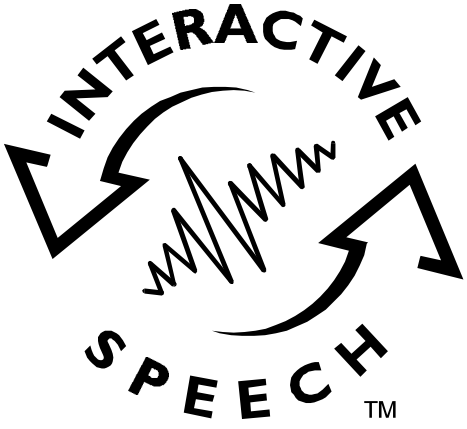
Voice Direct
Data Book



SENSORY

P/N 80-0037-3

Technical Support
For product support and questions:

Marketing Communications
521 East Weddell Drive
Sunnyvale, CA 94089-2164
Tel 408-744-9000

Technical Support
521 East Weddell Drive
Sunnyvale, CA 94089-2164
Tel 408-744-9000
techsupport@sensoryinc.com

Table of Contents

List of Figures

List of Tables

Voice Direct

Introduction

Sensory's Voice Direct delivers state-of-the-art speech recognition technology in a cost-effective, flexible IC format that is ideally suited for a variety of consumer products. Voice Direct is quickly and easily embedded into existing products or new designs because it can operate in Stand Alone Mode or as a slave IC to an external host processor. This flexibility makes Sensory's Voice Direct the value-added choice for countless applications. With a minimum of external memory and hardware, you can incorporate Voice Direct into electrical switches, desktop appliances, consumer electronics, and practically any consumer product that can be controlled using voice-activated technology.

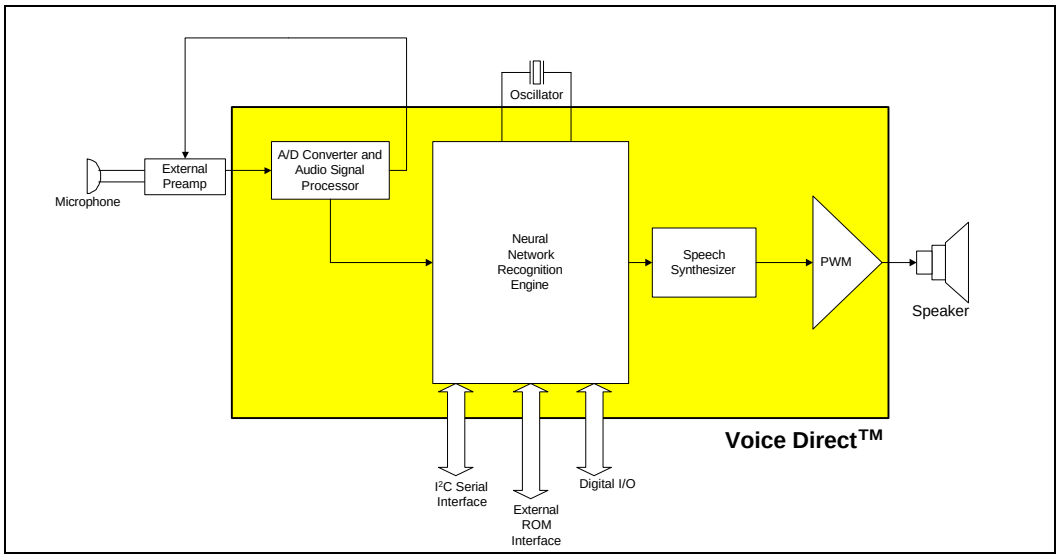
This data book describes the end-user voice recognition features, training approach, and command set, as well as the design, interface, and protocol parameters needed to implement Voice Direct technology.

> *For Voice Direct IC schematics, see page 39; for Voice Direct Module or Evaluation Kit schematics, go to page 69.*

Feature Overview

Voice Direct is a speaker-dependent speech recognition IC. It maps spoken commands to system control functions using sophisticated speech recognition technology. Using easy-to-learn, intuitive techniques, end users train Voice Direct, which performs speech and sound prompting and speaker-dependent (SD) speech recognition. Voice Direct can be controlled by an external host processor or operate in a pin-configurable Stand Alone Mode. Voice Direct comes in three product packages: the Voice Direct IC, the Voice Direct IC Module, or the Voice Direct Evaluation Kit.

Figure 1 - Block Diagram



As shown in Figure 1 - Block Diagram, Voice Direct incorporates analog-to-digital signal conversion, analog control signals, fully-automated speech prompting (neural network recognition engine and speech/sound generator), speech synthesis and PWM output, as well as serial EEPROM, external ROM and digital I/O interfaces. Voice Direct employs a sophisticated neural network to recognize trained words or phrases with greater than 99% accuracy. Among Voice Direct's capabilities are:

- Speaker-dependent recognition works in any language (support for language localization and custom prompts)
- Full support for speech and sound prompts
- Recognition of up to 60 words/phrases
- Phrase recognition up to 3.2 seconds
- Minimal memory: less than 100 bytes of external memory required per trained word
- Two operating modes: external host-controlled or pin-configurable stand alone
- User-friendly speech prompting
- A full command set for external host-driven programs

The Voice Direct IC's highly integrated design reduces the external parts count, so you can configure a complete recognition system with little more than a battery, speaker, external memory, microphone, crystal, and audio input circuitry. The Voice Direct Module comes with memory, crystal and audio input circuitry. Another option available is the Voice Direct Evaluation Kit -- a complete modular system that includes host interface and stand alone boards as well as all required external components.

Voice Direct records and accesses trained speech in voice recognition templates. It has private access to these voice recognition templates via external EEPROM memory (8 Kbytes). You can also configure an optional language/extended speech ROM.

Sensory's Voice Direct is designed to operate as a stand alone device or as a slave chip. As a slave chip, it receives and processes commands from a host controller or Master CPU (MCPU) and returns status information and data. Communication with the MCPU is through a 3-wire serial bus. Voice Direct uses a rich command set, so application programmers can implement complex voice recognition functions with a minimum of MCPU overhead. This allows the application software to focus on providing an intuitive and efficient user interface.

Implementation Overview

Implementing Voice Direct technology involves selecting external host or Stand Alone Mode and then:

- External Host Mode: Writing application control software for a host controller (MCPU) and physically embedding a Voice Direct IC into your product.

- Stand Alone Mode: Configuring Voice Direct pins to switches that you use to implement your program and physically embedding a Voice Direct IC into your product.

What follows is an overview of Voice Direct concepts and capabilities, its command set and serial interface, and its hardware requirements and features. Subsequent sections discuss these topics in detail.

Voice Direct maintains a template directory with the following information for each entry:

- A speaker-dependent speech template – a digitally constructed pattern of a spoken name
- An attribute byte – to identify the kind or type of entry during searches

Users can execute control functions associated with a name entry by voice recognition. One or more users can organize multiple directories or word groups – these capabilities make use of the attribute byte. The attribute byte is combined with a search mask during recognition and directory functions, which makes it possible to more effectively scan templates assigned by multiple users, searching by categories.

External Host Mode

In External Host Mode, all operations are controlled by the MCPU through the Master/Slave bus. You can program Voice Direct to provide system control functions. An extensive, easy-to-implement command set helps you write the MCPU application control software. Your program can incorporate speech from a long list of pre-recorded prompts to creatively guide the end user through the user interface. Commands are conveyed in data packets via the 3-wire serial interface that connects the host processor to Voice Direct.

In External Host Mode, Voice Direct requires minimal external circuitry (see the schematics in the appendix). The following components are required unless otherwise noted:

- Microphone
- Preamplifier
- External Serial EEPROM Memory Microchip 24C65 or compatible (for external template storage memory)
- Oscillator
- Speaker
- Speaker Amplifier (optional)
- External ROM for customized speech or foreign language (optional)

Stand Alone Mode

In Stand Alone Mode, the functional capability of the Voice Direct IC is determined by the configuration of a number of I/O pins. The pin inputs (Train and Recognize) generate specific chip outputs and actions. You can build a complete recognition system using only the Voice Direct IC and the following:

- External Serial EEPROM Memory Microchip 24C65 or compatible (for external template storage memory)
- Preamplifier
- Oscillator
- Passive components:
- Speaker
- Microphone

Voice Direct Product Line

A full line of Sensory Voice Direct products is available to help you develop the application:

- *The Voice Direct IC* - The Voice Direct integrated circuit without the external components described above.
- *The Voice Direct Module* - A pre-configured IC for prototype development. The Voice Direct Module comes with all of the required external components except for a microphone and speaker.
- *The Voice Direct Evaluation Kit* - A fully configured Voice Direct Host Interface Board (for external host-driven applications) and the Stand Alone Evaluation Board. Both boards incorporate the Voice Direct Module, speaker, microphone, and battery pack. The kit also includes Sensory's graphical VoiceHost software to evaluate, demonstrate, and train Voice Direct using the Host Interface Board. The kit includes three Voice Direct Modules and a parallel port card.

Voice Direct Concepts and Functional Capabilities

The following sections describe core concepts and capabilities that are useful in creating the application control program that controls Voice Direct capabilities. Differences between External Host Mode and Stand Alone Mode are noted. In the descriptions that follow, "user" refers to the end user of your application control program. The following topics are covered:

- Speech Recognition
- Voice Direct Entry Structure
  - Speech Templates
  - Voice Recordings
  - Attribute Bytes
- Masks
- Current Pointer
- Mandatory Words and Phrases
- Optional Words and Phrases
- Development Parameters
- External Host Mode
- Stand Alone Mode

Speech Recognition

Voice Direct performs speaker-dependent discrete word recognition by comparing a pattern that it generates in real time with previously-trained word templates. The pattern Voice Direct generates is based on a digital reconstruction of the voice command.

Each word to be recognized must first be *trained.* During training, Voice Direct builds up a *template* representing the individual speaker's unique sound pattern for each specific word or phrase to be recognized. Templates are stored in serial EEPROM memory. During recognition a new pattern is produced and compared to the stored templates to determine which word was spoken.

Voice Direct features integrated speech-prompting for both training and recognition operations, allowing the development of sophisticated interactive products with little or no programming. As always, the focus of Voice Direct control programs should be simplicity and ease-of-use.

In both stand alone and External Host Modes, Voice Direct performs the following operations when recognizing a word:

1. The audio signal (spoken word) is externally amplified and filtered and then supplied to the analog inputs of the Voice Direct, which converts the analog waveforms to digital samples.

2. Voice Direct analyzes the speech signal samples and generates a pattern of information representing significant speech elements.

3. Voice Direct increases or decreases the gain of the external amplifier as needed to maintain signal quality.

4. Using a neural network, the pattern is compared with previously stored template patterns; a small number of candidate templates is selected.

5. The candidate templates are further processed to determine the one template that provides the best match to the unknown pattern.

6. If the best match template gives a score above a pre-defined threshold, Voice Direct chooses the word associated with that template. If no template provides a match above threshold, a special "no match" value is chosen.

Steps 1-3 above are repeated for each word during training. Voice Direct stores the average of two training patterns for each word to improve accuracy. Before storing a new template it is compared with the existing templates in the set. The new candidate word will not be accepted if it is too similar to an existing word (for example, "Bill Smith" and "Jill Smith").

Voice Direct, like other speech recognition systems, is necessarily subject to two types of errors: *rejects* (failure to recognize a word in the vocabulary) and *substitutions* (confusion of two vocabulary words, or recognition of a non-vocabulary word). The relative importance of each type of error may depend on the application.

Voice Direct provides selectivity levels that allow the user to optimize the tradeoff between these two types of recognition errors. When the recognition selectivity level is set to its highest value, Voice Direct minimizes substitution errors but may produce more rejection errors. When the training selectivity level is set to its highest value, Voice Direct minimizes both substitution and rejection errors by rejecting words that are too similar sounding, potentially increasing recognition accuracy. See Developer Parameters (p. 15).

In External Host Mode the recognition selectivity level is set to using the "Store Parameter" (41h) command. In Stand Alone Mode , a subset of these parameters is available for selection via resistor programming. See Stand Alone Mode (p. 16).

Voice Direct Entry Structure

During training, Voice Direct creates and stores entries in the serial EEPROM. Entries have the following elements:

- A speaker-dependent speech template of the name
- An attribute byte

These are described below.

Speaker-Dependent Speech Templates

Within the context of the product application, a user must train Voice Direct to recognize each word in the directory. During this training process, speaker-dependent speech templates are created and then stored for comparison during the recognition process.

Training is simple. It consists of speaking a word and then repeating it for confirmation. Each time the word is spoken a template is created. The two templates thus created during the training process must closely match (the speech patterns should be similar). If the two templates match sufficiently, then a third template composed of the average of the first two templates is created. This third average template is then stored in external memory and used during the recognition process.

During the voice recognition process, a fresh template of the spoken word is produced. This new template is then compared to the stored templates to determine which word was spoken.

Attribute Bytes

When an entry is created, an attribute byte is associated with the voice template. Attribute bytes allow users to distinguish between different groups or kinds of entries.

The attribute byte will always have one bit position set to one, while the rest remain at 0. This allows for differentiating up to eight different kinds of templates.

This method of dividing trained words into categories makes the most efficient use of template memory because Voice Direct dynamically allocates memory rather than using fixed blocks. Consider an application that supports two different users. Rather than dividing the memory in half, it can be allocated according to actual usage. The user with more words to recognize can train templates as needed, while the other user fills up the remainder. This avoids leaving empty memory slots in one user's directory, while the second directory is needlessly and inconveniently restricted to fewer slots.

Masks

During searches, the MCPU control program (External Host Mode) or Voice Direct (Stand Alone Mode) generates a mask. The mask is logically 'and'ed against each attribute byte. Masks support flexible applications with multiple directories or word lists. In fact, masks help such applications perform more responsively by reducing search times. During recognition and directory functions, a mask is constructed to select and/or exclude words in specific groups.

> *During the training process, the mask is used to determine which existing word groups should be included when testing for similar words. The mask also has an effect on how other directory commands such as "Increment Current Pointer," "Decrement Current pointer," and "Query Lexicon Status" will function. For more information, see the Command Summary.*

Current Pointer

The Current Pointer points to the current entry of a directory and is used to select and operate on entries within a directory. The entry selected by the Current Pointer is referred to as the Current Entry. Many of Voice Direct commands operate on the Current Entry.

Standard Words and Phrases

Voice Direct has 32K bytes of general-purpose Standard English words and phrases onboard. Two word lists, a Mandatory Prompt List and an Optional Prompt List, are available. Each list has its own set of indices.

The Mandatory Prompt List contains speech that is accessed directly by the chip during the training and recognition process. Speech in this list is also accessed to issue specific error messages (such as "Please talk louder"), thereby providing user feedback during training and voice recognition. Words and phrases in this list can be also accessed via the "Say A Prompt" command.

The Optional Prompt List contains speech that can only be accessed via the “Say A Prompt”, Train, or Recognize Commands. These are standard prompts useful in creating the user interface for applications. Examples include prompting to review operations, and to play error messages (such as “memory full” or “memory empty”) at the application level.

Custom Words and Phrases

You can replace the onboard standard word list with a customized word list for English or foreign languages via an external ROM. This external ROM can contain either 32K or 64K of speech synthesis. The only limitation on the external synthesis is that it must be divided into two lists -- the Mandatory Prompt and Optional Prompt Lists described above. The Mandatory Prompt List in the external ROM must contain the same number of indices as that of the onboard Mandatory Prompt List, and the synthesis at each index must have a meaning (in the new language) identical to that of the onboard Mandatory Prompt List (in standard English). This discipline is required because the Mandatory Prompt List indices reflect functionality that is hard-coded into Voice Direct. For example, sentence 61 of the Mandatory Prompt List Table (“Say word one” -- see below) could be replaced with a custom voice synthesis equivalent in a different language, or with an extended prompt such as “Please say the name of the first person to be added to the directory.”

Since the Optional Prompt List is accessed solely by the application control code, there are no requirements to either the number of indices, or the semantic meaning of any index. Your are free to create a Optional Prompt List of any desired synthesis required by your application – you are limited only by the available memory. **For more information on creating custom Prompt Lists, contact Sensory.**

> *A rule of thumb: about 800 bytes of memory per word are required for good quality synthesis, and about 400 bytes per word for acceptable quality synthesis. This rule of thumb varies by language.*

Developer Parameters

The developer parameters let you control and configure your Voice Direct application. You can select settings that anticipate the proximity of users to the microphone, the kind of audio output (DAC and/or PWM) to select and so on. You set these parameters using the Store Parameter Command (see this command in the Command Set Overview p. 21). See Developer Parameters (p. 30).

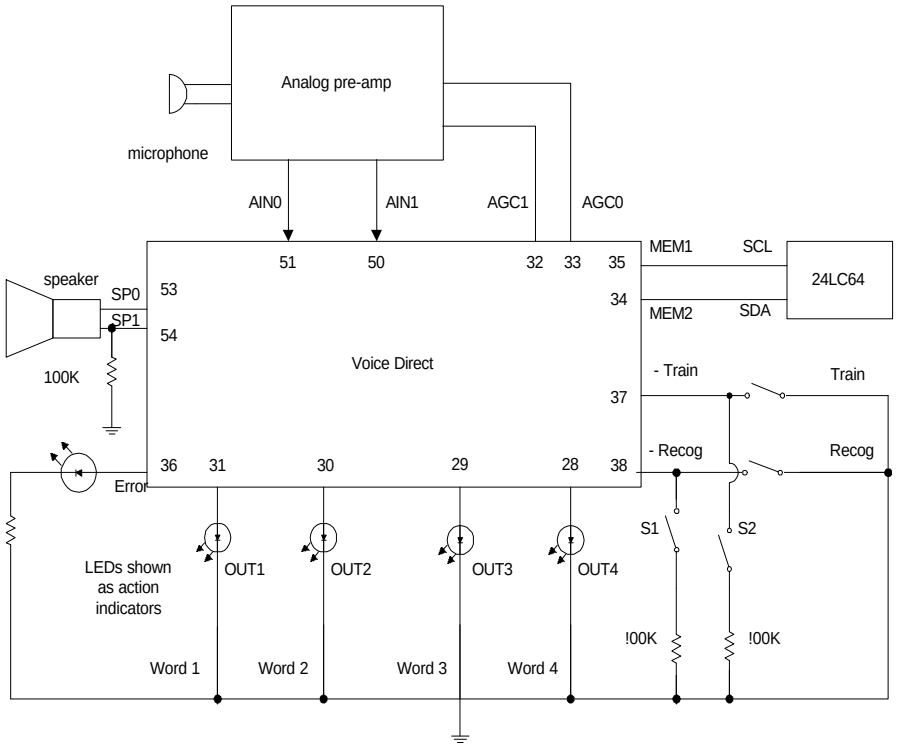
Stand Alone Mode

If the MODE signal is resistively pulled to Ground, the Voice Direct will operate in the Stand Alone Mode. This allows the chip to work in an environment without any external host elements. Pin inputs will generate specific chip outputs and actions. The functionality is fixed and determined by the configuration of several I/O pins.

Stand alone operating mode is designed to provide a complete recognition system using only the Voice Direct, external template storage memory, an audio preamplifier, and passive components. Two switches must be connected to the TRAIN and RECOG pins to control operation. The descriptions below describe this configuration and refer to these switches as the RECOG and TRAIN switches. Decoded outputs provide direct action controls for small-vocabulary implementations. Other passive components can be used to configure the chip for operation as described below. Figure 5 shows a block diagram of a typical stand-alone implementation.

Figure 2 – Stand Alone Block Diagram

(Voice Direct Serial Memory 4-Word Application Shown)



Power On

At power-on Voice Direct checks the level on the MODE/SP1 pin. An external 100K resistor pulls this pin to either Vdd or GND. When the pin is pulled to GND, the Voice Direct begins operation in Stand Alone Mode. The chip continues to operate in Stand Alone Mode as long as the power remains on. After this power-on check, the MODE/SP1 pin is automatically configured for use as a driver pin for speaker output.

At the beginning of its Stand Alone Mode power-on tests, the Voice Direct activates all of the OUTx signals, the HIGH signal, and the ERROR signal. Then Voice Direct performs the power-on configuration operations. The configurations remain set until power is turned off.

Memory Configuration

Voice Direct uses the SCL and SDA signals to communicate with the Serial EEPROM. If this template memory has valid data, Voice Direct retrieves internal information about the number of words. If this template memory has invalid data, the Serial EEPROM is initialized with Voice Direct default values.

Training and Recognition Selectivity Pins

The RECOG and TRAIN pins are checked to see if either or both are pulled to Gnd by 100K programming resistors. These bits determine the selectivity level for training and recognition.

- **TRAIN high** selects optimum selectivity level in training. In this case, Voice Direct will accept words that are similar sounding to each other, possibly degrading recognition accuracy.
- **TRAIN low** selects optimum selectivity level in training. In this case, Voice Direct will reject words that are too similar sounding to each other, potentially increasing recognition accuracy.
- **RECOG high** selects minimum selectivity level in recognition. In this case, Voice Direct will produce less rejection errors of words in the list, but may produce more substitution errors of correct words.
- **RECOG low** selects maximum selectivity level in recognition. In this case, Voice Direct will produce less substitution errors of correct words, but may produce more rejection errors of words in the list.

Stand-Alone Selectivity Levels

	Open	Closed
S1	Optimum Selectivity	Maximum Selectivity
S2	Optimum Selectivity	Maximum Selectivity

- The Train and Recog pins are pulled high with strong pullups. In lieu of switches S1 and S2, these pins can be directly connected to 100K pull-down resistors, or left floating.

After some further internal checking, Voice Direct is ready to begin operation. It de-activates all output signals and serves notice that it is ready with one or two short audible "beeps". If the unit is in Stand Alone Mode, a single beep will be generated, if it is in External Host Mode, two beeps will be generated (the synthesis of these beeps can be suppressed by altering an External Host Developer Parameter). It then begins monitoring for commands to train, recognize, or erase.

Training

Pressing the TRAIN switch for at least 1/10 second initiates automatic training of up to 15 individual words or phrases. A word or phrase must be shorter than 3.2 seconds and may not contain silences longer than 0.5 seconds. For example, the name "John Smith" would be an acceptable phrase as long as the two words are not separated by a large pause. Training terminates when no word is spoken in response to a prompt, when either switch is pressed a second time during training, when three errors have occurred during training, *or after all words have been trained.*

Pressing the TRAIN switch at a later time resumes training. New words are added to the end of the set already recorded. New words may be added to the set at any time, up to a maximum of 15 words. Individual words from the set may not be deleted or overwritten, but the entire set can be erased (see "Erasing" below).

Recognition

Recognition is initiated when the RECOG button is pressed. A prompt is spoken and the Voice Direct listens to the word being spoken. If the spoken word matches a stored template, one (or two) of 8 category pins is activated (pulled high for approximately one second) and a voice message indicates the matching category number. If the set contains 8 or fewer elements, these pins may be used to control actions directly. If the set contains more than 8 elements, hardware decoding is necessary. If words nine to fifteen are recognized, output pin 8(HIGH) will be pulled high, in addition to output pins 1 to 7 respectively. If the spoken word is not recognized for any reason, none of the category pins is activated and an appropriate voice message is synthesized. Since this can occur if Voice Direct has not been trained on the word, the "word not recognized" output is not considered an error condition, and the ERROR bit remains inactive.

Voice Direct monitors the background noise level and gives a voice warning if the noise is too high for recognition. Voice Direct works well with high (above 80 dB) levels of *steady* background noise (such as a fan), but it may make errors at lower levels if the background noise is not steady (for example, a TV set). Best results will be obtained in a relatively quiet location.

Warnings may also appear if the word is spoken too softly, too loudly, or too quickly after the prompt. The ERROR signal goes high when these conditions occur.

Erasing
Users can erase all stored templates by pressing the TRAIN and RECOG buttons together for at least one second. A voice message confirms the operation.

Errors
Any error occurring during training or recognition produces a spoken message. The ERROR signal is also activated (for approximately 1 second) to indicate an error.

In addition, the error code is placed on the Memory Data bus, and the -WR (Write Result) signal is strobed.

A non-zero value in the data bus indicates that recognition or training did not succeed. See Command Set Overview p. 21.

Attributes and the Mask
All words trained in Stand Alone Mode are trained with an attribute of 01h, and all recognition is performed using a Mask of FFh. Thus, in Stand Alone Mode, the possible recognition set consists of all trained words.

External Host Mode

If the MODE signal is resistively pulled to Vdd, Voice Direct will operate in the External Host Mode, which requires an external control device (MCPU). You program the MCPU to issue commands that control the Voice Direct IC.

The external host operating mode is designed to provide a full-featured, multiple-set recognition system combined with high quality phrase synthesis in a chip that can easily be controlled by an External Host. Recognition templates are stored in non-volatile serial EEPROM accessed by Voice Direct. While providing a powerful high-level control interface, the External Host Mode Voice Direct allows the MCPU to control the flow of operations and to initiate training, recognition, or synthesis as appropriate. This contrasts with Stand Alone Mode, which requires a fixed sequence.

External Host Mode supports up to 60 words organized in flexible recognition sets. During recognition and directory functions, a mask is applied against each attribute byte to select and/or exclude words in specific groups.

In External Host Mode, Voice Direct allows the MCPU to control the five Voice Direct I/O lines directly via the Configure/Read/ Write I/O Expansion commands. For more on these commands, see page 30.

External Host Mode operation uses certain pins for functions that differ from those used in Stand Alone Mode functions. Alias names are given below for these pins to clarify their use in External Host Mode.

External Host Data Transfer Signal Description

Signal	Pin	Description	VoiceDirect=>MCPU	MCPU=>Voice Direct
DATA	28	Bi-directional Data line	Data valid when SHS = 0	Data valid when MHS = 0
MHS	29	Master Handshake (Host => VoiceDir)	0 = Data bit valid	0 = Data bit accepted
SHS	30	Slave Handshake (VoiceDir => Host)	0 = Data bit accepted	0 = Data bit valid

Miscellaneous I/O

Signal	Pin	Description	Use
Talk	31	Indicates that the Voice Direct is generating speech output	Switching audio paths when synthesizing speech

The Voice Direct Command Set

Command Set Overview

This section describes the set of Voice Direct commands and command responses. The commands are sent to Voice Direct by the MCPU via the serial interface (see Serial Interface Communications, p. 31).

> *This section applies only to Voice Direct in External Host Mode.*

Table 1 - Command Set

Command	Parameters	Data returned	Description	Class
00h	~	~	No Operation	6
01h	~	5 byte version code string	Get Version	6
02h	Prompt, Source, Tries	~	Train A Name	1
10h	Prompt, Source, Tries	~	Recognize Word	2
20h	~	~	Clear Current Pointer	3
21h	~	~	Increment Current Pointer	3
22h	~	~	Decrement Current Pointer	3
23h	~	~	Save Current Pointer	3
24h	~	~	Restore Current Pointer	3
25h	~	~	Swap Current Pointer	3
26h	~	Current Pointer (1 byte)	Return Current Pointer	3
27h	~	~	Delete Current Entry	
28h	0x55	~	Delete All Stored Words	3
29h	Mask, Attribute	~	Set Mask and Attribute	3
2Ah	~	Attribute, Environment	Get Entry Status	3
2Bh	~	Capacity, Free, Matching	Query Lexicon Status	3
30h	Prompt, Source	~	Say A Prompt	4
40h	~	~	Power Down	5
41h	Index, Parameter	~	Store Parameter	5
42h	Index	Parameter	Fetch Parameter	5
43h	5 bits of data[1]	~	Configure I/O Expansion	5
44h	~	5 bits of data[1]	Read I/O Expansion	
45h	5 bits of data[1]	~	Write I/O Expansion	5

Class definitions:

1 – Training
2 – Recognition
3 – Directory function
4 – Synthesis
5 – Control function
6 – Debug

[1]Data is sent as a single byte, with Data[4:0] corresponding to Expansion IO[4:0].

Command Responses

Table 2 - Command Responses

Response	Class	Description
00h	All	Command successfully executed to completion
01h	1,2	Time out
02h	1,2	Reserved
03h	1,2	Background too noisy
04h	1,2	Spoke too soft
05h	1,2	Spoke too loud
06h	1,2	Spoke too soon
07h	2	No match found
08h	1	Uncertain match found
09h	1	Entry too similar
0Ah- 0Fh		Reserved
10h	2,3	Memory empty
11h	1	Memory full
12h	3	Invalid pointer (No matching Attributes)
13h	3	No matching Attributes
14h – 1Fh		Reserved
20h	All	Unknown command
21h	All	Invalid parameter (value, digit, or index)
FFh	1,2,4	Command interrupted

Functional Description

What follows is a functional description of each command.

The invalid parameter response (21h) is always a possible command response, even when not explicitly cited in the descriptions that follow.

Command – 00h (No Operation)
Arguments: ~
Returns: ~
Responses: 00, 21

This command returns a response code of 00h (Success). It is provided primarily for testing the communications interface.

Command 01h - (Get Version)
Arguments: ~
Returns: Version String (56h, 44h, 49h, 01h, 01h)
Responses: 00, 21

This command returns a five-byte sequence indicating the Sensory product model and software version number. The first three bytes will always be 56h, 44h, and 49h (ASCII 'VDI') for the Voice Direct. The last two bytes are, respectively, the major and minor version numbers in binary format.

Command 02h - (Train A Name)
Arguments: Prompt, Source, Tries
Returns: ~
Responses: 00, 01, 02, 03, 04, 05, 06, 08, 09, 11, 21, FF

This command allocates directory space and launches a training session for new entries. The training session generates a speaker-dependent speech template. This command requires three parameters: an index indicating which spoken prompt is to be used, a byte indicating the source synthesis table to use (00=Mandatory Table, 01= Optional Table, see Prompt Lists, p.) and the number of retries allowed. This command may be aborted by the MCPU by transmitting FFh to the Voice Direct.

The user is prompted using the prompt index supplied, then asked again to "Repeat." The new word is tested for similarity with existing words selected by the current mask. When the entry is saved to permanent memory, the current attribute is also saved (see Set Mask and Attribute Command, p. 27). After Voice Direct saves the entry, the Current Pointer points to the newly created entry. If the training process is unsuccessful after the specified number of retries, the command response will reflect the last attempt (i.e., the reported failure will be that of the final try) and returned to the MCPU. If the trained name is too similar to an existing entry name, then the Current Pointer points to the entry that caused the error, but no spoken error is output to the speaker.

Table 3 - Training Response Codes

Command Response	Current Pointer
Success (00h)	Index of new entry
Too similar (09h)	Index of entry which is too similar
Any other error	Invalid

Command 10h - (Recognize Word)

Arguments:	**Prompt, Source, Tries**
Returns:	**~**
Responses:	**00, 01, 02, 03, 04, 05, 06, 07, 08, 10, 21, FF**

This command prompts the user for a name, creates a temporary word template and compares it against stored templates for a match. Specific word sets may be selected using Command 29h (Set Mask and Attribute.) If recognition is successful, the Current Pointer will be set to the index of the entry that matches the word. If the recognition is unsuccessful after the specified number of retries, the Current Pointer will be invalid and the command response will reflect the last recognition attempt. This command requires three parameters to execute successfully: an index indicating which spoken prompt is to be used, a byte indicating the source synthesis table to use (00= Mandatory Table, 01=Optional Table, see Prompt Lists, p. 47). and the number of retries allowed. This command may be aborted by the MCPU by transmitting FFh to Voice Direct.

Table 4 - Recognize Word Response Code

Command Response	Current Pointer
Success (00h)	Index of new entry
Confused (08h)	Best matching entry (Use Command 25h – Swap Current Pointer to select 2nd best matching entry)
Any other error	Invalid

Command 20h - (Clear Current Pointer)	
Arguments:	**~**
Returns:	**~**
Responses:	**00, 10, 13, 21**

This command resets the Current Pointer to the first matching entry based on the current mask; it should be used before Commands 21h (Increment Current Pointer) or 22h (Decrement Current Pointer) so that directory entries may be reviewed from a known position. If no entries match the current mask, then this command returns a Command Response of 13h (No Matching Attributes). No parameters are required.

Any template whose Attribute has a bit set to 1, which is also set to 1 in the Mask, will be recognized and operated upon. Any template whose Attribute does not have a bit set which matches a Mask bit will be ignored during all recognition and directory operations (see the Set Mask and Attribute Command, p. 27)

Command 21h - (Increment Current Pointer)	
Arguments:	**~**
Returns:	**~**
Responses:	**00, 10, 14, 21**

This command causes Voice Direct to increment (move forward) the Current Pointer to the next matching directory entry based on the mask. This command can be used in routines that browse the directory. There are no parameters required for this command.

> **Increment Current Pointer** *and* **Decrement Current Pointer** *are not necessarily linear functions – when used in conjunction with a mask byte they will skip over names that the mask does not match.*

Command 22h - (Decrement Current Pointer)	
Arguments:	**~**
Returns:	**~**
Responses:	**00, 10, 14, 21**

This command causes Voice Direct to decrement (move backward) the Current Pointer to the previous matching directory entry based on the mask. This command can be used in routines that browse the directory. There are no parameters required for this command.

Command 23h - (Save Current Pointer)	
Arguments:	**~**
Returns:	**~**
Responses:	**00, 21**

This command saves the Current Pointer in a temporary memory location. The saved Current Pointer value can be restored by Command 24h (Restore Current Pointer). This command requires no parameters.

> *Note that the saved Current Pointer is lost if power is removed from the Voice Direct.*

Command 24h - (Restore Current Pointer)	
Arguments:	**~**
Returns:	**~**
Responses:	**00, 12, 21**

This command restores the Current Pointer to its last saved value – the value saved by a Command 23h (Save Current Pointer). Note that if directory entry associated with the Current Pointer has been deleted, then the restored Current Pointer will no longer point to a valid directory entry. There are no parameters required for this command.

Command 25h - (Swap Current Pointer)	
Arguments:	**~**
Returns:	**~**
Responses:	**00, 21**

If, after a recognition attempt, there are two possible recognition candidates in the speaker-dependent templates, and the Current Pointer points to the best matching entry. This command switches the Current Pointer to point to the second-best matching entry. Issuing this command again switches the Current Pointer back to the best matching entry. There are no parameters required for this command.

Command 26h - (Return Current Pointer)	
Arguments:	**~**
Returns:	**Current Pointer**
Responses:	**00, 21**

This command causes the Voice Direct to return the one byte value for the Current Pointer. This command requires that the Current Pointer be valid. There are no arguments required for this command. The returned Current Pointer is the position in template memory where the current template resides. This is indexed with the first template recorded starting at Current Pointer=0.

Command 27h - (Delete Current Entry)	
Arguments:	**~**
Returns:	**~**
Responses:	**00, 12, 21**

This command deletes the entry that is associated with the Current Pointer. If the Current Pointer is not valid, then this command does not delete an entry and instead returns a Command Response of 12h (Invalid Pointer). There are no parameters required for this command. The deletion of an entry makes the associated memory available for new template storage.

Command 28h - (Delete All Stored Words)	
Arguments:	**0x55**
Returns:	**~**
Responses:	**00, 21**

This command deletes all stored names and returns the external SEEPROM to the initial state. This command requires a one-byte parameter of the value 055h.

Command 29h - (Set Mask and Attribute)	
Arguments:	**Mask, Attribute**
Returns:	**~**
Responses:	**00, 21**

This command is used to set the internal variables mask and attribute when using multiple word sets. Mask and attribute have power up defaults of FFh and 01h, respectively. This allows a single directory of words to be managed without concern for this command. See the sections on Attribute Bytes (p. 13) and Masks (p. 14).

Command 2Ah - (Get Entry Status)	
Arguments:	**~**
Returns:	**Attribute, Environment**
Responses:	**00, 12, 21**

This command returns the attribute and status of the entry associated with the Current Pointer. The first byte sent is the attribute byte for the entry and the next byte indicates the environment used for training the template. However, if the Current Pointer is not valid, this command does not return an entry status and instead returns a Command Response of 12h (Invalid Pointer). There are no parameters required for this command.

Table 5 – Environment Byte Definition

Bit	Description
Bit0=0	NEAR
Bit0=1	FAR
Bit2,3=00	Quiet
Bit2,3=11	Automotive

Command 2Bh - (Query Directory Status)	
Arguments:	**~**
Returns:	**Capacity, Free, Entries**
Responses:	**00, 21**

This command is used to return information about the directory. Three bytes are returned: Total word Capacity, Number of Free slots available, and Number of directory Entries selected by the current mask byte.

Command 30h - (Say A Prompt)	
Arguments:	**Prompt, Source**
Returns:	**~**
Responses:	**00, 12, 21**

This command plays a predefined voice prompt from the onboard ROM or an External Speech ROM. It requires two parameters to execute successfully: an index indicating which spoken prompt is to be used, and a byte indicating the source synthesis table to use (00=Mandatory Table, 01=Optional Table). See the Prompt Lists, p.47. This command may be aborted by the MCPU by transmitting FFh to the Voice Direct.

Command 40h - (Power Down)	
Arguments:	**~**
Returns:	**~**
Responses:	**00, 21**

This command causes Voice Direct to enter a low power sleep state. This command returns a "successful" response just before entering low power mode.

> *Note that the only method for exiting this low power state is a pulse on Voice Direct reset pin. Voice Direct cannot be woken up by a serial command.*

Command 41h - (Store Parameter)	
Arguments:	**Index, Parameter**
Returns:	**~**
Responses:	**00, 21**

This command stores a new value for a single parameter in the SEEPROM parameter block. The required parameters are an index and value. The SEEPROM parameter block provides a number of user configurable options, as well as 23 bytes of general purpose storage. See the section on Developer Parameters (p.30).

Command 42h - (Fetch Parameter)	
Arguments:	**Index**
Returns:	**Parameter**
Responses:	**00, 21**

This command returns a single parameter value from the SEEPROM parameter block. This command requires an index and returns a value. The SEEPROM parameter block provides a number of user configurable options, as well as 23 bytes of general purpose storage. See the section on Developer Parameters (p.30).

Command 43h - (Configure I/O Expansion)	
Arguments:	**5 bits of I/O data**
Returns:	**~**
Responses:	**00, 21**

This command configures the Expansion I/O pins. IO[4:0] is addressed as Data[4:0]. To configure an I/O as an output, set the bit to "1". To configure an I/O as input, set the bit to "0".

Command 44h - (Read I/O Expansion)	
Arguments:	**~**
Returns:	**5 bits of I/O data**
Responses:	**00, 21**

This command returns the values on the Expansion I/O pins that have been configured as inputs. IO[4:0] is returned as Data[4:0]. A high level returns "1", a low level returns "0". Bits programmed as outputs return the output values.

Command 45h - (Write I/O Expansion)	
Arguments:	**5 bits of I/O data**
Returns:	**~**
Responses:	**00, 21**

This command sets the I/O pins programmed as outputs to the values supplied by Data. Data[4:0] are written as IO[4:0]. Each output pin is set to a high level if the corresponding bit is "1"; the output pin is set to a low level if the corresponding bit is "0". Pins programmed as inputs are unaffected.

Developer Parameters

The developer parameters let you control and configure your Voice Direct application. You can select settings that anticipate the proximity of users to the microphone, the kind of audio output (DAC and/or PWM) to select, whether beeps are to be synthesized upon power-up or reset, and so on. You set these parameters using the Store Parameter Command (see the Store Parameter Command, p. 29). The table and text below summarize the choices.

Table 6 - Developer Parameters

Index	Parameter	Default	Range
00h	Environment	00h	Bit0 = 0, NEAR Bit0 = 1, FAR Bit2,3 = 00, Quiet Bit2,3 = 11, Automotive
01h	Reserved	FFh	00h – FFh
02h	Reserved	80h	00h – FFh
03h	Reserved	20h	00h – FFh
04h	Reserved	20h	00h – FFh
05h	Reserved	01h	00h – 01h
06h	Reserved	00h	00h – 01h
07h	Speech Output Select	03h	01h - 03h
08h	Startup Beep Select	01h	00h – 01h
09h-1Fh	Unused	00h	00h - FFh

Functional Description

07h - Speech Output Select Parameter This parameter allows the user to select the analog output device to be used with speech synthesis. A value of 01h will direct the output to the DAC and a value of 02h will direct the output to the PWM. Note that it is legal to select both the DAC and PWM simultaneously by setting this parameter to 03h. The default value for this parameter is 03h (both outputs).

08h – Startup Beep Select Parameter This parameter allows the user to select whether the Voice Direct will synthesize a beep (or series of beeps) upon power up or reset. A value of 00h will prevent synthesis of these beeps and a value of 01h will cause the beeps to be synthesized. The default value for this parameter is 01h (beep synthesis on).

Serial Interface

This section applies only to Voice Direct in External Host Mode.

Communications

Data communication and control are accomplished using a 3-wire synchronous serial interface. Since the serial interface lines are open-collector (open-drain), pull-up resistors should be attached to all signals. If application requirements call for Voice Direct to be powered down while the rest of the system is still powered up, take care to ensure that the external pull-up resistors are also powered down. Doing so will prevent current leakage into the Voice Direct I/O ports. The following table describes the 3 lines used for data and handshaking between the Voice Direct and the MCPU (Host). In the table and the timing diagrams, VDI refers to the Voice Direct IC.

Table 7 - Data Transfer Signal Description: External Host Mode

Signal	Pin	Description	VDI=>MCPU	MCPU=>VDI
DATA	28	Bi-directional Data line	Data valid when SHS = 0	Data valid when MHS = 0
MHS	29	Master Handshake (Host => VoiceDir)	0 = Data bit valid	0 = Data bit accepted
SHS	30	Slave Handshake (VoiceDir => Host)	0 = Data bit accepted	0 = Data bit valid

Miscellaneous I/O

Signal	Pin	Description	Use
Talk	31	Indicates that the Voice Direct is generating speech output	Switching audio paths when synthesizing speech

Data is transferred one bit at a time with full handshaking as described below.

1. When the MCPU has data to transmit to the VDI, the MCPU sets DATA to the data value, verifies that SHS (Slave Handshake) is in the *high* state, then sets MHS (Master Handshake) to the *low* state to request a transfer.

2. The VDI senses the *low* state of MHS and reads DATA, which then sets SHS to the *low* state to acknowledge the DATA.

3. The MCPU senses the *low* state of SHS, and sets MHS to the *high* state to indicate that DATA is no longer valid, and at the same time sets DATA high (releasing it).

4. The VDI then sets SHS to the *high* state to indicate that the cycle is complete. Both devices are now ready to transfer the next data bit.

Voice Direct remains busy (SHS = low state) after receiving the final bit of a command packet, until after that command has been completed and Voice Direct is ready to send a response. During this time, a time-consuming command can be interrupted by the MCPU with a low pulse on the MHS line. When Voice Direct has data to transmit to the MCPU, it follows the same procedure by setting SHS to the low state. (See Figures 2, 3, and 4.) The protocol is completely symmetrical. The first processor to set its HS signal to the low state is the transmitter; the other processor is the receiver.

Data is always transferred in 8 bit bytes, with the most significant bit transferred first.

> *Note: By convention, the slave (Voice Direct) never transmits data unless requested to by the master (MCPU), thus avoiding the possibility of a collision (both processors setting their HS signals low simultaneously).*

Figure 3 - Data Transfer – (Command to Response)

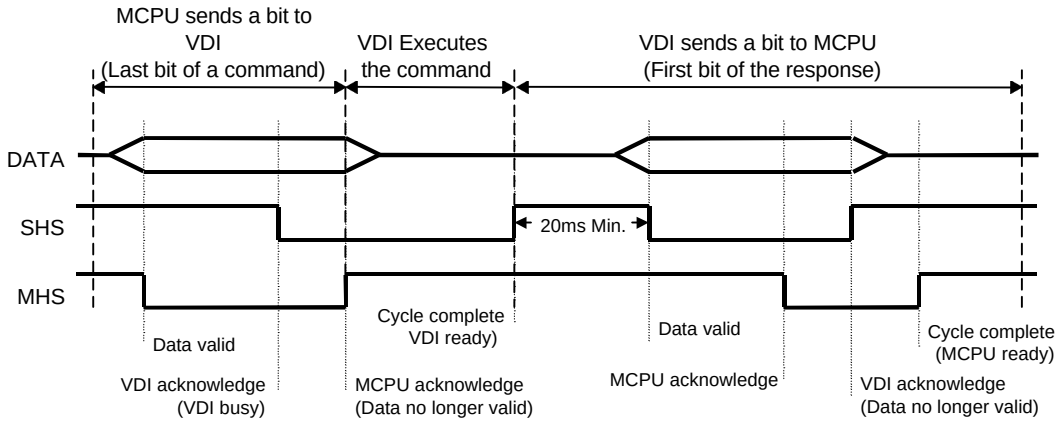


Figure 4 - Interrupt Timing Diagram

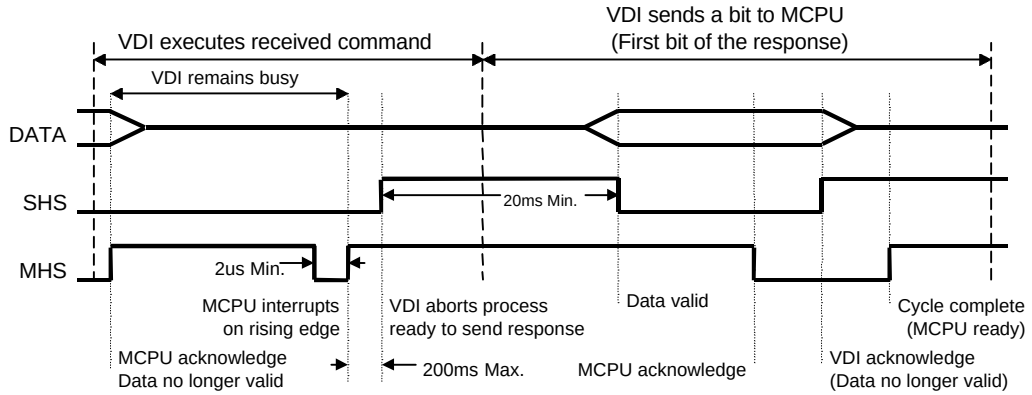
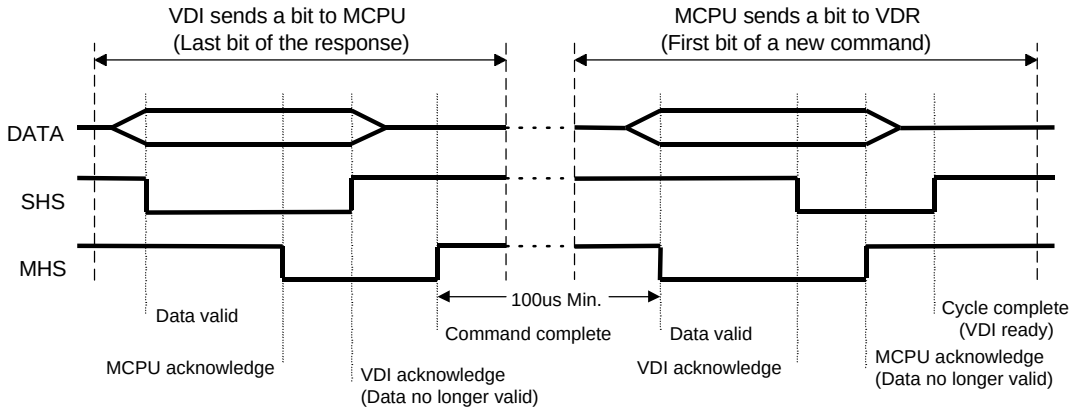


Figure 5 - Data Transfer (Response to Next Command)



Data Packets

This section applies only to Voice Direct in External Host Mode.

Overview

All data is transmitted in 8-bit bytes, and all messages are sent in packets. The packet format conveys error checking and byte synchronization information. A packet always starts with a "sync field" of FFh (8 or more 1 bits) followed by a length byte, then one or more data bytes, then a checksum byte. The length byte specifies the number of bytes to follow including the checksum. The checksum is the 8-bit additive, modulo-256 sum of all the data bytes and the length byte added together. For example the data 01, 02, 03, 04 would be sent as follows:

Table 8 - Sample Data Packet

Byte	Value	Notes
0	FFh	Sync Field, 8 or more consecutive 1 bits
1	05h	Packet length, count of bytes to follow
2	01h	First data byte
3	02h	Second data byte
4	03h	Third data byte
5	04h	Forth data byte
6	0Fh	Checksum, (5+1+2+3+4)

Implementation

Packet/byte synchronization is accomplished as follows:

1. The receiver shifts in bits until the accumulated byte value is equal to FFh (all bits are zero).

2. The receiver then continues shifting in data until the most significant bit resets to zero. This indicates that the currently accumulated byte is a valid length byte and the start of a packet.

For synchronization to work properly, the receiver must first initialize its shift register to all zeroes. Once a valid packet is found the receiver accepts (length) bytes and performs the checksum calculation (ascertains that the sum of all the data bytes plus the length byte is equal to the checksum byte). Note that modulo-256 arithmetic is used; the carry is discarded during the checksum calculation.

It is worth recalling that all communication is initiated by the MCPU; the Voice Direct never sends data unless requested by the MCPU. Accordingly, all commands are in one of the following formats:

- A request from the MCPU to Voice Direct. A request consists of a command byte and possible parameters or data.
- A response from Voice Direct to the MCPU. A response consists of a status byte and possible data.
-

After a command is issued to Voice Direct, the MCPU must wait until the VDI is ready (SHS = high) before receiving the command response. The MCPU can interrupt a command by pulsing the MHS line to the low state. This causes Voice Direct to abort the task and return appropriate status.

Audio Output

Voice Direct can directly drive a 32-Ohm speaker from the SP0 and SP1 pins, providing approximately 0.15 Watts of audio power. SP0 and SP1 are push-pull Pulse Width Modulation (PWM) outputs with an aliasing frequency well above the audio band, so the signals can be connected directly to a speaker's terminals with no additional filtering.

DAC output – In addition to the PWM audio output, Voice Direct also provides a high-impedance (22kOhm) analog audio output, DAC. This signal must be power amplified to drive a speaker, and should be low-pass filtered with a corner frequency around 20 KHz. This additional circuitry provides better speech quality than the PWM output. Sensory recommends using the DAC audio output for applications requiring either louder volume or better speech quality.

External Components and Specifications

This section includes a brief description and schematic of the external components needed to fully characterize a Voice Direct IC. You can acquire these components and then use the schematic to build the circuit. Or Sensory, Inc. can supply you with everything you need:

- *The Voice Direct Module:* A pre-configured circuit for prototype development. The Voice Direct Module has all of the external components listed in this section and shown in the schematic, except for a microphone and speaker.

- *The Voice Direct Evaluation Kit:* A fully configured Voice Direct Host Interface Board (for external host-driven applications) and a Stand Alone Evaluation Board. Both boards incorporate the Voice Direct Module, speaker, microphone, battery pack and more. Three Voice Direct modules are included. The kit also includes a parallel port card and Sensory's graphical VoiceHost software to evaluate, demonstrate, and train Voice Direct for external host-driven applications. For more information on these products, contact Sensory.

As the block diagram (Figure 1) illustrates, Voice Direct needs very little external circuitry to operate in either mode.

Microphone

The microphone is used for speech inputs. It is an inexpensive omni-directional electret model with a minimum sensitivity of -60dB. Required for both stand alone and External Host Modes. *Included in the Voice Direct Evaluation Kit.*

Preamplifier

The preamplifier amplifies and filters the microphone signal to a level suitable for Voice Direct. It is either a 2-bit AGC or non-AGC. See the preamplifier schematic. Required for both stand alone and External Host Modes. *Included in the Voice Direct Module and the Voice Direct Evaluation Kit.*

When used with an inexpensive omni-directional electret microphone, the input audio amplifier and filter must provide approximately 58 dB of low-noise mid-band gain, 2-bit AGC controllability, and a first order bandpass filter response with 3dB points at roughly 700 Hz and 3300 Hz. The recommended audio input schematic is shown in Figure 6- Host Interface Board Schematic

. Sensory strongly recommends the use of this circuit.

External Memory

Voice Direct uses non-volatile memory to store speech templates. It is a 8 Kbyte serial EEPROM (MicoChip 24C65) required for both stand alone and External Host Modes. *Included in the Voice Direct Module and the Voice Direct Evaluation Kit.*

Oscillator

The oscillator provides a high frequency clock (14.32 MHz) for Voice Direct. It must be a crystal ceramic resonator or LC circuit. Required for both stand alone and External Host Modes. *Included in the Voice Direct Module and the Voice Direct Evaluation Kit.*

Speaker

The speaker is used to output audio prompts from Voice Direct. The type of speaker depends on whether you use the IC or the module. The IC requires an amplified DAC 8 ohm speaker with high input impedance. If you are using the module, then you can use either a DAC output with an amplifier and an appropriate speaker, or use the PWM directly with a 32 ohm speaker. Required for both stand alone and External Host Modes. *Included in the Voice Direct Evaluation Kit.*

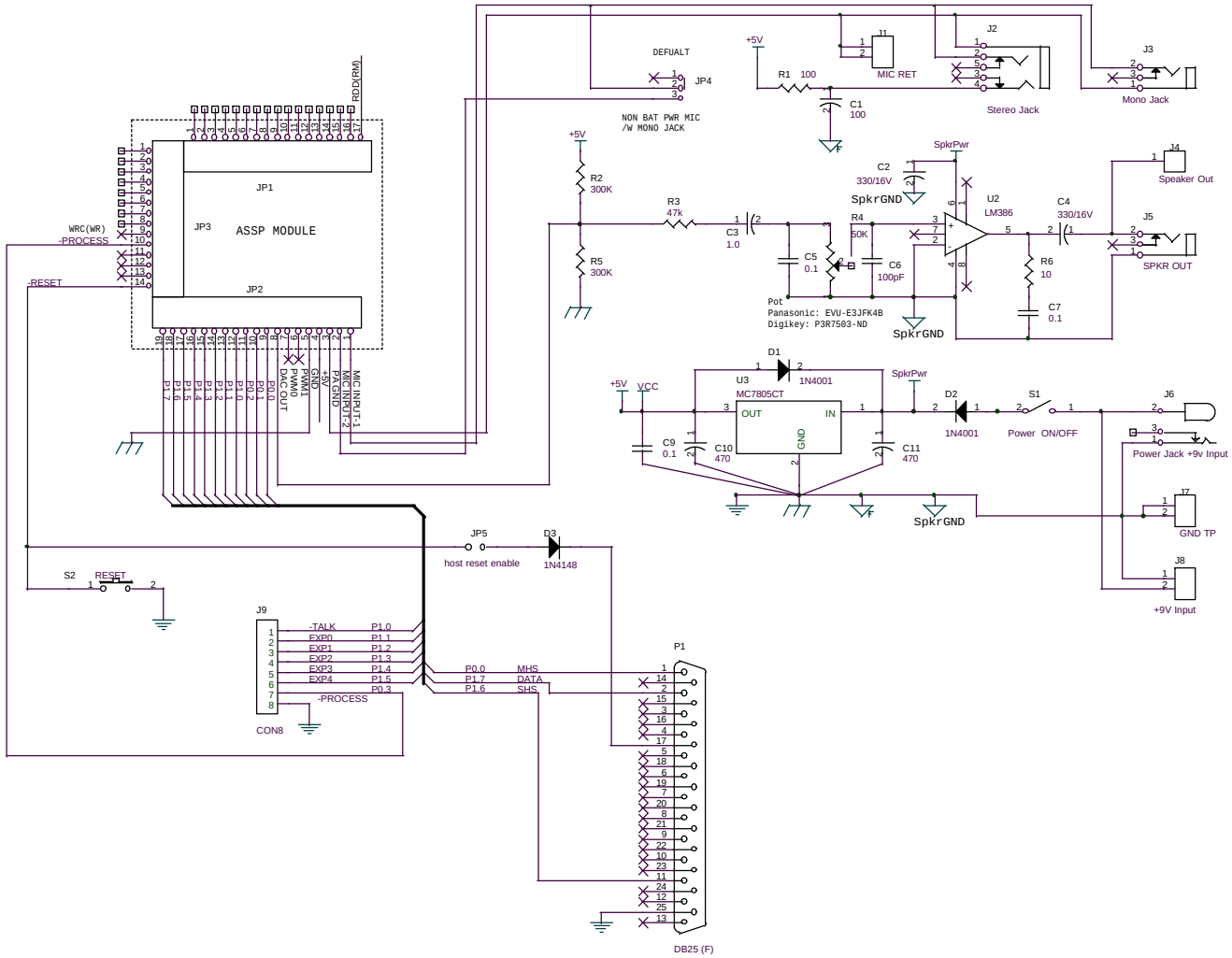
External ROM (Optional)

The external ROM is optional. It is required for customized speech or non-standard English language applications. For more information, contact Sensory. *A slot for external ROM is included in the Voice Direct Evaluation Kit.*

This page is intentionally blank.

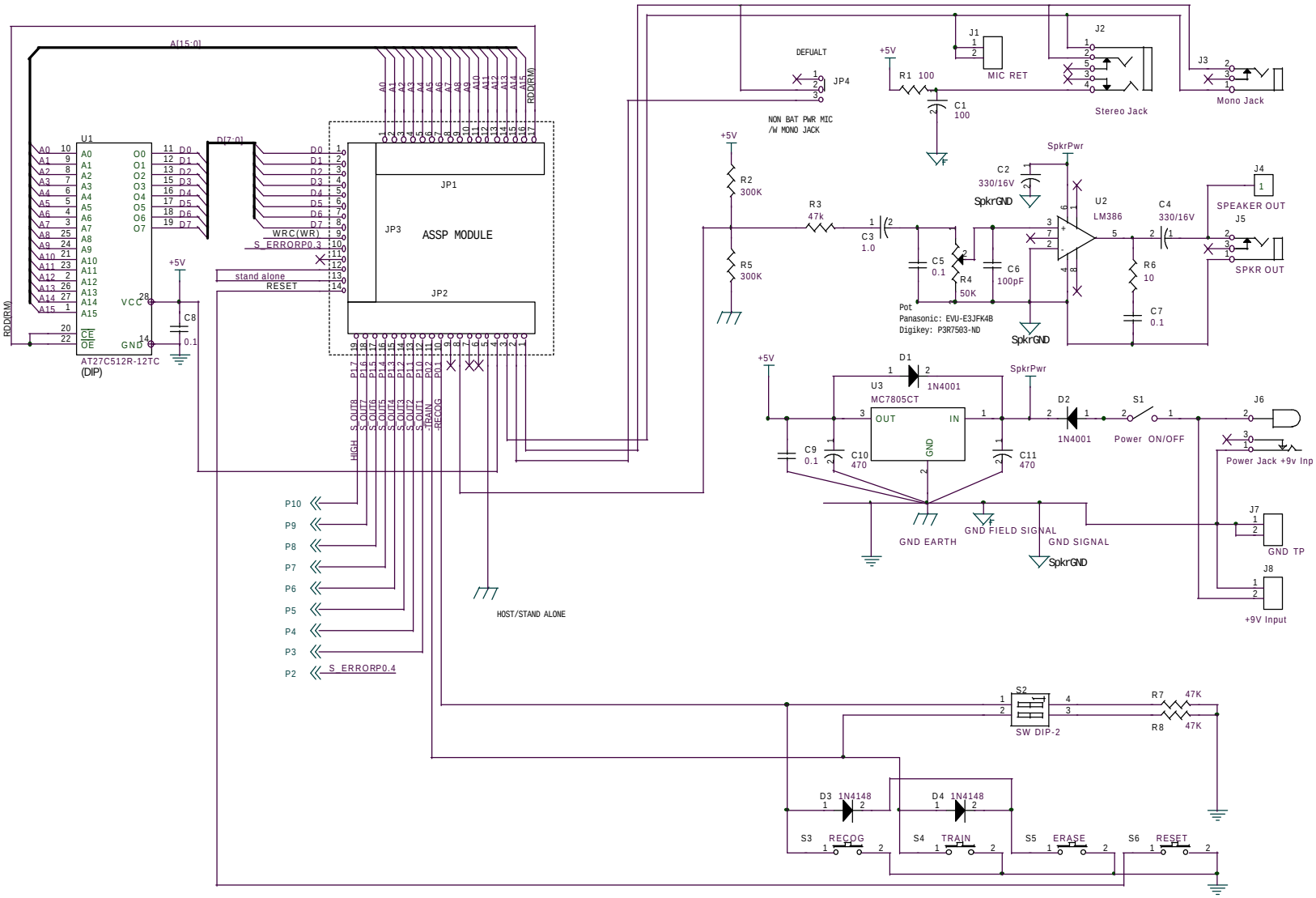
Schematics

Figure 6- Host Interface Board Schematic



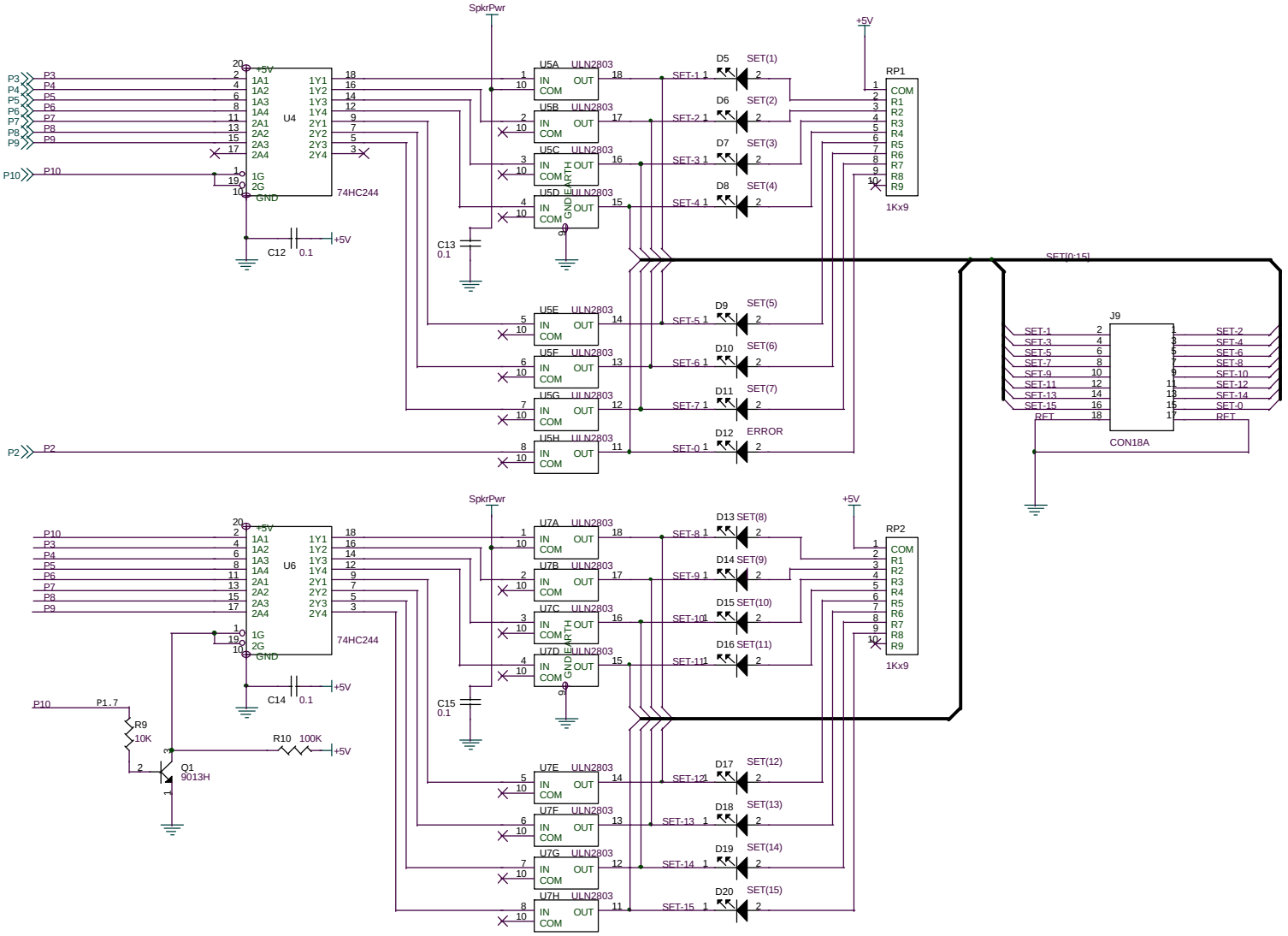
This page is intentionally blank.

Figure 7 – Voice Direct Stand Alone Board (1of 2)



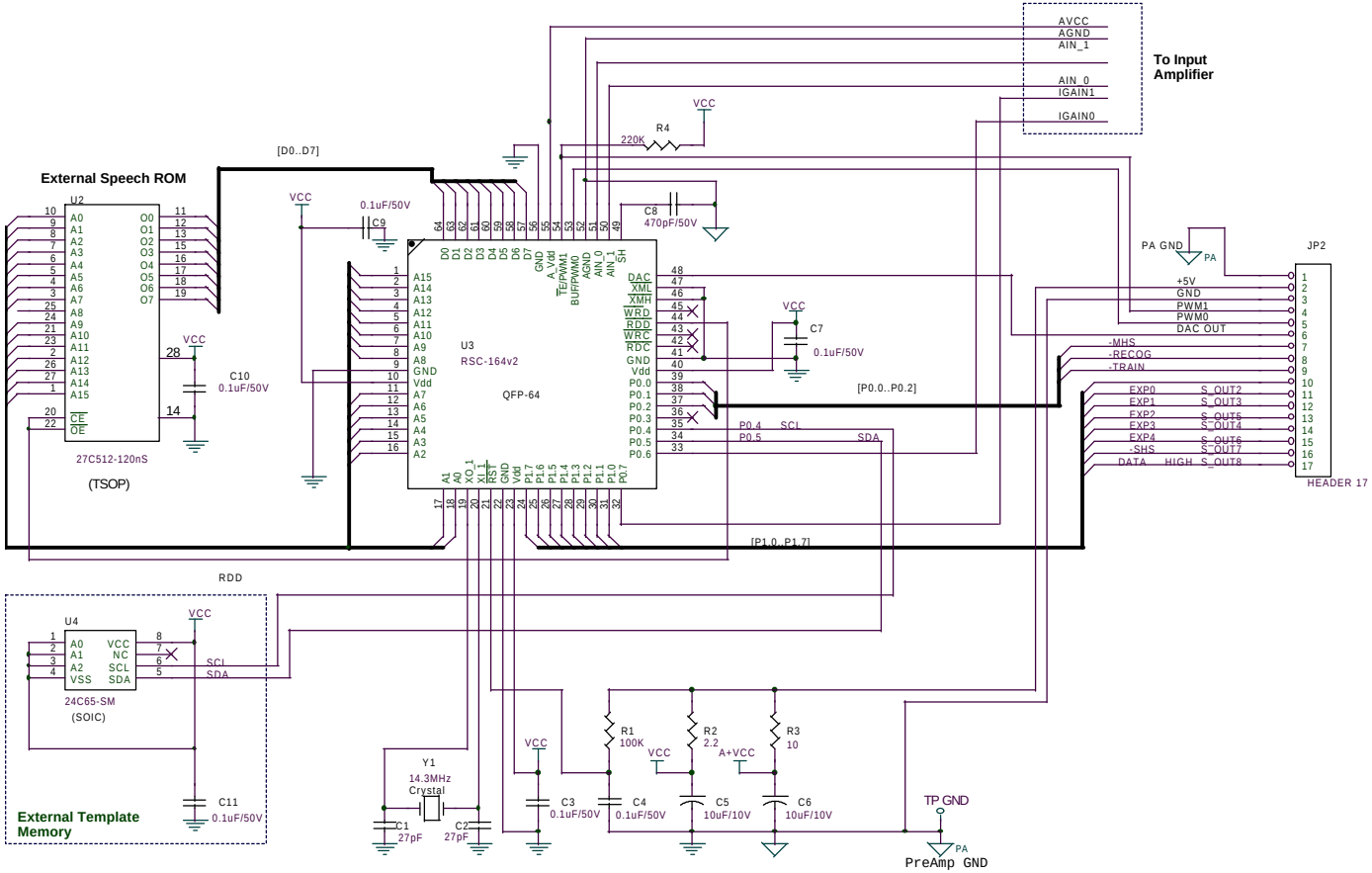
This page is intentionally blank.

Figure 8 – Voice Direct Stand Alone Board (2 of 2)



This page is intentionally blank.

Figure 9 – Voice Direct and External Components



This page is intentionally blank.

User Interface

Mandatory Prompt List

While running training and recognition processes, Voice Direct has direct and exclusive access to speech in the Mandatory Prompt List. The Mandatory Prompt List is actually a set of pointers to speech elements listed in the table below.

> *The spoken prompts in the Mandatory Prompt List can be duplicated in the Optional Prompt list with little memory penalty.*

Table 9 – Mandatory Prompt List

Prompt #	Spoken Prompt Description
1	one
2	two
3	three
4	four
5	five
6	six
7	seven
8	eight
9	nine
10	ten
11	eleven
12	twelve
13	thirteen
14	fourteen
15	fifteen
16	sixteen
17	seventeen
18	eighteen
19	nineteen
20	twent0y
21	twenty one
22	twenty two
23	twenty three
24	twenty four
25	twenty five
26	twenty six
27	twenty seven
28	twenty eight
29	twenty nine
30	thirty

31	thirty one
32	thirty two
33	thirty three
34	thirty four
35	thirty five
36	thirty six
37	thirty seven
38	thirty eight
39	thirty nine
40	forty
41	forty one
42	forty two
43	forty three
44	forty four
45	forty five
46	forty six
47	forty seven
48	forty eight
49	forty nine
50	fifty
51	fifty one
52	fifty two
53	fifty three
54	fifty four
55	fifty five
56	fifty six
57	fifty seven
58	fifty eight
59	fifty nine
60	sixty
61	say word one
62	say word two
63	say word three
64	say word four
65	say word five
66	say word six
67	say word seven
68	say word eight
69	say word nine
70	say word ten
71	say word eleven
72	say word twelve
73	say word thirteen
74	say word fourteen
75	say word fifteen
76	say word sixteen
77	say word seventeen
78	say word eighteen
79	say word nineteen
80	say word twenty

81	say word twenty-one
82	say word twenty-two
83	say word twenty-three
84	say word twenty-four
85	say word twenty-five
86	say word twenty-six
87	say word twenty-seven
88	say word twenty-eight
89	say word twenty-nine
90	say word thirty
91	say word thirty-one
92	say word thirty-two
93	say word thirty-three
94	say word thirty-four
95	say word thirty-five
96	say word thirty-six
97	say word thirty-seven
98	say word thirty-eight
99	say word thirty-nine
100	say word forty
101	say word forty-one
102	say word forty-two
103	say word forty-three
104	say word forty-four
105	say word forty-five
106	say word forty-six
107	say word forty-seven
108	say word forty-eight
109	say word forty-nine
110	say word fifty
111	say word fifty-one
112	say word fifty-two
113	say word fifty-three
114	say word fifty-four
115	say word fifty-five
116	say word fifty-six
117	say word fifty-seven
118	say word fifty-eight
119	say word fifty-nine
120	say word sixty
121	say a word
122	repeat
123	repeat to confirm
124	training complete
125	reconfigure or erase
126	memory
127	memory empty
128	memory full

129	memory erased
130	word not recognized
131	error
132	training error
133	try again
134	please talk louder
135	please talk softer
136	spoke too soon
137	accepted
138	rejected
139	similar to a prior word
140	similar to word one
141	similar to word two
142	similar to word three
143	similar to word four
144	similar to word five
145	similar to word six
146	similar to word seven
147	similar to word eight
148	similar to word nine
149	similar to word ten
150	similar to word eleven
151	similar to word twelve
152	similar to word thirteen
153	similar to word fourteen
154	similar to word fifteen
155	similar to word sixteen
156	similar to word seventeen
157	similar to word eighteen
158	similar to word nineteen
159	similar to word twenty
160	similar to word twenty-one
161	similar to word twenty-two
162	similar to word twenty-three
163	similar to word twenty-four
164	similar to word twenty-five
165	similar to word twenty-six
166	similar to word twenty-seven
167	similar to word twenty-eight
168	similar to word twenty-nine
169	similar to word thirty
170	similar to word thirty-one
171	similar to word thirty-two
172	similar to word thirty-three
173	similar to word thirty-four
174	similar to word thirty-five
175	similar to word thirty-six
176	similar to word thirty-seven
177	similar to word thirty-eight
178	similar to word thirty-nine

179	similar to word forty
180	similar to word forty-one
181	similar to word forty-two
182	similar to word forty-three
183	similar to word forty-four
184	similar to word forty-five
185	similar to word forty-six
186	similar to word forty-seven
187	similar to word forty-eight
188	similar to word forty-nine
189	similar to word fifty
190	similar to word fifty-one
191	similar to word fifty-two
192	similar to word fifty-three
193	similar to word fifty-four
194	similar to word fifty-five
195	similar to word fifty-six
196	similar to word fifty-seven
197	similar to word fifty-eight
198	similar to word fifty-nine
199	similar to word sixty
200	*low beep*
201	*double low beep*
202	*high beep*
203	*double high beep*
204	*low beep high beep (rising)*
205	*high beep low beep (falling)*
206	*long beep*
207	*high beep*
208	*double high beep*

Optional Prompt List

The speech list contains optional prompts that may be used to create the application's user interface. The Optional Prompt List is accessed via the Say A Prompt Command. Phrases can be substituted or added to this list based on memory availability. See External ROM (Optional), p. 37.

Table 10 – Optional Prompt List

Prompt #	Spoken Prompt Description
1	one
2	two
3	three
4	four

5	five
6	six
7	seven
8	eight
9	nine
10	ten
11	eleven
12	twelve
13	thirteen
14	fourteen
15	fifteen
16	sixteen
17	seventeen
18	eighteen
19	nineteen
20	twenty
21	twenty one
22	twenty two
23	twenty three
24	twenty four
25	twenty five
26	twenty six
27	twenty seven
28	twenty eight
29	twenty nine
30	thirty
31	thirty one
32	thirty two
33	thirty three
34	thirty four
35	thirty five
36	thirty six
37	thirty seven
38	thirty eight
39	thirty nine
40	forty
41	forty one
42	forty two
43	forty three
44	forty four
45	forty five
46	forty six
47	forty seven
48	forty eight
49	forty nine
50	fifty
51	fifty one
52	fifty two
53	fifty three
54	fifty four

55	fifty five
56	fifty six
57	fifty seven
58	fifty eight
59	fifty nine
60	sixty
61	user one
62	user two
63	user three
64	user four
65	accepted
66	rejected
67	a
68	b
69	c
70	d
71	directory a
72	directory b
73	directory c
74	directory d
75	memory set a
76	memory set b
77	memory set c
78	memory set d
79	reconfigure or erase
80	memory
81	memory empty
82	memory full
83	memory erased
84	memory set a erased
85	memory set b erased
86	memory set c erased
87	memory set d erased
88	repeat
89	repeat to confirm
90	training complete
91	you said word one
92	you said word two
93	you said word three
94	you said word four
95	you said word five
96	you said word six
97	you said word seven
98	you said word eight
99	you said word nine
100	you said word ten
101	you said word eleven
102	you said word twelve

103	you said word thirteen
104	you said word fourteen
105	you said word fifteen
106	you said word sixteen
107	you said word seventeen
108	you said word eighteen
109	you said word nineteen
110	you said word twenty
111	you said word twenty-one
112	you said word twenty-two
113	you said word twenty-three
114	you said word twenty-four
115	you said word twenty-five
116	you said word twenty-six
117	you said word twenty-seven
118	you said word twenty-eight
119	you said word twenty-nine
120	you said word thirty
121	you said word thirty-one
122	you said word thirty-two
123	you said word thirty-three
124	you said word thirty-four
125	you said word thirty-five
126	you said word thirty-six
127	you said word thirty-seven
128	you said word thirty-eight
129	you said word thirty-nine
130	you said word forty
131	you said word forty-one
132	you said word forty-two
133	you said word forty-three
134	you said word forty-four
135	you said word forty-five
136	you said word forty-six
137	you said word forty-seven
138	you said word forty-eight
139	you said word forty-nine
140	you said word fifty
141	you said word fifty-one
142	you said word fifty-two
143	you said word fifty-three
144	you said word fifty-four
145	you said word fifty-five
146	you said word fifty-six
147	you said word fifty-seven
148	you said word fifty-seven
149	you said word fifty-eight
150	you said word fifty-nine
151	you said word sixty
152	word not recognized

153	error
154	training error
155	try again
156	please talk louder
157	please talk softer
158	spoke too soon
159	similar to a prior word
160	similar to word one
161	similar to word two
162	similar to word three
163	similar to word four
164	similar to word five
165	similar to word six
166	similar to word seven
167	similar to word eight
168	similar to word nine
169	similar to word ten
170	similar to word eleven
171	similar to word twelve
172	similar to word thirteen
173	similar to word fourteen
174	similar to word fifteen
175	similar to word sixteen
176	similar to word seventeen
177	similar to word eighteen
178	similar to word nineteen
179	similar to word twenty
180	similar to word twenty-one
181	similar to word twenty-two
182	similar to word twenty-three
183	similar to word twenty-four
184	similar to word twenty-five
185	similar to word twenty-six
186	similar to word twenty-seven
187	similar to word twenty-eight
188	similar to word twenty-nine
189	similar to word thirty
190	similar to word thirty-one
191	similar to word thirty-two
192	similar to word thirty-three
193	similar to word thirty-four
194	similar to word thirty-five
195	similar to word thirty-six
196	similar to word thirty-seven
197	similar to word thirty-eight
198	similar to word thirty-nine
199	similar to word forty
200	similar to word forty-one

201	similar to word forty-two
202	similar to word forty-three
203	similar to word forty-four
204	similar to word forty-five
205	similar to word forty-six
206	similar to word forty-seven
207	similar to word forty-eight
208	similar to word forty-nine
209	similar to word fifty
210	similar to word fifty-one
211	similar to word fifty-two
212	similar to word fifty-three
213	similar to word fifty-four
214	similar to word fifty-five
215	similar to word fifty-six
216	similar to word fifty-seven
217	similar to word fifty-eight
218	similar to word fifty-nine
219	similar to word sixty
220	*low beep*
221	*double low beep*
222	*high beep*
223	*double high beep*
224	*low beep high beep (rising)*
225	*high beep low beep (falling)*
226	*long beep*
227	*high beep*
228	*double high beep*
229	*0 msec silence*
230	*20 msec silence*
231	*40 msec silence*
232	*50 msec silence*
233	*75 msec silence*
234	*100 msec silence*
235	*160 msec silence*
236	*320 msec silence*
237	*500 msec silence*
238	*1000 msec silence*
239	*2000 msec silence*
240	*4000 msec silence*

Training/Recognition Flow Charts

Figure 10 - Training (1 of 2)

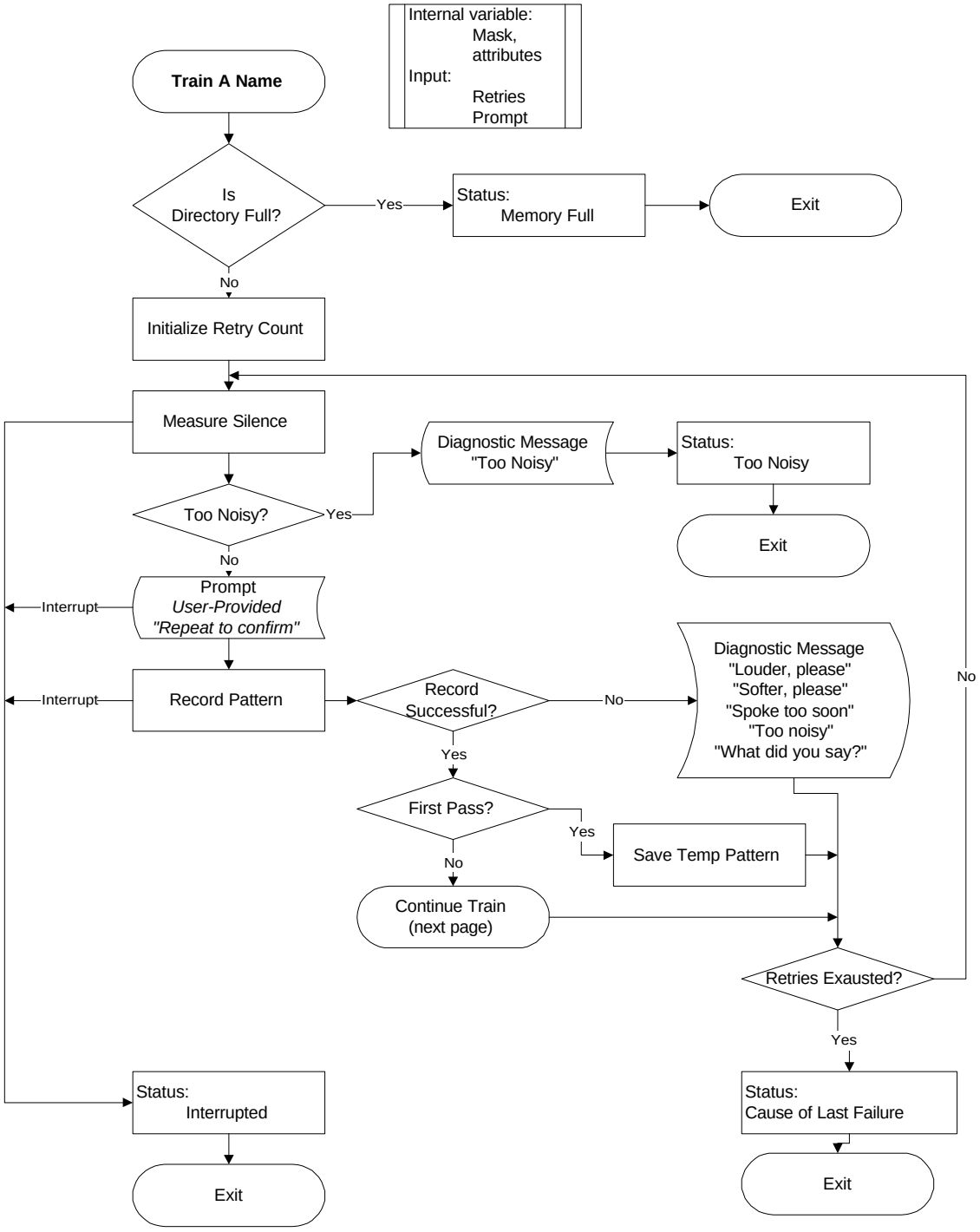


Figure 11 - Training (2 of 2)

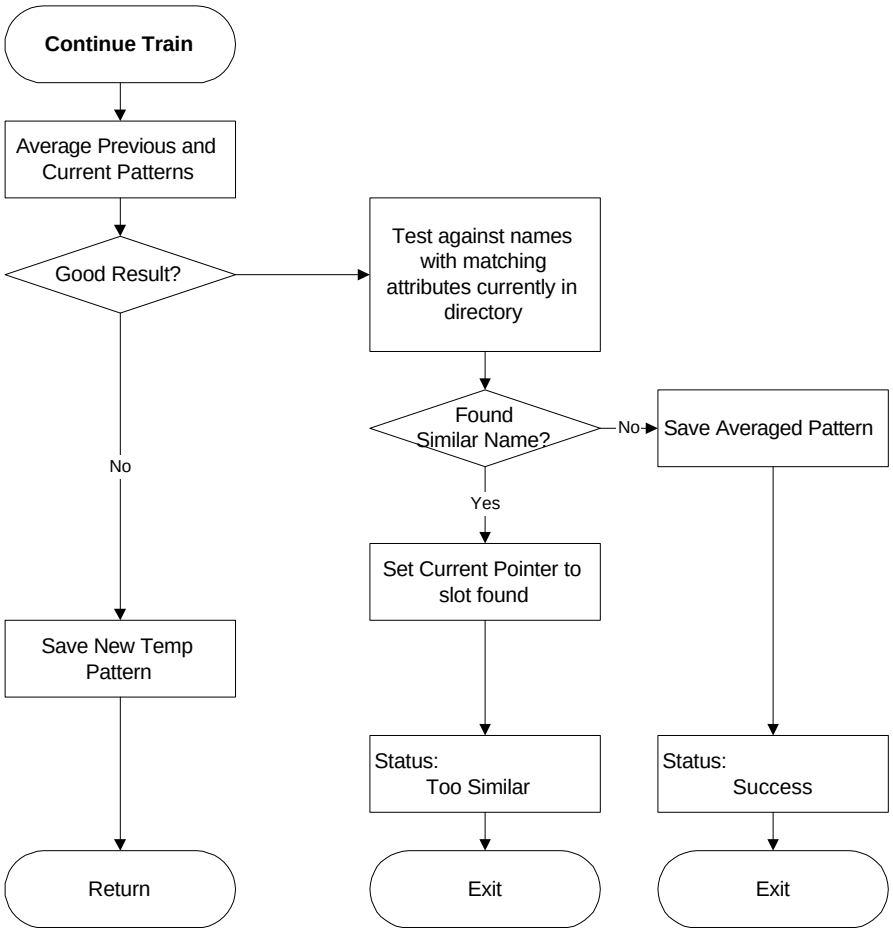
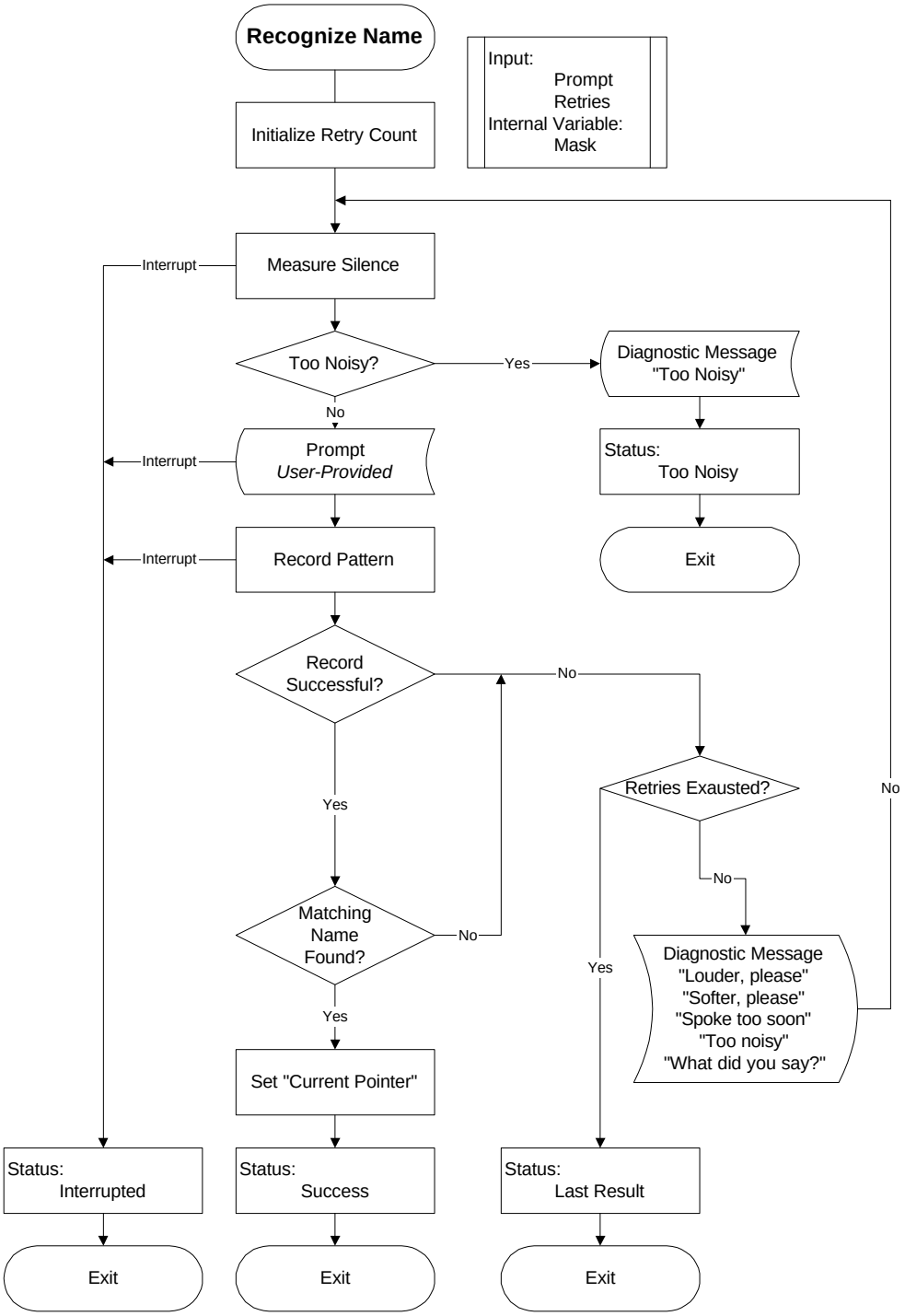


Figure 12 - Recognition Flowchart



Voice Direct IC Specifications

IC Pin Descriptions

Figure 13 - 64-pin MQFP pin assignments

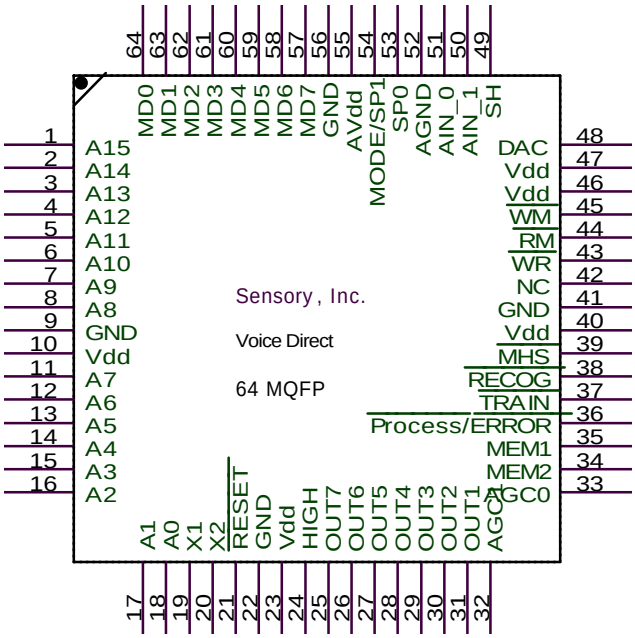


Table 11 - Pin Identification

Name	Pin	Description	I/O
A[15:0]	1-8, 11-18	External ROM Memory Address Bus	O
AGC0	33	AGC control 0. The Voice Direct controls the amplifier gain with this signal.	O
AGC1	32	AGC control 1	O
AGND	52	Analog Ground. For noise reasons, analog and digital grounds should connect together only at the RSC-164.	-
AIN0	51	Analog In, low gain. (range AGND to AVDD/2.)	I
AIN1	50	Analog In, hi gain (8X input amplitude of AIN0, same range)	I
AVDD	55	Analog Supply Voltage. For noise reasons, keep this supply independent of digital circuitry.	-
GND	9, 22, 41, 56	Digital Ground	-
MD[7:0]	57-64	External ROM Memory Data Bus	I/O
SCL	35	Serial Clock for Serial EEPROM.	O

SDA	34	Serial Data for Serial EEPROM.	O
MODE/SP1	54	The MODE pin is connected via a 100K resistor to VDD or GND to select External Host or Stand Alone Mode. This pin is also Speaker Connect1. A 32-Ohm speaker can be connected directly to this pin. (Input for MODE at power up.)	O
MHS	39	Master handshake pin	I
RECOG	38	Trigger recognition pin	I
TRAIN	37	Trigger training pin	I
PROCESS/ ERROR	36	Process pin in external host mode/error pin in stand-alone mode	O
-RESET	21	Reset	I
-RM	44	Read Memory Strobe. Can control -OE pin of External ROM.	O
SH	49	Sample and Hold. Connect a 470 pF capacitor from here to AGND.	I
DAC	48	DAC output	O
SP0	53	Speaker Connect0. A 32-Ohm speaker can be connected directly to this pin.	O
VDD	10,23,36,40, 46, 47	Digital Supply Voltage (core).	-
-WR	43	Write Result. When a recognition sequence is complete the chip will place the result on the memory data bus MD[7:0] and strobe this signal to latch the result into external devices.	O
X1, X2	19,20	Crystal connect. A 14.312 MHz crystal is connected to these pins.	O, I
NC	*42*		

Absolute Maximum Ratings

- Minimum voltage on any pin V_{ss}-0.6V
- Maximum voltage on any pin V_{dd}+0.6V
- Operating temperature (T_O) -20°C to +70°C
- Soldering temperature 260°C for 10 sec
- Maximum voltage 7.5V
- Power dissipation 1W
- Storage Temperature -65°C to +150°C
- Minimum Operating Voltage 3.5V
- Maximum Operating Voltage 5.0V

> ***WARNING**: Stressing Voice Direct beyond the "Absolute Maximum Ratings" may cause permanent damage. These are stress ratings only. Operation beyond the "Operating Conditions" is not recommended and may affect device reliability.*

DC Characteristics

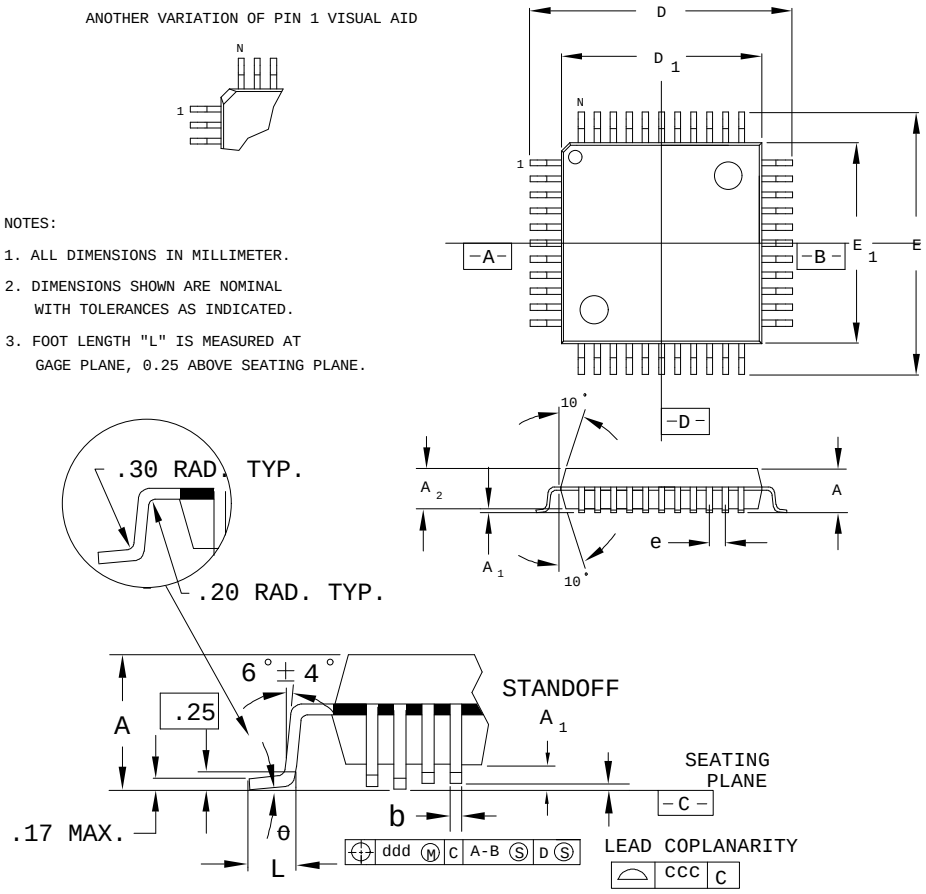
(T_O = -20°C to +70°C, Vdd = 5V)

SYMBOL	PARAMETER	MIN	TYP	MAX	UNITS	TEST CONDITIONS
V_{IL}	**Input Low Voltage** XML, XMH, TE1, I/O pins XI1 RESET	 -0.1 -0.1 -0.1		 0.75 0.2 V_{DD} 0.60	 V V V	
V_{IH}	**Input High Voltage** XML, XMH, TE1, I/O pins XI1 RESET	 2.5 0.7 V_{DD} 3.0		 Vdd+0.3 Vdd+0.3 Vdd+0.3	 V V V	
V_{OL}	**Output Low Voltage** I/O pins			 0.5	 V	 I_{OL}= 2.0 mA
V_{OH}	**Output High Voltage** I/O pins	 4.0			 V	 I_{OH}= -2.0 mA
I_{IL}	**Input Leakage Current** XML, XMH, TE1, I/O pins XI1 RESET AIN0, AIN1, SH		 <1 <1 <1 <1	 5 5 5 5	 uA uA uA uA	 $V_{ss}<V_{pin}<V_{dd}$ $V_{ss}<V_{pin}<V_{dd}$ $V_{ss}<V_{pin}<V_{dd}$ $V_{ss}<V_{pin}<V_{dd}$
C_I	**Input Pin Capacitance**		6		pF	
I_{cc1}	**Supply Current, Operating**		7	20	mA	Hi-Z outputs
I_{cc2}	**Supply Current, Quiescent**		300	600	uA	Hi-Z outputs
R_{PU}	**Internal Pull-up Resistance, I/O pins**		4.5, 200,HiZ		kOhms	Software selected

Analog Characteristics

SYMBOL	PARAMETER	MIN	TYP	MAX	UNITS	TEST CONDITIONS
V_{ILA}	AIN0, AIN1	-0.5		0	V	
V_{IHa}	AIN0, AIN1		Avdd/ 2		V	
C_{SH}	SH- capacitance		470 ±10%	pF	V	I_{OL}= 2.0 mA
V_{DO}	DACOUT Voltage Swing	Agnd		Avdd	V	

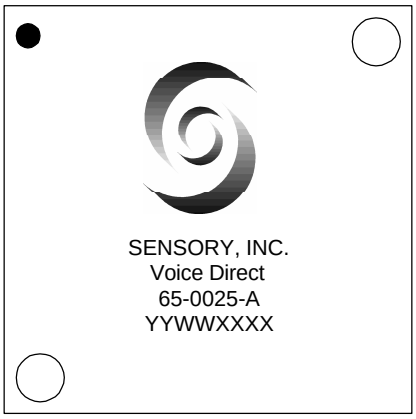
IC Metric Drawings



		Body + 3.20 mm FOOTPRINT
Package Thickness		2.00
Dimensions	Tolerance	64L
A	M	2.35
A_1	M	0.25
A_2	+	2.00
D	±	17.20
D_1	±	14.00
E	±	17.20
E_1	±	14.00
L	+	.88
e	BASIC	.80

b	±.05	.35
θ		0-7°
ddd		.20 NOM
ccc	MAX.	.10

Marking



Pin #1 of: .70 mm
Height of Top Mark: Top Left
Height of Character Logo: 4.475mm
"YYWWXXXX": YYWW - Date Code (year and week)
XXXX - Sensory Identifier

IC Packaging Specification

Package: 64L QFP (14x14x2.0)
Tray: Peak Thin Bakeable, Black, Static Dissipative
84 positions.
Max Bake Temperature: 180° C.
DWG #: ND 1414 2.0 0614 8 Rev A.

Bundle (Typical): 5 + 1 (420 parts)

Packaging Description

Each bundle consists of 5 full trays plus 1 empty tray. Each bundle is strapped with 3 nylon straps. The strappings are approximately at the ¼, ½, and ¾ positions along the long length of the tray. After the trays are bundled, each bundle is placed in a conductive bag with silica gel. The bag is then sealed after air is evacuated from the bag. A label affixed to the outside of the bag contains the following information:

Customer	(S039)*
P.O. No.	
P/Traveller	
Package	(64 LEAD QFP)*
Device/Type	(Voice Direct)*
Die Lot No.	
Date Code	
Quantity	

*Please Note: Information in ( ) does not change.

The drypacked bags will be individually wrapped in plastic sealed air bubble wrap and placed in a box appropriate to the size of the shipment (8400 = 10 drypacked bags). This box will then be placed into another shipping box and will be packaged with padding (i.e., foam popcorn, sealed air bubble wrap) and labeled accordingly. The shipping box will then be sealed with 2" thick fiberglass reinforced tape.

Dry Bag Recommendations

This device has been qualified to meet JEDEC Moisture Sensitivity Level 3 requirements. Level 3 specifies that the exposure time at the customer site after opening the units is 168 hours in an environment less than 30C and 60% RH.

Manufacturing Information

Manufacturer: Taiwan Semiconductor Manufacturing Corporation (TSMC)
Country of Origin: Taiwan

Voice Direct Module Specifications

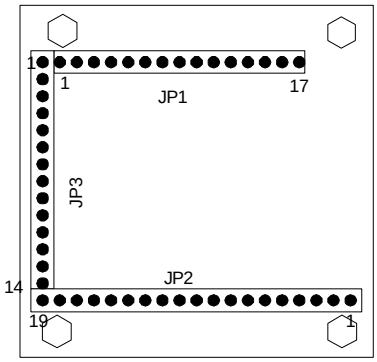
The Voice Direct Module is a pre-configured circuit that integrates required external components for prototype development. The Voice Direct Module comes with all of the required external components except for a microphone and speaker. The Voice Direct Module integrates the following:

- Preamplifier
- Microchip Serial EEPROM 24C65 (8K)
- Oscillator (14.32 MHz)
- A slot for an optional external ROM (for customized speech or non-English language)

The Voice Direct Module is 2 inches x 2 inches.

What follows are Voice Direct Module pin descriptions, schematics and a bill of materials.

Module Pin Descriptions



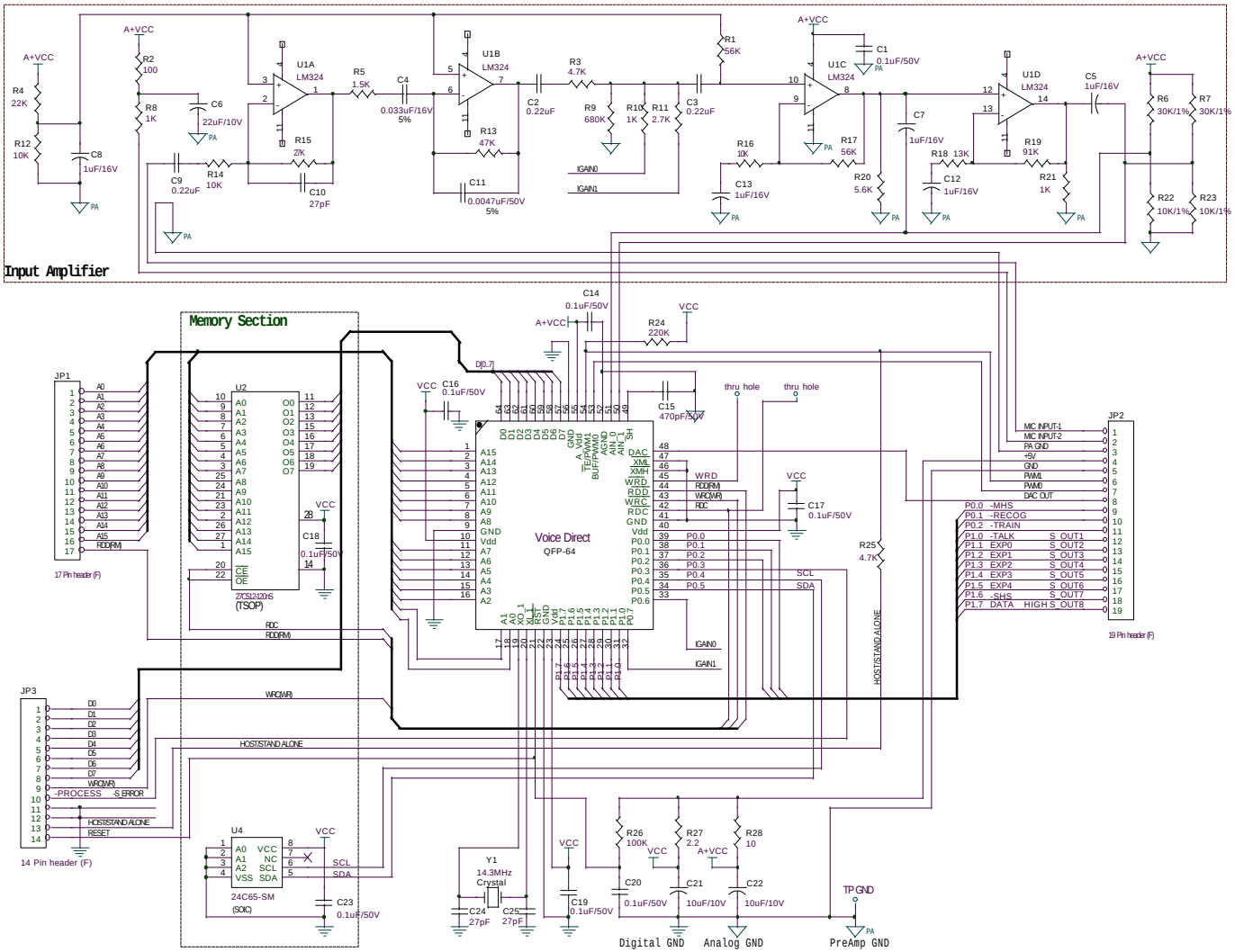
Voice Direct™ Module

Table 12 - Pin Identification

Name	Module	Description	I/O
	JP1 - 1-16, JP2 - 1-8, 10, 13	Unused	-
-RDD	JP1 – 17	External Data Read Strobe	O
PREAMP IN	JP2 - 1	Microphone Input Connection	I
MIC BIAS	JP2 - 2	Mic Bias (Electret microphone)	I
AGND	JP2 - 3,5	Analog Ground. For noise reasons, analog and digital grounds should connect together only at the Voice Direct (VDI)	-
+5V	JP2 - 4	5 Volt (+) Power Supply Connection	-
PWM1	JP2 - 6	Pulse Width Modulator Output1 (multiplexed)	O
PWM0	JP2 - 7	Pulse Width Modulator Output0	O
DACOUT	JP2 - 8	Analog Output (unbuffered).	O
MHS	JP2 - 9	Master Handshake. Driven by host.	I
- RECOG	JP2 – 10	Recognition sensitivity selection and activate recognition	
P1.2, P1.3, P1.4,, P1.5	JP2 -14, 15, 16, 17	Reserved Port Pins, No Connections should be made here	-
- TRAIN	JP2 – 11	Training sensitivity selection and activate training	
-TALK/SOUT1	JP2 - 12	Audio talk signal. Active low during speech synthesis	O
EXP 0-4/SOUT 2-6	JP2–13, 14, 15, 16, 17	IO expansion port 0-4 *or* Stand Alone Mode output port 2 -6	
SHS/SOUT7	JP2 – 18	Slave Handshake. Driven by Voice Direct or Stand Alone Mode output port 7.	O
DATA/HIGH/ SOUT8	JP2 – 19	Serial Data between Master and Slave. Bi-directional *or* Stand Alone Mode output high *or* output port 8	I/O

-WRC	JP3 – 9	External Code Write Strobe	O
PROCESS/S_ERROR	JP3 – 10	Active low when VDI processing command *or* Stand Alone Mode error signal	
GND	JP3 - 11,12	Digital Ground, CPU core (pins 1 and 33) and I/O (pins 18and 52)	-
HOST/ STANDALONE	JP3 – 13	VDI mode selection.	
-RESET	JP3 - 14	Reset	I

Figure 14 – Module Schematic



This page is intentionally blank.

Bill of Materials for Voice Direct Module

Table 13 – Bill of Materials

Qty	Reference	Part
8	C1,C14,C16,C17,C19,C20, C23,C18	0.1uF 50V
3	C2,C3,C9	0.22uF 16v
1	C4	0.033uF 16V
5	C5,C7,C8,C12,C13	1uF/16V
1	C6	22uf/10v
3	C10,C24,C25	27pF 50v
1	C11	0.0047uF 50V
1	C15	470pF 50V
2	C22,C21	10uF/10V
1	JP1	17P F header
1	JP2	19P F header
1	JP3	14P F header
2	R1,R17	56K
1	R2	100
2	R3,R25	4.7K
1	R4	22K
1	R5	1.5K
2	R6,R7	30K/1%
3	R8,R10,R21	1K
1	R9	680K
1	R11	2.7K
3	R12,R14,R16	10K
1	R13	47K
1	R15	27K
1	R18	13K
1	R19	91K
1	R20	5.6K
2	R23,R22	10K/1%
1	R24	220K
1	R26	100K
1	R27	2.2
1	R28	10
1	U1	LM324M
1	U2	
1	U3	RSC-164v2
1	U4	24C65-SM (SOIC)
1	Y1	14.3MHz
1	U2	27C512 (TSOP)