

Emic 2 Text-to-Speech Module (#30016)

Designed in conjunction with Grand Idea Studio (www.grandideastudio.com), the Emic 2 Text-to-Speech Module is an unconstrained, multi-language voice synthesizer that converts a stream of digital text into natural sounding speech output. Using the universally recognized DECTalk text-to-speech synthesizer engine, Emic 2 provides full speech synthesis capabilities for any embedded system via a simple command-based interface.

Features

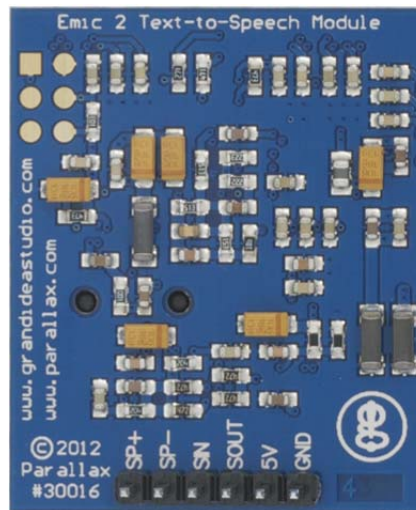
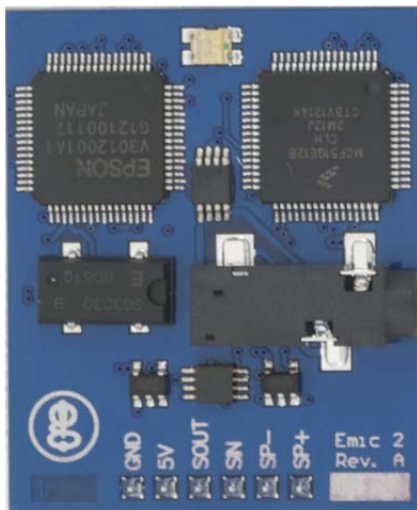
- High-quality speech synthesis for English and Spanish languages
- Nine pre-defined voice styles comprising male, female, and child
- Dynamic control of speech and voice characteristics, including pitch, speaking rate, and word emphasis
- Industry-standard DECTalk text-to-speech synthesizer engine (5.0.E1)
- On-board audio power amplifier and 1/8" (3.5 mm) audio jack
- Single row, 6-pin, 0.1" header for easy connection to a host system

Key Specifications

- Power requirements: +5 VDC, 30 mA idle, 46-220 mA active (depending on speech parameters and output load)
- Communication: asynchronous 9600 bps serial
- Operating temperature: -20 to +70 °C (-4 to +158 °F)
- Dimensions: 1.25" W x 1.5" L x 0.37" H (3.17 W x 3.81 L x 0.94 H cm)

Application Ideas

- Reading Internet-based data streams (such as e-mails or Twitter feeds)
- Conveying status or sensor results from robots, scientific equipment, or industrial machinery
- Language learning or speech aids for educational environments



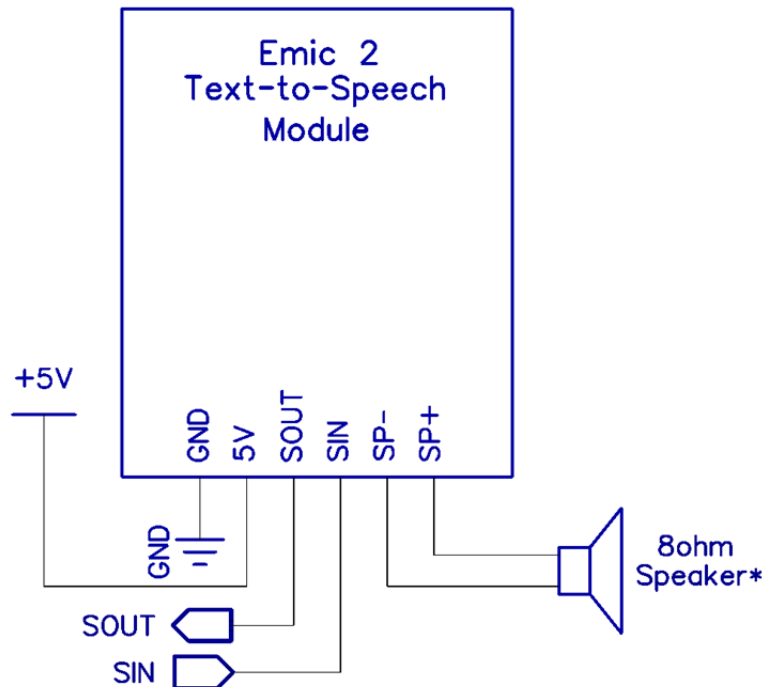
Connections

Emic 2 interfaces to a host microcontroller or computer system using only four connections (GND, 5V, SOUT, SIN). Additional connections (SP+, SP-) are available for direct interfacing to an 8Ω speaker. A 1/8" (3.5mm) audio jack provides a single-ended, monaural output for easy connection to headphones, amplified speakers, or other audio equipment.

Pin	Pin Name	Type	Function
1	GND	G	System ground. Connect to power supply's ground (GND) terminal.
2	5V	P	System power, 5 VDC input.
3	SOUT	O	Serial output to host. 5 V TTL-level interface, 9600 bps, 8 data bits, no parity, 1 stop bit, non-inverted.
4	SIN	I	Serial input from host. 3.3 V to 5 V TTL-level interface, 9600 bps, 8 data bits, no parity, 1 stop bit, non-inverted.
5	SP-	O	Differential audio amplifier output, bridge-tied load configuration, negative side. Connect directly to 8 Ω speaker.
6	SP+	O	Differential audio amplifier output, bridge-tied load configuration, positive side. Connect directly to 8 Ω speaker.

Type: I = Input, O = Output, P = Power, G = Ground

Use the following example circuit for connecting the Emic 2 Text-to-Speech Module:



*Note: For audio output, a connection needs to be made to either SP+/SP- or the 1/8" audio jack. Audio quality may be affected if both outputs are used at the same time.

Usage

Emic 2 is controlled by the host via a serial communications interface. To use, simply send the desired command to Emic 2 and listen for audio output from the SP+/SP- speaker connection or 1/8" audio jack.

The serial interface is configured for 9600 bps, 8 data bits, no parity, 1 stop bit (8N1). When Emic 2 is ready to receive commands, it will send a ":" to the host. It will then wait in an idle state until it receives a valid command, at which time it performs the command and returns any command-specific response. Emic 2 will return a "?" upon receiving an invalid command.

On power-up, Emic 2 loads its default text-to-speech settings consisting of voice type, audio volume, speaking rate, language, and parser. These settings can be configured by the user to vary the audio output. See the Command Set section below for more details.

Status Indicator

A visual indication of Emic 2's operating state is given with the on-board light-emitting diode (LED):

1. **Green:** Idle state. Waiting for a valid command to be sent by the host.
2. **Red:** Active state. For example, during a text-to-speech conversion.
3. **Orange (Solid):** Initialization state. Occurs on power-up only. Emic 2 takes approximately three seconds to properly initialize on power-up before it is ready to receive commands.
4. **Orange (Blinking):** Error state. Emic 2 has malfunctioned due to an on-board communication error. If a power cycle of Emic 2 does not remedy the situation, please contact Parallax technical support for further assistance.

If the LED is OFF, Emic 2 may not be receiving power.

Command Set

All commands are ASCII-based printable characters and are not case-sensitive (upper case and lower case will both work). Each command must be terminated with a CR or LF.

Sx	Convert text-to-speech: x = message (1023 characters maximum)
Dx	Play demonstration message: x = 0 (Speaking), 1 (Singing), 2 (Spanish)
X	Stop playback (while message is playing)
Z	Pause/un-pause playback (while message is playing)
Nx	Select voice: x = 0 to 8
Vx	Set audio volume (dB): x = -48 to 18
Wx	Set speaking rate (words/minute): x = 75 to 600
Lx	Select language: x = 0 (US English), 1 (Castilian Spanish), 2 (Latin Spanish)
Px	Select parser: x = 0 (DECTalk), 1 (Epson)
R	Revert to default text-to-speech settings
C	Print current text-to-speech settings
I	Print version information
H	Print list of available commands

Command Details

Sx: Convert text-to-speech

Convert the passed text string into synthesized speech. The text string is limited to 1023 characters and should terminate on a clause or sentence boundary as indicated by a full stop '.' or comma ',' punctuation mark. If the text is longer than the allowable limit, it will be truncated and may result in unintelligible speech output.

Emic 2 expects characters that conform to the ISO-8859-1 Latin character set (http://en.wikipedia.org/wiki/ISO/IEC_8859-1). See the Special Characters section (p. 9) for details on entering accents, foreign characters, and symbols.

The audio will be output from both the SP+/SP- speaker connection and 1/8" audio jack. The LED will remain RED while the text-to-speech message is being played.

Example:

```
:SHello there! My name is Emic 2. Nice to meet you.  
<audio output>  
:
```

Dx: Play demonstration message

Play one of Emic 2's built-in demonstration messages:

- 0: English Introduction
- 1: Singing "Daisy Bell" (http://en.wikipedia.org/wiki/Daisy_Bell)
- 2: Spanish Introduction

Note that each demonstration message is fixed with specific voice, audio volume, speaking rate, language, and parser settings and cannot be modified by the user. All user-configured settings will be saved prior to demonstration playback and restored afterwards.

See the Sample Text Strings section (p. 9) for the actual text strings used for these demonstration messages.

Example:

```
:D0  
<audio output>  
:
```

X: Stop playback (while message is playing)

Immediately stop the currently playing text-to-speech message. This command is only valid while a message is playing.

Example:

```
:D0  
<audio output>  
X  
:
```

Z: Pause/unpause playback (while message is playing)

Immediately pause or unpause the currently playing text-to-speech message. Emic 2 will respond with a "." indicating that the command has successfully been received. While the playback is paused, the LED will remain RED. This command is only valid while a message is playing.

Example:

```
:D0
<audio output>
Z.
<playback paused>
Z.
<audio output>
:
```

Nx: Select voice

Select the speaking voice:

- 0: Perfect Paul (Paulo)
- 1: Huge Harry (Francisco)
- 2: Beautiful Betty
- 3: Uppity Ursula
- 4: Doctor Dennis (Enrique)
- 5: Kit the Kid
- 6: Frail Frank
- 7: Rough Rita
- 8: Whispering Wendy (Beatriz)

Each voice has a different baseline amplitude. As such, your volume settings may need to be adjusted to suit your particular application.

This setting will remain in effect until another value is entered or Emic 2 is powered off. Upon power-up, the default value is 0 (Paul).

Example:

```
:N5
:
```

Vx: Set audio volume (dB)

Set the audio output volume in decibels from -48 (softest) to 18 (loudest).

This setting will remain in effect until another value is entered or Emic 2 is powered off. Upon power-up, the default value is 0.

Example:

```
:V-10
:
```

Wx: Set speaking rate (words/minute)

Set the speaking rate in words per minute from 75 (slowest) to 600 (fastest).

This setting will remain in effect until another value is entered or Emic 2 is powered off. Upon power-up, the default value is 200.

Example:

```
:W150  
:
```

Lx: Select language

Select the language and/or dialect to be used for text-to-speech conversion:

- 0: US English
- 1: Castilian Spanish
- 2: Latin Spanish

This setting will remain in effect until another value is entered or Emic 2 is powered off. Upon power-up, the default value is 0 (US English).

Example:

```
:L2  
:
```

Px: Select parser

Select the text parsing engine to be used during text-to-speech conversion:

- 0: DECTalk
- 1: Epson

See the Parser Selection section (p. 8) for usage information and parser details.

This setting will remain in effect until another value is entered or Emic 2 is powered off. Upon power-up, the default value is 1 (Epson).

Example:

```
:P0  
:
```

R: Revert to default text-to-speech settings

Resets the user-configurable text-to-speech settings to their default, power-up configuration:

Voice = 0 (Paul)
Volume = 0 dB
Rate = 200 words/minute
Language = 0 (US English)
Parser = 1 (Epson)

Example:

```
:R
:
```

C: Print current text-to-speech settings

Displays the current values of the user-configurable text-to-speech settings.

Example:

```
:C
Emic 2 Text-to-Speech Settings:
Voice = 3 (Ursula)
Volume = 0 dB
Rate = 160 words/minute
Language = 0 (US English)
Parser = 0 (DECTalk)
:
```

V: Print version information

Displays version information for the Emic 2 Text-to-Speech Module. This data is useful for troubleshooting and debugging.

- EMIC FW: Firmware version (major.minor)
- S1V30120 HW and FW: Hardware and firmware version (major.minor.revision) of the Epson S1V30120 Voice Guidance IC used on-board Emic 2

Example:

```
:V
Parallax Emic 2 Text-to-Speech Module
Designed by Grand Idea Studio [www.grandideastudio.com]
Manufactured and distributed by Parallax [support@parallax.com]
EMIC FW = 1.0
S1V30120 HW = 4.2
S1V30120 FW = 2.1.6 [0x2551_25]
:
```

H: Print list of available commands

Lists all available commands for the Emic 2 Text-to-Speech Module.

Example:

```
:H
Emic 2 Command List:
Sx Convert text-to-speech: x = message (1023 byte maximum)
Dx Play demonstration message: x = 0 (Speaking), 1 (Singing), 2 (Spanish)
< more commands listed, but not shown in this example >
:
```

Parser Selection

Emic 2 provides a choice of text parsing engines: Epsion or DECTalk.

Both will process incoming text strings and generate synthesized speech, but there are differences in control and customization of the resulting output. Each parser's specific functionality is incompatible with the other. Choosing which parser to use ultimately depends on what sort of speech output you desire for a particular text string.

See the Sample Text Strings section (p. 9) using specific features of the Epsion and DECTalk parsers.

Epsion Parser

The Epsion parser is the default setting upon Emic 2 power-up. It allows on-the-fly, dynamic changes of emphasis, pitch, voice selection, and speaking rate to take place within a text using embedded mark-up control symbols:

\/	Decrease pitch
\/	Increase pitch
>>	Increase speaking rate
<<	Decrease speaking rate
—	Emphasize the next word
##	Whisper the next word
:-)x	Select voice (x = 0-8) (See the "Nx" command in the Command Set section (p. 5) for corresponding voice names)

DECTalk Parser

The DECTalk parser is intended for advanced users and allows the finest control and customization of speech output by providing direct access to the internal parameters of the DECTalk 5.0.E1 text-to-speech synthesizer engine.

Dynamic customization using the DECTalk parser requires the correct usage of specialized commands and their associated parameters. Incorrect usage of the DECTalk commands may result in improper or unintelligible speech output. Please refer to the Epsion/Fonix DECTalk 5.0.E1 User Manual available on the Emic 2 product page for full command details and phonetic symbols.

Only a subset of the DECTalk commands are supported by the Emic 2 Text-to-Speech Module:

[:comma]	Set the length of a comma pause (in milliseconds)
[:dv]	Customize voice parameters (save option not supported)
[:mode]	Set how text is processed/parsed (no e-mail parsing supported)
[:name]	Set the current speaking voice (numbers only: 0-8)
[:period]	Set the length of a period pause (in milliseconds)
[:phoneme]	Enable phonemic interpretation of subsequent text
[:pitch]	Modify the pitch of uppercase letters
[:pronounce]	Set the type of pronunciation for the subsequent word
[:punct]	Set how punctuation marks are handled
[:rate]	Set speaking rate
[:say]	Specify when speaking begins
[:skip]	Skip a selected part of text pre-processing
[:sync]	Makes a command synchronous to allow it to be processed before synthesis continues

Spanish Introduction

(Built-in demonstration message #2) (Language must be set to Castilian Spanish)

:-)0 Hola. Me llamo Emic numero 2. Ahora puedo hablar espa\xF1o! Soy la pr\xF3xima generacion de texto a voz modulo creado por Grand Idea Studio. Yo puedo ##susurrar ##en ##voz ##muy ##baja. Yo puedo cambiar a una de las nueve voces. Por ejemplo, de Paulo :->1 a Francisco :->4 a Enrique :->8 a Beatriz. :->0 Tambien puedo /\ / aumentar mi tono. /\ / y aumentar mi tono otra vez. >> Entonces hablar mas r\xE1pido. >> y aun mas r\xE1pido >> y aun mas r\xE1pido otra vez. <<<<<< \V\V\V Y luego volver a la normalidad.

DECTalk Parser

Robotic monotone

[:rate 200][:n0][:dv ap 90 pr 0] All your base are belong to us.

Singing "Daisy Bell"

(Built-in demonstration message #1), http://en.wikipedia.org/wiki/Daisy_Bell

[:phone arpa speak on][:rate 200][:n0][dey<650,22> ziy<600,19> dey<650,15> ziy<600,10> gih<200,12>v miy<200,14> yurr<200,15> ae<400,12> nsax<200,15>r duw<750,10> _<400,10>][:n0][ay<600,17>m hxae<500,22>f kr ey<650,19> ziy<600,15> ao<200,12>ll fao<200,14>r dhax<200,15> llah<400,17>v ao<200,19>v yu<750,17> _<400,17>][:n0][ih<200,19>t wow<200,20>nt biy<200,19> ax<200,17> stay<500,22>llih<200,19>sh mae<350,17>rih<400,15>jh<150,15>][:n0][ay<200,17> kae<400,19>nt ax<200,15>fow<400,12>rd ax<200,15> kae<350,12>rih<400,10>jh<150,10>][:n0][bah<200,10>t yu<500,15>d lluh<200,19>k swiy<400,17>t ah<200,10>p ao<500,15>n dhax<200,19> siy<200,17>t ao<200,17>v ah<200,19> bay<200,22>six<200,19>kel<200,15> bih<400,17>llt fao<200,10>r tuw<800,15>][:n0]

Singing "Happy Birthday"

[:phone arpa speak on][:rate 200][:n0][hxae<300,10>piy<300,10> brrrx<600,12>th<100>dey<600,10>tuw<600,15> yu<1200,14> _<300> hxae<300,10>piy<300,10> brrrx<600,12>th<100>dey<600,10> tuw<600,17>yu<1200,15> _<300> hxae<300,10>piy<300,10> brrrx<600,22>th<100>dey<600,19> dih<600,15>r eh<600,14>m<100,12>ih<350,12>k_<120> _<300> hxae<300,20>piy<300,20> brrrx<600,19>th<100>dey<600,15> tuw<600,17> yu<1200,15>][:n0]

Singing "The Star-Spangled Banner"

(The text exceeds the maximum allowable 1023 characters per message, so it must be split into two messages)

[:phone arpa speak on][:rate 100][:n0][ow<200,18>ow<200,15> sey<400,11> kae<400,15>n yu<400,18>w siy<600,23> _<300> bay<350,27> dhah<50,25> dao<400,23>nz rr<400,15>ll iy<400,17> llay<600,18>t _<300> wah<200>t sow<200> praw<600,27>dliiy<200,25> wiy<400,23> hxey<400,22>eld _<300> ae<300,20>t dhah<100,22> tway<400,23> llay<400>ts llae<400,18>st gliiy<400,15>m iy<200,11>nx _<300> hxuw<300,18>z brao<100,15>d stray<400,11>ps ae<400,15>nd bray<400,18>t stah<600,23>rz _<300> thruw<300,27> dhah<100,25> peh<400,23> rrel<400,15> ah<400,17>s fay<600,18>t _<300> ow<200,18>r dhah<200,18> rrae<600,27>mpah<200,25>rts wiy<400,23> waa<600,22>cht wrw<300,20> sow<100,22> gae<400,23>llah<400>ent lliiy<400,18> striy<400,15>miy<200,11>nx _<300>][:n0]

[:phone arpa speak on][:rate 100][:n0][ae<300,27>nd dhah<100> raa<400> keh<400,28>ts r eh<400,30>d glley<700>r _<300> dhah<100,28> baa<400,27>mz brr<400,25>stih<400,27>nx ih<400,28>n ey<600>r _<300> gey<400>v pruw<600,27>f thruw<200,25> dhah<400,23> nay<900,22>t dhae<300,20>d aw<100,22>rr flae<400,23>g wah<400,15> stih<400,17>ll dheh<600,18>r _<300> ow<400> sey<400,23> dah<400>z dhae<200,23>ae<200,22>t stah<400,20>r spae<400>ngel<400>d bae<400,25>nrr<200,28>rr<200,27> yxeh<200,25>eh<200,23>t wey<800,23>ey<150,25>ey<150,23>ey<1200,22>v _<900> fow<300,18>rdhah<300> llae<900,23>ae<400,25>nd ah<300,27>v dhah<300,28> friy<1000,30>iy<1000,35> _<900> ae<300,23>nd dhah<300,25> hxow<1000,27>m _<600> ah<300,28>v dhah<1000,25> brrey<1500,23>v _<900>][:n0]

Certain characters are not spoken

The following characters are not spoken by Emic 2 during a text-to-speech conversion. If you'd like any of these specific characters to be spoken, simply use the text equivalent in your text string (For example, "less than" instead of "<").

Character	ISO-8859-1 Hex Code	Description
(0x28	Left parenthesis
<	0x3C	Less than
[0x5B	Left square bracket
{	0x7B	Left curly bracket
	0x7C	Vertical bar
	0xA0	No-break space
¤	0xA4	Currency sign
‡	0xA6	Broken bar
¨	0xA8	Diaeresis
¬	0xAC	Not sign
SHY	0xAD	Soft hyphen

Character	ISO-8859-1 Hex Code	Description
®	0xAE	Registered sign
–	0xAF	Macron
´	0xB4	Acute accent
¸	0xB8	Cedilla
º	0xBA	Masculine ordinal indicator
Ð	0xD0	Latin capital letter eth
×	0xD7	Multiplication sign
Þ	0xDE	Latin capital letter thorn
÷	0xF7	Division sign
þ	0xFE	Latin small letter thorn
ÿ	0xFF	Latin small letter y with diaeresis

Electrical Characteristics

At VCC = +5.0V and TA = 25°C unless otherwise noted

Parameter	Symbol	Test Conditions	Specification			Unit
			Min.	Typ.	Max.	
Supply Voltage	V _{CC}	---	3.6	5.0	5.5	V
Supply Current, Idle	I _{IDLE}	---	---	30	---	mA
Supply Current, Active	I _{CC}	Demo #0, R _L = 8 Ω	---	---	46	mA
		Volume = -38dB	---	---	65	
		Volume = 0dB	---	---	220	
Volume = +18dB	---	---	---	---	---	
SP+/SP- Continuous Average Output Power	P _O	R _L = 8 Ω	---	---	300	mW
Audio Jack Impedance	J _L	No load	---	10	---	Ω
Audio Jack Peak-to-Peak Output Voltage	J _{VPP}	Demo #0, No load	---	---	0.85	V
		Volume = -38dB	---	---	1.75	
		Volume = 0dB	---	---	4.1	
Audio Jack Peak Output Voltage	J _{VPK}	Demo #0, No load	---	---	0.4	V
		Volume = -38dB	---	---	1.0	
		Volume = 0dB	---	---	2.5	
Volume = +18dB	---	---	---	---	---	

Absolute Maximum Ratings

Condition	Value
Operating Temperature	-20°C to +70°C (-4°F to +158°F)
Storage Temperature	-55°C to +125°C (-67°F to +257°F)
Supply Voltage (V _{CC})	+6.0V
Ground Voltage (V _{SS})	0V

NOTICE: Stresses above those listed under *Absolute Maximum Ratings* may cause permanent damage to the device. This is a stress rating only and functional operation of the device at those or any other conditions above those indicated in the listings of this specification is not implied. Exposure to maximum rating conditions for extended periods may affect device reliability.

Open Source Files and Example Code

The following engineering materials are released as open source under a Creative Commons Attribution 3.0 United States license (<http://creativecommons.org/licenses/by/3.0/us/>), allowing free distribution and reuse. The materials are posted on the Emic 2 product page; search "30016" at www.parallax.com:

- Schematic
- Bill-of-Materials
- Assembly Drawing
- Example Code for BASIC Stamp 2, Propeller P8X32A, and Arduino Uno

Revision v1.1: Updated note below connection diagram and added row 6 to Connections table; page 2.