



## I. Introduction

Proper hardware design is essential to creating a high quality speech recognition product. The Sensory Interactive Speech™ ICs have the best combination of speech recognition and cost available today, but without an effective hardware design, the performance of these ICs will be severely degraded. This design note focuses on hardware issues, guiding the reader through important hardware design considerations. This note should be used in conjunction with the other Interactive Speech™ design notes to ensure the creation of a high quality speech recognition product.

**Sensory highly recommends that this procedure be used to design and test ALL new product designs.**

## II. Microphone

### A. Selecting a Suitable Microphone

1. **Microphone Types** - A microphone works by converting sound pressure into a varying voltage. There are two basic types of microphones: omni-directional and uni-directional. For most applications, an inexpensive omni-directional electret microphone is acceptable. In some applications, a directional microphone might be more suitable if the desired signal comes from a different direction than the audio noise. Since directional microphones have a frequency response that depends on their distance from the sound source and angle from center position, such microphones should be used with caution. If the product is intended to be used in a noisy environment, care should be taken to design around the noise. Improving the signal-to-noise ratio will help increase a product's speech recognition accuracy.

**Microphone Sensitivity** – A microphone with a minimum sensitivity of -36 to -46 dB ([0dB=1v/Pa @ 1KHz](#)) is usually required. For RSC-4x designs, Please refer to Sensory Application Note 80-0259, "Selecting A Microphone" for a list of recommended microphones. For NLP-5x designs, Please refer to Sensory Application Note 80-0318, "Selecting a Microphone and Gain Resistor for the NLP-5X" for a list of recommended microphones.

### B. Design of Microphone Housing

1. **Objective** - Proper design and consistent manufacturing of the microphone housing is important, because improper acoustic positioning of the microphone will reduce recognition accuracy. This section describes several important considerations that must be carefully followed in designing the microphone mounting and housing. Many mechanical arrangements are possible for the microphone element, and some will work better than others. Sensory recommends the following guidelines for designing the microphone housing:
2. **Flush Mounting** - In the product, the microphone element should be positioned as close to the mounting surface as possible and should be fully seated in the plastic housing. There must be NO airspace between the microphone element and the housing. Having such an airspace can lead to acoustic resonance, which can reduce recognition accuracy.

**Good:** Mic. element flush with surface

**Bad:** Air cavity between mic. element and surface.



Figure 1 -- Microphone Mounting (1)

- No Obstructions, Large Hole** - The area in front of the microphone element must be kept clear of obstructions to avoid interference with recognition. The diameter of the hole in the housing in front of the microphone should be at least 5 mm. Any necessary plastic surface in front of the microphone should be as thin as possible, being no more than 0.7 mm, if possible.

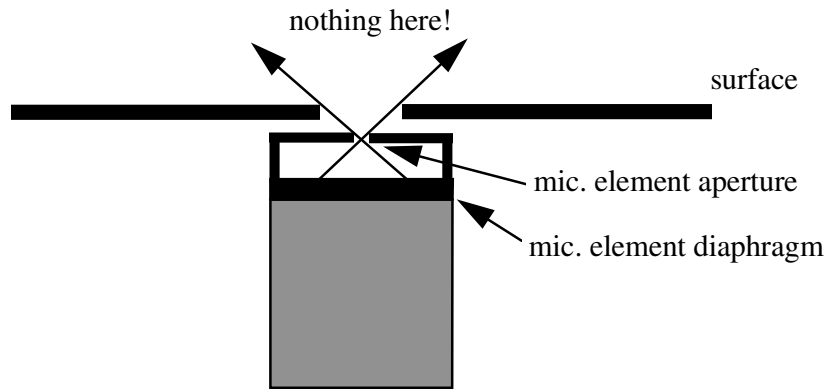


Figure 2 -- Microphone Mounting (2)

- Insulation** - The microphone should be acoustically isolated from the housing if possible. This can be accomplished by surrounding the microphone element with a spongy material such as rubber or foam. Mounting with a non-hardening adhesive such as RTV is another possibility. The purpose is to prevent auditory noises produced by handling or jarring the product from being "picked up" by the microphone. Such extraneous noises can reduce recognition accuracy.

**Good:** Absorbent material surrounding element.

**Bad:** Element fastened directly to housing.

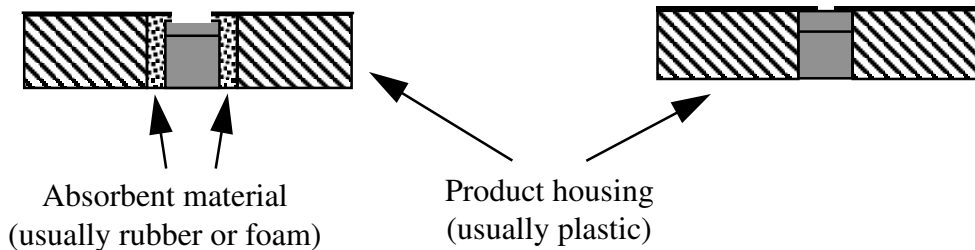


Figure 3 -- Microphone Mounting (3)

- Distance** - If the microphone is moved from 6 inches to 12 inches from the speaker's mouth, the signal power decreases by a factor of four. The difference between a loud and a soft voice can also be more than a factor of four. Although the internal preamplifier of the RSC-4x compensates for a wide dynamic range of input signal strength, if its range is exceeded, software can provide auditory feedback to the speaker about the voice volume. The product can achieve this by saying such as "please talk louder" or "please don't talk so loud."

### III. Schematic Design

#### A. Audio Amplifier

- RSC\_4x Reference Circuit** - The diagram in Figure 4 shows the microphone input circuit and the required external components and their connections to the RSC\_4x for the internal pre-amplifier circuit. The components for the circuit are listed in Table 1.

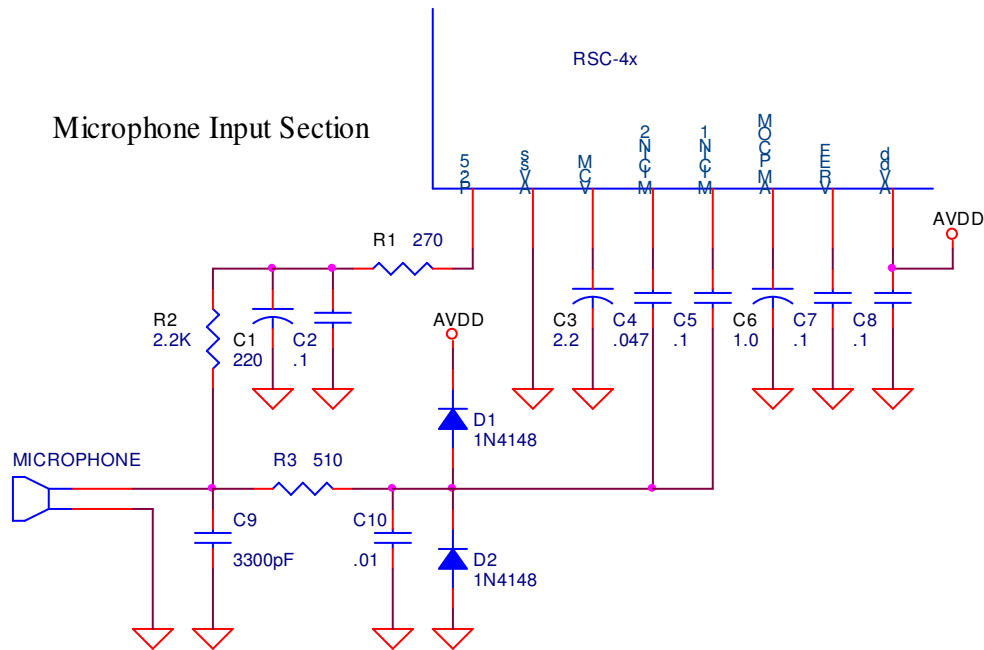


Figure 4 - Diagram for RSC\_4x Analog Section

Table 1 – Components for Microphone Input Circuit for RSC-4x

Resistors		
<b>R1</b>	270, 5%	Recommended value – Should be between 270 and 2000 ohms. Along with C1 & C2 provides power filtering – Actual values of C1 and R1 must be selected to give 60 millisecond RC time.
<b>R2</b>	2.2K, 5%	Actual value must match the microphone sensitivity rating. Please refer to Sensory Design Note 80-0259 (“Selecting a Microphone”).
<b>R3</b>	510, 5%	ESD protection – may be omitted if microphone is mounted on PCB.
Ceramic Capacitors		
<b>C2</b>	.1uF	Critical value.
<b>C4</b>	.047uF	Optional (required for audio wakeup only).
<b>C5</b>	.1uF	Critical value.
<b>C7</b>	.1uF	Critical value.
<b>C8</b>	.1uF	Decoupling between AVdd and Avss.
<b>C9</b>	3300pF	Recommended value – can be omitted in some designs.
<b>C10</b>	.01uF	ESD protection – may be omitted if microphone is mounted on PCB.
Electrolytic Capacitor		



Table 2 – Components for Microphone Input Circuit for NLP-5x

Resistors		
<b>R1</b>	270, 5%	Recommended value – Should be between 270 and 2000 ohms. Along with C12 & C13 provides power filtering – Actual values of C12 and R1 must be selected to give 60 millisecond RC time.
<b>R2, R3</b>	5.6K, 5%	Critical value. R2 is optional*.
<b>R4, R6</b>	2.2K, 5%	Actual value must match the microphone sensitivity rating. Please refer to Sensory Design Note 80-0318 (“Selecting a Microphone”). R4 is optional*.
<b>R5, R7</b>	510, 5%	ESD protection – may be omitted if microphone is mounted on PCB. R5 is optional*.
Ceramic Capacitors		
<b>C1, C2</b>	100uF (.1nF)	Critical value. C1 is optional*.
<b>C3</b>	.1uF	Decoupling between AVdd and Avss.
<b>C4</b>	.01uF	Critical value.
<b>C5</b>	.1uF	Critical value.
<b>C7</b>	.01uF	Critical value.
<b>C9, C11</b>	220uF (2.2nF)	Critical value. C9 is optional*.
<b>C13</b>	.1uF	Critical value.
<b>C14, C15</b>	.22uF	Critical value. C14 is optional*.
<b>C16, C17</b>	4700pF (4.7nF)	ESD protection – may be omitted if microphone is mounted on PCB. C16 is optional*.
Electrolytic Capacitor		
<b>C12</b>	220uF	Recommended value – should be between 30 and 220 uF – Select the largest practical value given the size/cost constraints of the application.
<b>C6</b>	2.2uF	Critical value.
<b>C8, C10</b>	1.0uF	Critical value. C8 is optional*.
Diodes		
<b>D1-D4</b>	1N4148	ESD protection – may be omitted if microphone is mounted on PCB. D1 and D2 are optional*.

\* Optional components may be omitted if only one microphone input is used.

### B. PWM and Oscillator Power Isolation Networks

1. The diagram in Figure 6 shows recommended “star” power and ground hookups and reference isolation networks for PWM and oscillator power supplies. Isolation on PWM and oscillator power supplies, while not mandatory, is recommended to prevent cross-coupling noise between the various power supplies.

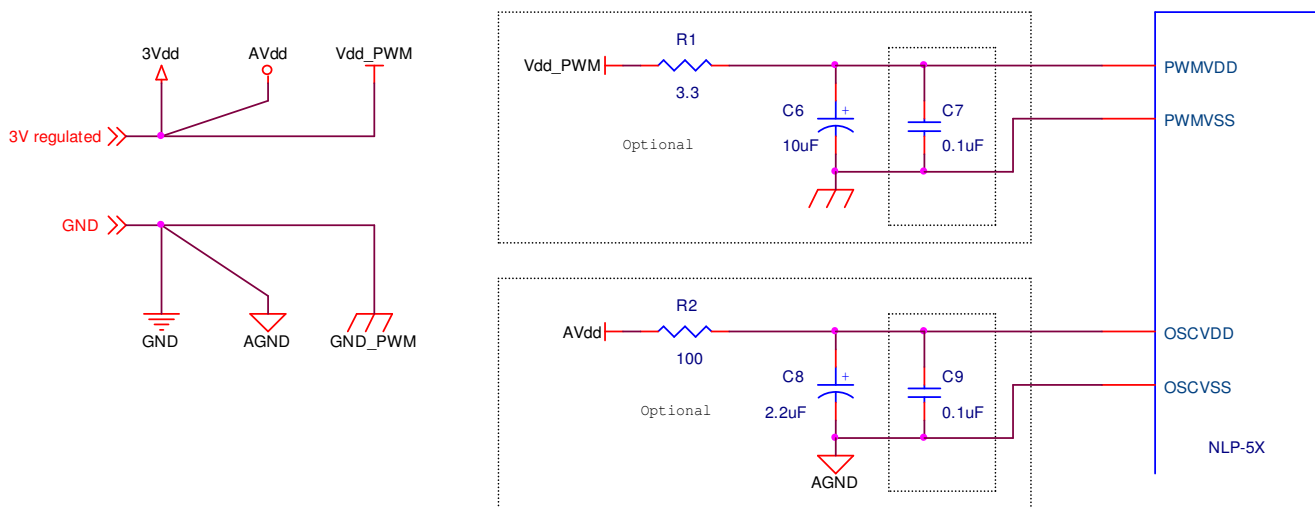


Figure 6 - Power Isolation Networks

## IV. PCB Design

A. **General Guidelines** - The following is a list of principles that if followed, will help ensure the best possible voice recognition accuracy.

1. **Separation of Digital and Analog** - The circuit board should be physically organized to allow the largest practical separation of the analog from the digital portions of the system. All of the grounds shown above in figure 4 are analog. They should be kept separate from all other grounds. The analog and digital grounds should only meet up at the main ground source (usually a battery or regulator). This technique is commonly referred to a “star grounding”. Digital traces can be made around the power traces and the data line traces, but should avoid analog traces.
2. **Short Microphone Path / Shielded Cable** – While it is a good practice to keep all analog PCB traces as short as possible, in particular, the main audio signal path starting at the positive side of the microphone and going into the RSC\_4x MICIN1 and MICIN2 pins - or the NLP\_5x MICP1 and MICP2 pins - **MUST be kept as short as practical!** The audio signal is a high impedance signal with very low amplitude of only a few millivolts peak-to-peak. It is critical to keep this line short to avoid an antenna effect from digital noise and EMI if good recognition accuracy is to be obtained. If the microphone is mounted externally, then shielded cable must be used for hookup to the PCB.
3. **Ground Plane** - If a multi-layer PCB can be used, then create a separate ground plane. Divide the ground plane into digital and analog sections which meet up only at the main ground source.
4. **Copper Pour Grounding** – If a ground plane is not practical or possible, then use copper pour grounding on the PCB. Divide the copper pour into digital and analog sections which meet up at only the main ground source. Do not leave any isolated copper pour “islands” unconnected to anything else as they negatively affect recognition accuracy. If practical, use copper pour grounding on both sides of a double sided PCB. Sprinkle plenty of feed through vias to make sure there is no difference in voltage potential between top and bottom planes.
5. **Decoupling Capacitors** - A 0.1  $\mu\text{F}$  bypass capacitor should be installed immediately next to each digital IC and near the Vdd pins of the RSC\_4x and NLP-5x chip. The bypass capacitors should be a monolithic ceramic type, rated at 50 volts. If a three-terminal voltage regulator (such as a LM2930-3.3V) is used, Aluminum electrolytic or Tantalum capacitors should be connected close to the regulator between the output pin and ground, and also the input pin and ground.
6. **Data and Address Bus** - The data lines should be routed for the minimal length, and be as wide as practical to reduce impedance. The bi-directional nature of data lines causes large instantaneous currents to get switched around. With inductive loading, these currents can cause data lines to ring and generate large switching spikes. Although these spikes and the ringing will have dampened when the data is sampled, they could affect the operation of other devices connected to the same bus. The address, strobe and control signals are outputs of the RSC\_4x and NLP\_5x and have a nominal capacitance; their trace widths and length aren’t critical, with a preferred width of more than 0.012”.
7. **ESD Mitigation Techniques** –There are a number of additional recommendations for ESD mitigation. Please refer to Sensory Design Note 80-0276 (“ESD Mitigation Techniques”).

## B. Locating and Mounting Interactive Speech ICs

1. **RSC-4x Package Type** - Some Interactive Speech ICs are available in packaged form or as bare die. Bare die may be wire bonded directly to the main PCB or, in some cases, may be bonded to a separate chip-on-board (COB) circuit board.
2. **Using COB** - In production this COB assembly may be functionally tested, then attached to the main board as a working module. There are several methods of attaching the COB to the main board, and a careful

choice should be made by the designer. Since cost is always a consideration, COB boards may often be designed as single-sided PCBs.

The simplest way of attaching a single-sided COB to a main board is to lay it, chip side up, on the main board and make solder bridges from the main board up along the thickness of the COB to the electrical contacts on the top of the COB board. However, in production this is not a reliable technique.

A second technique is to use wires or pins to connect thru-holes between the main board and the COB. This reliable technique may be more time consuming.

A third technique is to put a hole in the main board at the center of the COB location and to mount the COB to the main board UPSIDE DOWN, that is, with the chip facing down into the hole. Then the COB and main boards may be soldered together. In this case, the orientation of the signal leads of the RSC-4x is different from that of the other arrangements.

A fourth technique is to put a slot in the main board that is the same width as the thickness as the COB, and the same length as the COB. The COB can then be mounted perpendicular to the main board, and the pads soldered together. This arrangement offers dimensional advantages in certain applications.

3. **Using QFP Packaged Parts** - All RSC-4x and NLP-5x parts in QFP packages are shipped in dry-packed trays, and meet JEDEC level-3 moisture sensitivity standards. Level-3 specifies that once the bag has been opened, the devices should be board mounted within 168 hours (1 week). If not, they must be re-baked at 125C for 24 hours. Profiling of the IR-Reflow ovens should be carefully monitored, with settings established in accordance with JEDEC standards.

## V. Software Design

### A. Configuration File Parameters

Every FluentChip application includes a configuration file called config.mca which defines various program configuration parameters. One parameter in particular has repercussions for hardware design.

1. **MIC\_DISTANCE** – This parameter controls the AGC range used by the internal chip hardware for voice recognition functions. The program should select the most appropriate setting for the application. Three settings are provided:

Setting	Amp. Gain	Initial AGC	AGC Range
HEADSET	Reduced	Low	Low End Range
ARMS_LENGTH	Normal	Medium	Full Range
FAR_MIC	Maximum	High	High End Range

## VI. Circuit Verification Procedure

Three methods of verifying a circuit are described: Karaoke, Diagnostic Output, and Triangle Wave.

### A. Karaoke Test (RSC-4x only)

1. **Description** – Karaoke is a quick but subjective “How good does it sound?” test.
2. **Usage** - Karaoke is simple to use. Just add 1 line of code to your application:  
Karaoke #N ; where N is the number of seconds 1-255  
This will place the program into a special mode where the microphone input will internally be connected to the DAC and PWM output. Speak into the microphone and listen to the output. Now listen to the output when not speaking. Try disconnecting the microphone. If any noise is heard from the speaker then there may be a problem in the circuit.

## B. Diagnostic Output Test (RSC-4x and NLP-5x)

- Description** – Diagnostic output is designed to give real-time diagnostics on how well the RSC-4x or NLP-5x are performing speech recognition. Each time the program calls any training or recognition function or macro, diagnostic information can be automatically transmitted serially on a dedicated I/O pin.

For full details, refer to the Sensory FluentChip\_Reference.chm file in the Programmer's Guide:Diagnostic Output (TxDiag) section. This file is normally located in the \docs\ folder of the FluentChip installation.

**WARNING:** Sensory STRONGLY recommends that ALL designs using the RSC-4x and NLP-5x have the capability for diagnostic output. If the final design has any serious problems with recognition the diagnostic output will be crucial to an accurate diagnosis and solution to the problem. The diagnostic output does not have to be enabled on every production board, but it must be present in case it is needed.

- Usage (RSC-4x)** – On the RSC-4x chip, two I/O pins are nominally required for diagnostic output, one for the actual serial output, and one for the diagnostic enable. Diagnostic output can be enabled by the DiagEnable pin. When this pin is asserted low during power on or reset, then the TxDiag pin will be used for diagnostic output. The developer can also programmatically enable or disable diagnostic output by means of the DiagOn and DiagOff macros/ functions.

Note that diagnostic output is at TTL levels (0-3V) and has opposite polarity to the RS232 standard. It must therefore be inverted and passed through a level shifter before sending it to a PC running a terminal program. Three suggested RS232 interface circuits are provided here:

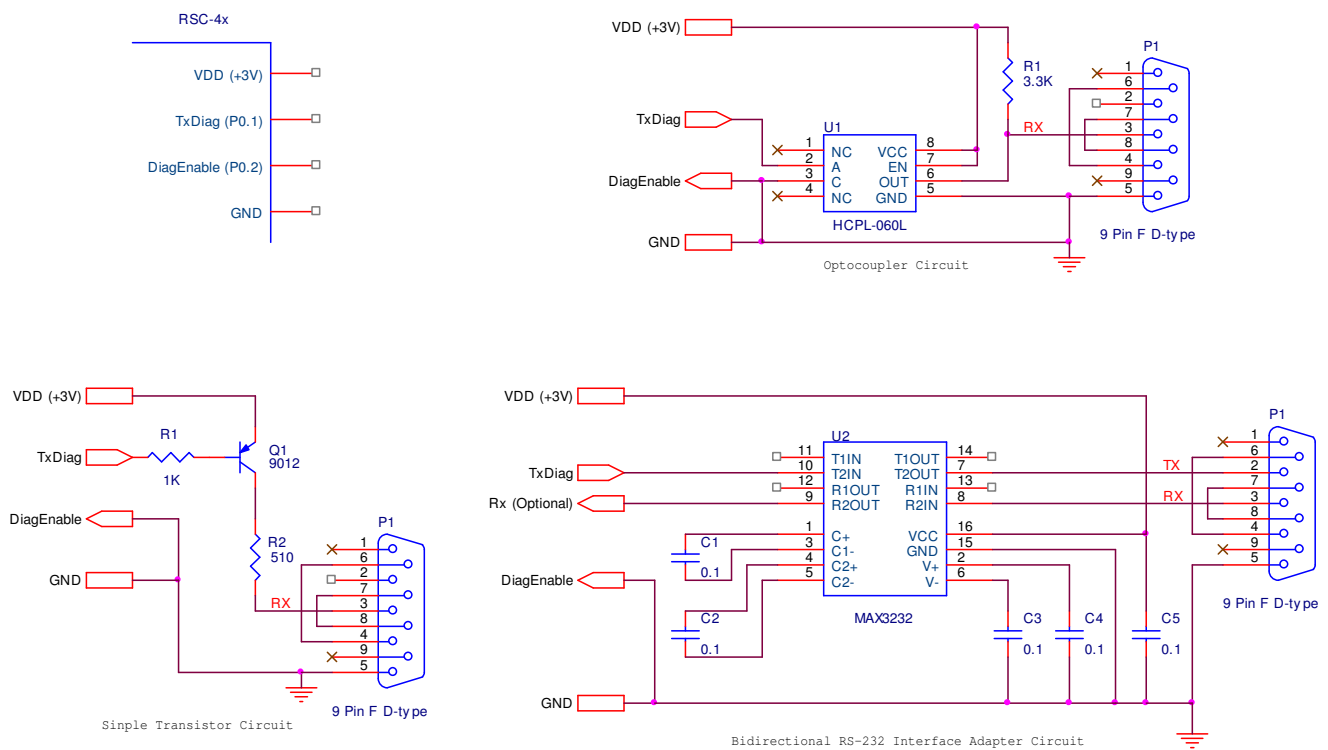


Figure 7 - Three Sample RSC-4x Diagnostic Output PC Interface Circuits

- Usage (NLP)** - On the NLP-5x chip, a nine pin header is required for diagnostic output, refer to the figure 7 below for a description of the header pinouts. In addition, a Diagnostic Output Interface Cable and Sensory Program/Verification board is required. On the NLP-5x, Diagnostic Output is enabled by –TRST. When this pin is asserted low at any time, then the NLP-5x will output JTAG messages for diagnostic output.

Note that it is possible to remap diagnostic output on the NLP-5x to use GPIO pins in a similar fashion to the RSC-4x. To do this, the developer needs to change the `_printf()` function to use the UART software driver and add a an RS-232 interface circuits as described above:

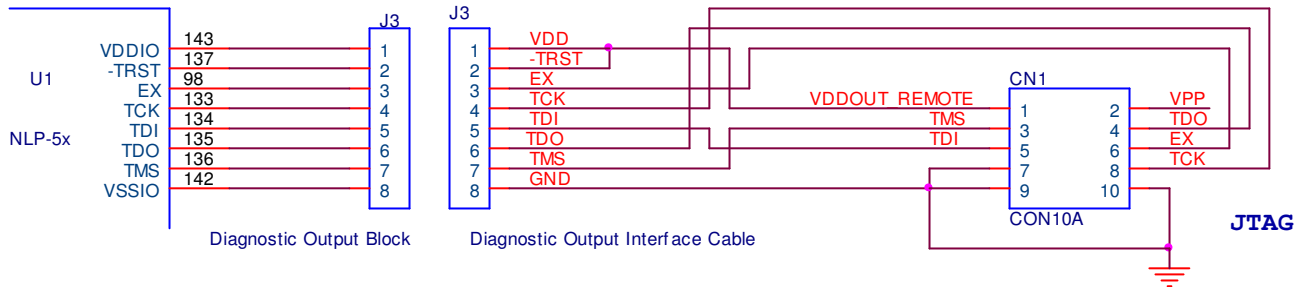


Figure 8 - NLP JTAG Diagnostic Output Interface Circuit

### C. Triangle Wave AFE Test (RSC-4x only)

1. **Description** – This test measures the RSC-4x’s Analog Front End (AFE), including most of the microphone circuit, for distortion and noise. For this test, a 1KHz, 2V p-p triangle wave is attenuated down to 20mV p-p and is then injected directly into the microphone input. Then a special test routine is invoked which measures the amplitude of the signal source triangle wave’s fundamental, 3rd and 5th harmonics (1KHz, 3KHz, and 5KHz). It also measures the amplitude of signals outside those frequencies. If all measured signal levels are within predefined limits, the test passes.
2. **Requirements** – You will need the following:
  - a) The RSC-4x to be tested with recommended microphone circuitry as show in Figure. 4,
  - b) A 1KHz, 2V p-p, <50 ohm impedance triangle wave source with accurate and stable frequency, amplitude, and symmetry,
  - c) One 1000 ohm, 1% resistor, one 10 ohm, 1% resistor, one 1 uF capacitor.
  - d) The AFETEST object code, available from the “downloads” page at [www.sensoryinc.com](http://www.sensoryinc.com).
3. **Usage** – Complete the following steps:
  - a) Connect the wave source to the RSC-4x as show in Figure 5 (don’t forget to remove R2 for this test),
  - b) Add AFESUBS.MCO to your Phytion project,
  - c) In the main program, issue a call to AFETest, which will return “0” if pass and “1” if fail.

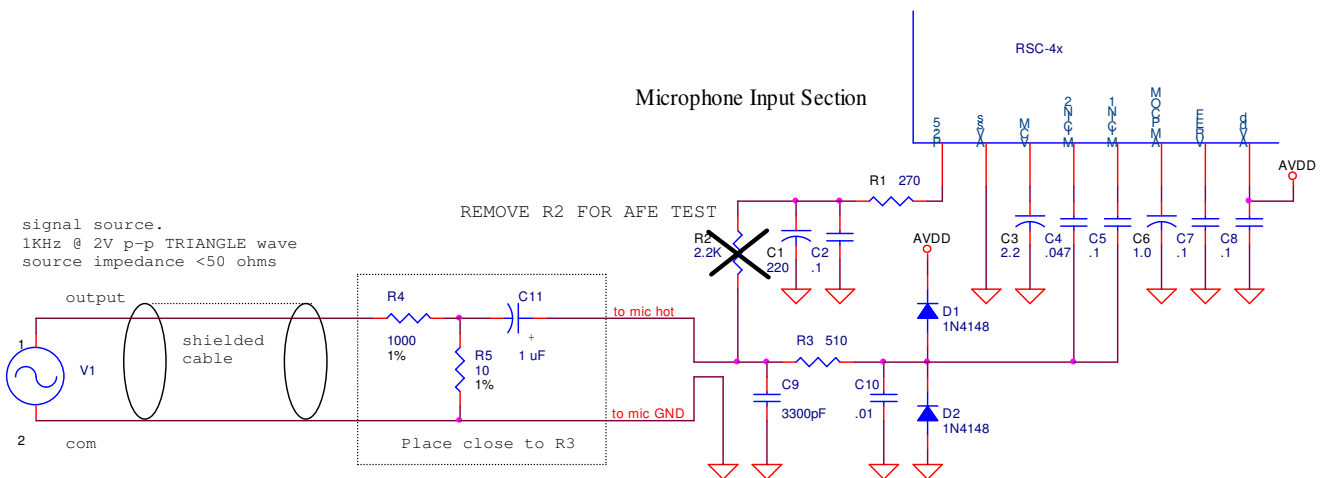


Figure 9 – Triangle Wave AFE test setup

## VII. Sensory's Free Design Review Service

Sensory offers a free design review service for virtually any product that incorporates the RSC-4x or NLP-5x microprocessors. At each stage of the review process, Sensory's technical support staff will work with you to help ensure the success of your design.

If the developer must maintain product security, then Sensory is willing in most cases to sign confidentiality or mutual non-disclosure agreements with the developer. Sometimes full details of a product cannot be released even under a written agreement. In these cases, Sensory can review as much of the product as possible, but strongly recommends that all non-proprietary aspects of the product be released for full review to help ensure the best possible design.

Since product development is an ongoing process, the design review must also be a process. For ease of description, the process has been broken down into three stages, called R1, R2 and R3.

### A. R1 Review

1. **Purpose** – The purpose of an R1 review is to examine and review each part of the product at various stages of development, including initial concepts and specifications, block diagrams, program flowcharts, recognition words and speech and music output. As the product progresses along its development path, completed components of the design including schematics, PCB layouts, BOM, and source code can be reviewed when made available to Sensory.
2. **Things required for an R1 review** –
  - a) An initial written product concept,
  - b) A system specification, including block diagrams and proposed flowcharts,
  - c) A list of proposed recognition words that will be used in the product,
  - d) A list of proposed output speech and music.
  - e) Data sheets and specifications for any proposed special subsystems (LCD, motors, etc),
  - f) Schematics, PCB layout files and BOM (when made available)
  - g) Source code (when made available)
3. **Sensory will** –
  - a) Review the concept, specifications flowcharts and check for things that may be difficult or impractical for the RSC-4x or NLP-5x to handle,
  - b) Review the recognition words and suggest alternates if any words are found which may be difficult for the RSC-4x or NLP-5x to recognize. Also suggest organizing the recognition words into subsets to maximize overall recognition accuracy,
  - c) Provide an estimate of the memory requirements of the product.
  - d) Review the schematics, PCB files, BOM and source code, especially those sections that deal with voice recognition and speech output.

### B. R2 Review

1. **Purpose** – The purpose of an R2 review is to thoroughly examine one or more completed **pre-production** units **prior** to mass production. This is a one-time, thoroughly extensive review of the product.
2. **Things required for an R2 review** –
  - a) A finalized system specification, including full schematics, gerber files, BOMs, etc.,
  - b) Full program flowcharts and source code if available (if not, then all non-proprietary flowcharts and source code snippets relating to Sensory voice recognition and other Sensory technologies are required),
  - c) At least one working prototype (more may be better),
  - d) User instruction manuals, quick start guides, etc.
3. **Sensory will** –
  - a) Do a complete review of the product including hardware, software, plastics and function,
  - b) Offer and suggestions to improve the product's function, reliability or ease of use.

**C. R3 Review**

1. **Purpose** – The purpose of an R3 review is to examine a fully released “ready for sale” production unit, preferably one that has been chosen at random from the production line, to ensure that no changes which could affect the function have been made to the product or manufacturing processes since the R2 review.
2. **Things required for an R3 review** –
  - a) At least one production unit chosen at random (more may be better).
3. **Sensory will** –
  - a) Compare the production unit(s) to the notes and unit(s) from the R2 review,
  - b) Make sure that all accepted recommendations from the R2 review have been correctly followed,
  - c) Inform the developer of any unauthorized changes from the R2 review,
  - d) Notify the developer of any problems encountered or other noteworthy comments.

## The Interactive Speech™ Product Line

Sensory's **Interactive Speech™** product line makes consumer electronics more intelligent by enabling them to talk and hear with speech synthesis, voice recognition, and other advanced audio and interactive technologies. It is designed for integration into cost-sensitive consumer electronic applications such as home electronics, smart toys, music players and personal communication devices. The hardware line includes the award-winning RSC-4x family of mixed signal processors, the *VR Stamp™* 40-pin DIP module, and the SC-691 music and speech synthesis slave processor. Embedded software options include our *FluentSoft™* Recognizer, which enables speech recognition on non-Sensory processors and DSPs. Sensory's *BlueGenie™* Voice User Interface, the first Voice Recognition and Synthesis option for Bluetooth enabled devices, offers user friendly control of headsets, music players and other BT devices requiring hands-free operation.

### 1. RSC Microcontrollers and Tools

The RSC product family contains low-cost 8-bit speech-optimized microcontrollers that are fully integrated and include A/D, pre-amplifier, D/A, RAM, and ROM circuitry. With Sensory's *FluentChip™* firmware, the RSC family offers speech recognition, speaker verification, speech and music synthesis, voice recording and playback, and an entire suite of interactive robotic and sonic networking technologies. The family is supported by a complete suite of evaluation and development toolkits that include the ability to quickly create speaker independent recognition sets in many languages.

### 2. Speech Recognition Modules and Tools

The *VR Stamp™* is a complete speech recognition module based on the RSC-4x and is ideal for fast design and easy production. A low-noise audio channel and standardized 40-pin DIP footprint allow rapid prototyping, less debugging, and shorter time to market. The *VR Stamp Toolkit* includes everything needed to get started today, including *VR Stamps*, *Module Programming Board*, sample applications, and a complete set of development tools featuring the *Phyton IDE* and limited-life *C compiler*, *QuickSynthesis™ 4* and *Quick T2SI-Lite™* speech tools.

### 3. SC6 Slave Processor and Tools

The SC-691 is a standard slave synthesizer that accepts compressed speech data from other microprocessors or microcontrollers and converts it to speech. The chip operates up to 12.32 MIPS, and provides high-quality, low data-rate speech compression and MIDI music synthesis, with unlimited speech duration using external memory. Sensory offers hardware and software tools for analyzing speech files, editing speech data and generating coded speech.

### 4. FluentSoft™ Recognizer

The *FluentSoft™* Recognizer is the engine powering the *FluentSoft™* SDK. It provides a noise-robust, large-vocabulary, speaker-independent solution with continuous digit recognition and word-spotting capabilities. This small-footprint software recognizes thousands of words and runs on non-Sensory processors including Intel XScale, TI OMAP, and ARM9, and supports operating systems such as MS Windows, Linux, and Symbian.

### 5. BlueGenie™ Voice User Interface

The *BlueGenie* Voice Interface software suite runs on CSR's BC-5 MM Kalimba DSP, and enables manufacturers of Bluetooth products to integrate full voice control and synthetic speech output without the need for visual displays or complex user interfacing. It frees designers to pack functionality onto small form factor Bluetooth devices and answers consumer demand for a truly hands-free experience.

### 6. Important notices:

Sensory Incorporated (Sensory, Inc.) reserves the right to make changes, without notice, including circuits, standard cells, and/or software, described or contained herein in order to improve design and/or performance. Sensory, Inc. assumes no responsibility nor liability for the use of any of these products, conveys no license or title under any patent, copyright, or mask-work right to these products, and makes no representations or warranties that these products are free from patent, copyright, or mask-work right infringement, unless otherwise specified. Applications that are described herein for any of these products are for illustrative purposes only. Sensory, Inc. makes no representation or warranty that such applications will be suitable for the specified use without further testing or modification.



S E N S O R Y®

575 N. Pastoria Ave., Sunnyvale, CA 94085  
Tel: (408) 625-3300 Fax: (408) 625-3350

© 2006-2010 SENSORY, INC. ALL RIGHTS RESERVED.

Sensory is registered by the U.S. Patent and Trademark Office.

All other trademarks or registered trademarks are the property of their respective owners.