Opening the chassis could potentially give even better results, but this was not tested.

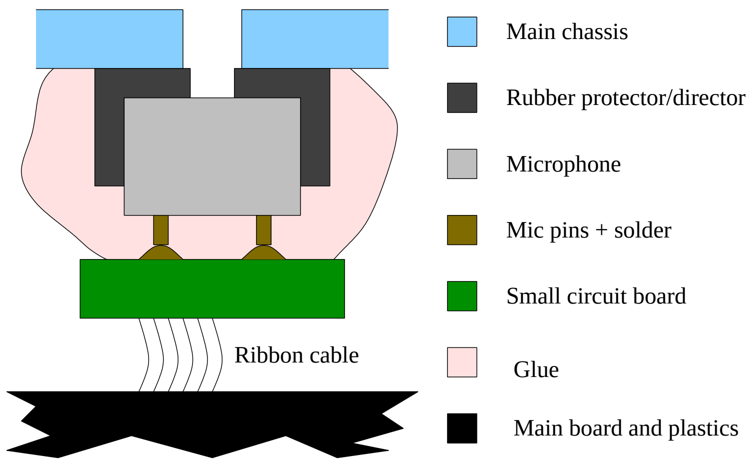

Figure 2.16: Microphone Board Glued to Chassis

2.3.11 Conclusions

The basic problem with the original product was that it transmitted too much audio signal from the speaker to the microphone (Figure 2.7 on page 10), which prevented echo cancellation from working efficiently. The main mechanisms appeared to be Equipment Vibration FES and Internal Airspace FES (see Chapter 2.1 for different types of FES feedback).

The research done to the product was not exhaustive, and further experiment would almost certainly result in lower FES feedback levels, which in turn would lead to even better sound quality. It must be remembered that echo cancellation is not a miracle cure for everything: the device that intends to use echo cancellation muste have good acoustic design to begin with. Only when these things are in order can echo cancellation do the thing it is designed for: filtering out Reflected FES and traces of the other FES types.

3 Pass-Through VoIP Software

3.1 What the Software Does

The Pass-Through VoIP Software is an alternative firmware to the default VoIP phone firmware. It doesn't have automatic gain control, echo cancellation or extraneous filtering except for a DC removal filter. The software has been intended for testing acoustic feedback as presented in Figure 3.1.

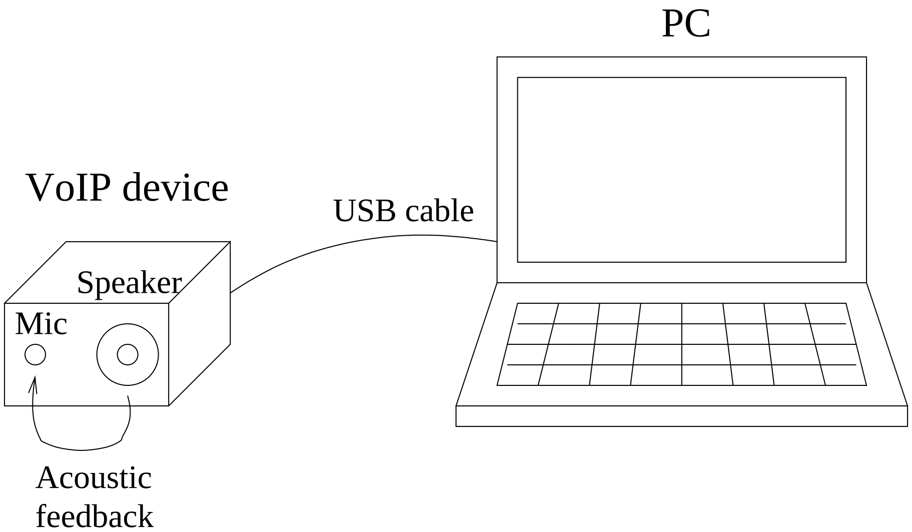

Figure 3.1: PC Setup for Audio Feedback Measurement

A test signal sampled at the VoIP application's native output sample rate of 32 kHz is played back from the PC to the VoIP device's speaker. At the same time the signal from the microphone is recorded with a PC recorder application. The resulting file is an 16-bit 8 kHz mono file, that consists of the microphone input. *Note that some PC recording applications record at a default of mono 8 bits. Remember to set your PC recording application to save 8 kHz mono with 16 bits!*

The file can be used for feedback analysis. As an example, if the test signal is wideband white noise the input can be analyzed with any good spectral analyzer program that can also display total signal power.

In this package there is a file called `Acoustic/Noise32k.wav`, which consists of 10 seconds of wideband white noise sampled at 32 kHz and encoded with 8 bits at a signal level of -20 dB of maximum amplitude. By recording 5. . . 10 seconds of this signal, both speaker-to-mic feedback spectrum and power can be resolved.

3.2 How to Load the Software

To update to a new software version, you need a PC/Windows computer with an RS232 port, an RS232 cable and an RS232 adapter. If you don't have an RS232 adapter, you can order one from VLSI Solution or build it yourself: the adapter consists of one single MAX3232 compatible RS232 signal converter.

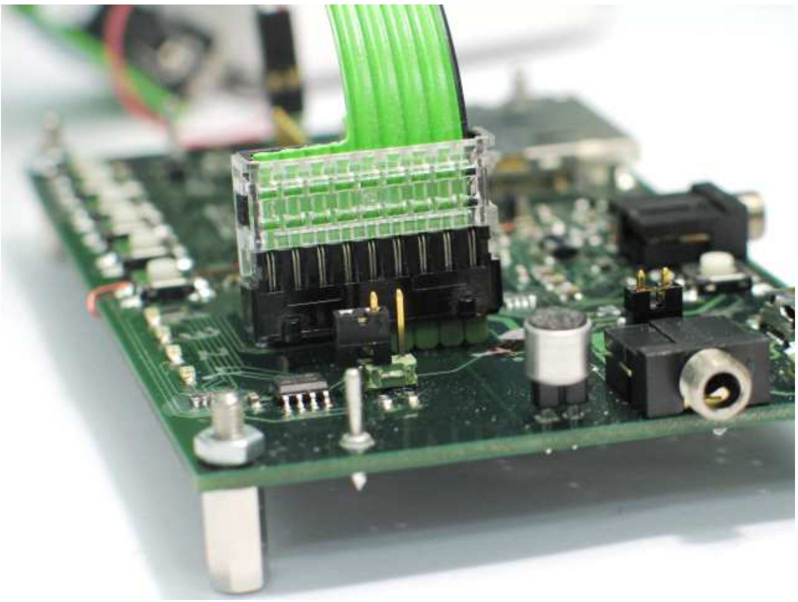

Figure 3.2: How to Connect an RS232 Adapter to VLSI's VoIP Board

If loading a program to VLSI Solution's example boards, connect the adaptor cable to the main board as shown in Figure 3.2. The black thread should go to pin 1. If you use a custom board like the Fujitsu unit, you will have to use your own adapter.

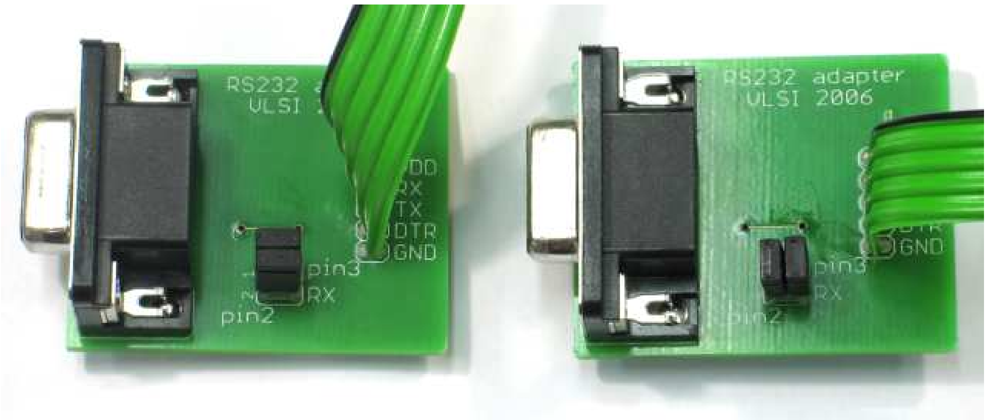

Figure 3.3: Two Possible RS232 Converter Jumper Configurations

This package includes command files named `prom1.bat` through `prom4.bat`. They are intended to be used from serial ports COM1 through COM4, respectively.

To load the new code to the board, perform the following steps:

1. Open the .zip package that contains the files. As you are reading this document, you have probably already done that.
2. Open a command prompt, and use the CD command to go to the `Acoustic/` directory where the files are.
3. Connect a serial cable between the PC and the VoIP Speaker Phone.
4. Turn the VoIP phone on. It doesn't matter what mode it is in.
5. Check which COM port it is on.
6. If COM1, run prom1.bat by typing `prom1` and pressing the Enter key. For COM2, run prom2.bat etc.
7. You should see roughy the following text appearing on the screen:

```
VSEMU 2.1 Nov 28 2007 11:50:01(c)1995-2007 VLSI Solution Oy
Clock 11999 kHz
Using serial port 1, COM speed 115200
Waiting for a connection to the board...
Caused interrupt
Chip version "1000"
Stack pointer 0x19e0, bpTable 0x7d4d
User program entry address 0x4083
eeprom.bin: includes optional header, 16 sections, 539 symbols
Section 1: code       page:0 start:80 size:3 relocs:1 fixed
[... many similar lines deleted ...]
Section 16: VS_stdiolib$0  page:0 start:689 size:134 relocs:37
0000
 status
...............................................................
Finished!!
```

   Never interrupt programming!
8. Sometimes you may only see

```
VSEMU 2.1 Nov 28 2007 11:50:01(c)1995-2007 VLSI Solution Oy
Clock 11999 kHz
Using serial port 1, COM speed 115200
Waiting for a connection to the board...
```

   In such a case you have either used a wrong COM port in your script or you must change the jumper configuration as shown in Figure 3.3.

4 Document Version History

This chapter includes a detailed document and VoIP package version history.

2008-09-18 Version 2.11

- Added Chapter 2.3.10, *+15.8 dB: Microphone Glued to Chassis*. This alternative to just glueing the microphone to the main chassis gave the best results for the example product.

2008-09-12 Version 2.10

- First release.

5 Contact Information

VLSI Solution Oy
Entrance G, 2nd floor
Hermiankatu 8
FIN-33720 Tampere
FINLAND

Fax: +358-3-3140-8288
Phone: +358-3-3140-8200
Email: sales@vlsi.fi
URL: http://www.vlsi.fi/

For support, ask:
support@vlsi.fi