Catalog Number 276-1786

ARCHER

TECHNICAL DATA

AN EXCLUSIVE RADIO SHACK SERVICE TO THE EXPERIMENTER

CTS256A-AL2 CODE-TO-SPEECH PROCESSING CHIP

Features:

- Unlimited vocabulary
- Utilizes letter-to-sound rules
- Serial or parallel interface
- Outputs directly to the SP0256A-AL2 Speech Processor
- User can add to existing letter-to-sound rules

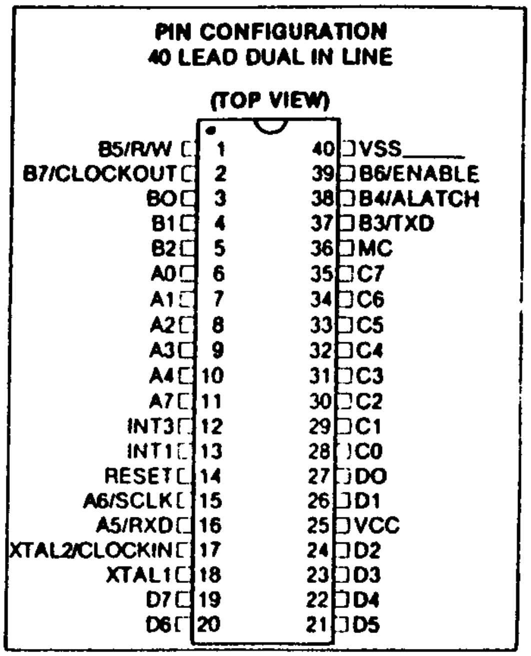
General Description

The CTS256A-AL2 is an 8-bit microprocessor programmed with letter-to-sound rules. This built-in algorithm allows for the automatic real-time translation of English ASCII characters into alloenaddresses compatible with the SP0256A-AL2 (276-1784) Narrator® Speech Processor.

The CTS256A-AL2 is functionally equivalent and pin-for-pin compatible with the PIC7041. As such, it retains its ability as an 8-bit microcomputer for execution of user-defined programs resident externally. Its uniqueness is the internal code-to-speech algorithm which converts English text (in the form of standard ASCII characters) into alloenaddresses, using letter-to-sound rules.

The chip can be used with virtually any microcomputer, since it utilizes either standard serial (TTL Level) or parallel input, with full hardware (DTR-DSR) and software (XON-XOFF) handshaking.

Upon power-up or use of the hardware reset, the system initializes itself and then speaks 'O.K.' to demonstrate that it is ready for input.



ELECTRICAL CHARACTERISTICS

Maximum Ratings*

All pins with respect to Vss	-0.3 to 8.0V
Storage Temperature	-25°C to 125°C

Standard Conditions

Clock—Crystal Frequency	10.00 MHz
Operating Temperature (T)	0°C to 70°C

DC CHARACTERISTICS/CTS256A-AL2

Exceeding these ratings could cause permanent damage to the device. This is a stress rating only and functional operation of this device at these conditions is not implied. Operating ranges are specified in Standard Conditions. Exposure to absolute maximum rating conditions for extended periods may affect device reliability.

Data labeled "typical" is presented for design guidance only and is not guaranteed.

CUSTOM PACKAGED IN U.S.A BY RADIO SHACK, A DIVISION OF TANDY CORPORATION

Schematic 1

Schematic 1 shows a minimum component configuration. It uses serial input whose serial parameters are 7 bits per character, 2 stop bits, no parity, and uses only the internal RAM input buffer. This buffer accommodates words that are no greater than 19 characters in length followed by a delimeter; with an output buffer that accomodates an allophone translation of not more than 26 bytes. Since the translation more often than not results in the output buffer contents consisting of two times that of the input buffer, words no longer than 13 characters in length should be used as a rule of thumb. If a translation results in an over-flowed output buffer, the system reset may have to be used to clear the system. Serial baud rate is selected per Table 1.

Schematic 2

Schematic 2 shows the configuration necessary to incorporate a parallel port, a 1792 byte input buffer (External RAM), and switch selectable serial parameters. The portions of the schematic surrounded by a dotted line may be deleted according to which options you may wish to incorporate.

The parallel port accepts standard English ASCII characters. The parallel data strobe signifies that the data is valid, and latches the data.

The External RAM mode can be used to increase the size of the input buffer to 1792 bytes. In this mode, two pages of text can be loaded into the system. The remainder of the 2K X 8 RAM is 256 bytes which is used to extend the output buffer. (The output buffer is the area where the strings of allophone addresses are held prior to transfer to the SP0256A-AL2.)

The serial parameters option may be selected if the host terminal or computer cannot be set to accomodate the default parameters. Note also that if you are using RS-232 level signals, and the transmission rate is to be under 4800 baud, the outlined components may be deleted.

Schematic 2 Option Switches

Internal/External RAM Mode

(Expansion Buffer Option) In either Internal or External RAM Mode, the input buffer is protected from overflow by handshaking which signals the host when the input buffer is full, and again when it is ready for more input. BUSY (pin 3) toggles LO when the input buffer becomes 86.5% full. The serial output (pin 37) will also send an XOFF character. In this way, the host may use its discretion to complete that transmission or a part thereof. If the input buffer becomes 100% full, the parallel and serial port interrupts are disabled to prevent input buffer overwrite. The interrupts are not reenabled until the BUSY condition has dissipated. Busy will toggle HI and Serial Output will send XON when the input buffer becomes 50% empty.

Any Delimiter/Carriage-Return Only Option

In the any-delimiter mode, the code-to-speech algorithm will process and speak words or phrases as soon as they are followed by any delimeter. These include . , ; : ? space, carriage return, etc. In the carriage-return only mode, the algorithm will process and speak only after a carriage-return is received. The latter mode is meant for use with a slow input device such as a terminal, where the user wishes to buffer-up a complete phrase so that it is spoken with fluency.

If the carriage-return only option is chosen (with external RAM), limit to 160 characters the length of the phrase which is entered before the carriage return is entered. This will allow a two line phrase to be spoken with fluency while ensuring that the 256 byte output buffer does not overflow.

Control Characters

'ESC' (1B Hex) - The Escape key code will cause the CTS256A-AL2 to dump the current input buffer. It may also be used to silence speech output which is in progress.

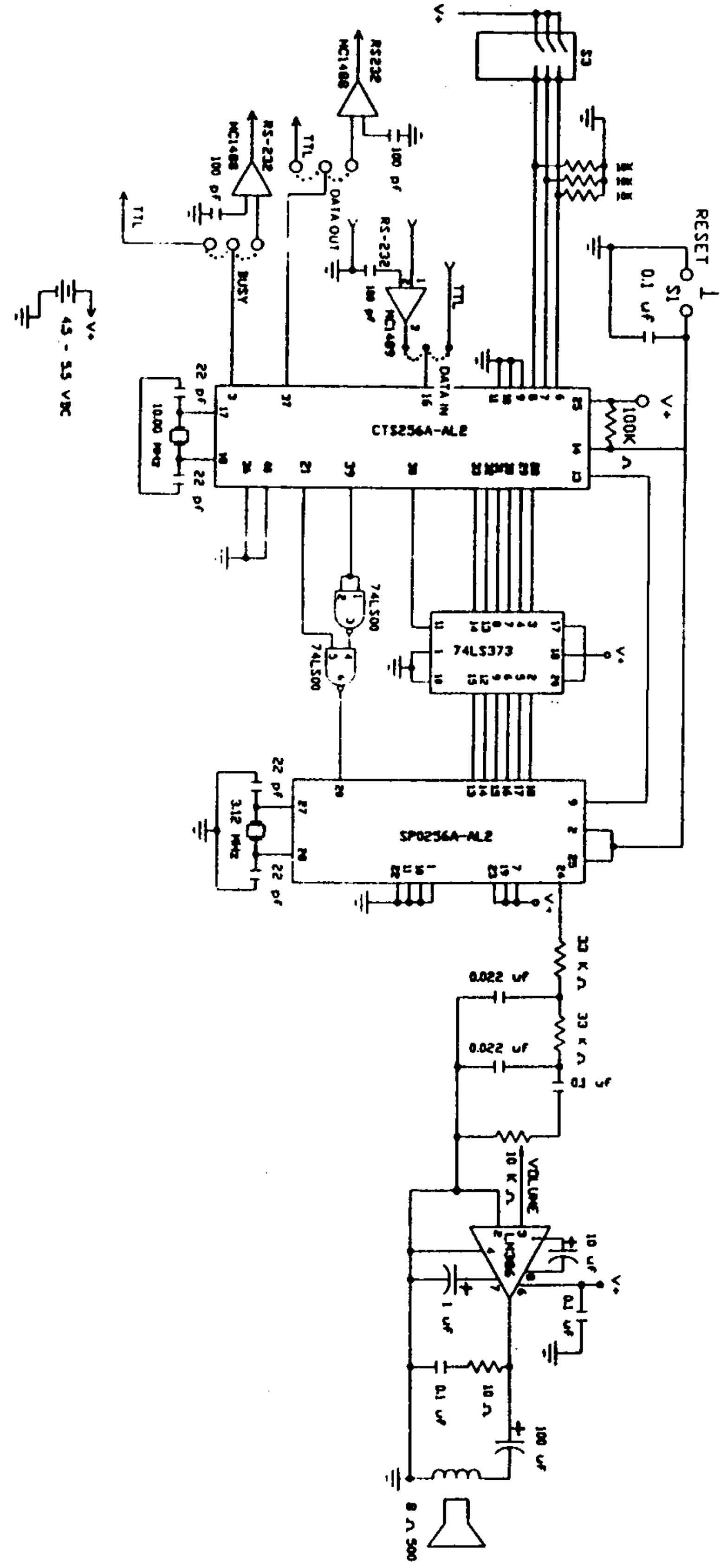
BACKSPACE (08 Hex) - The Backspace key code erases the input buffer one character at a time, beginning with the latest entry.

TABLE 1
SPECIAL PIN FUNCTIONS

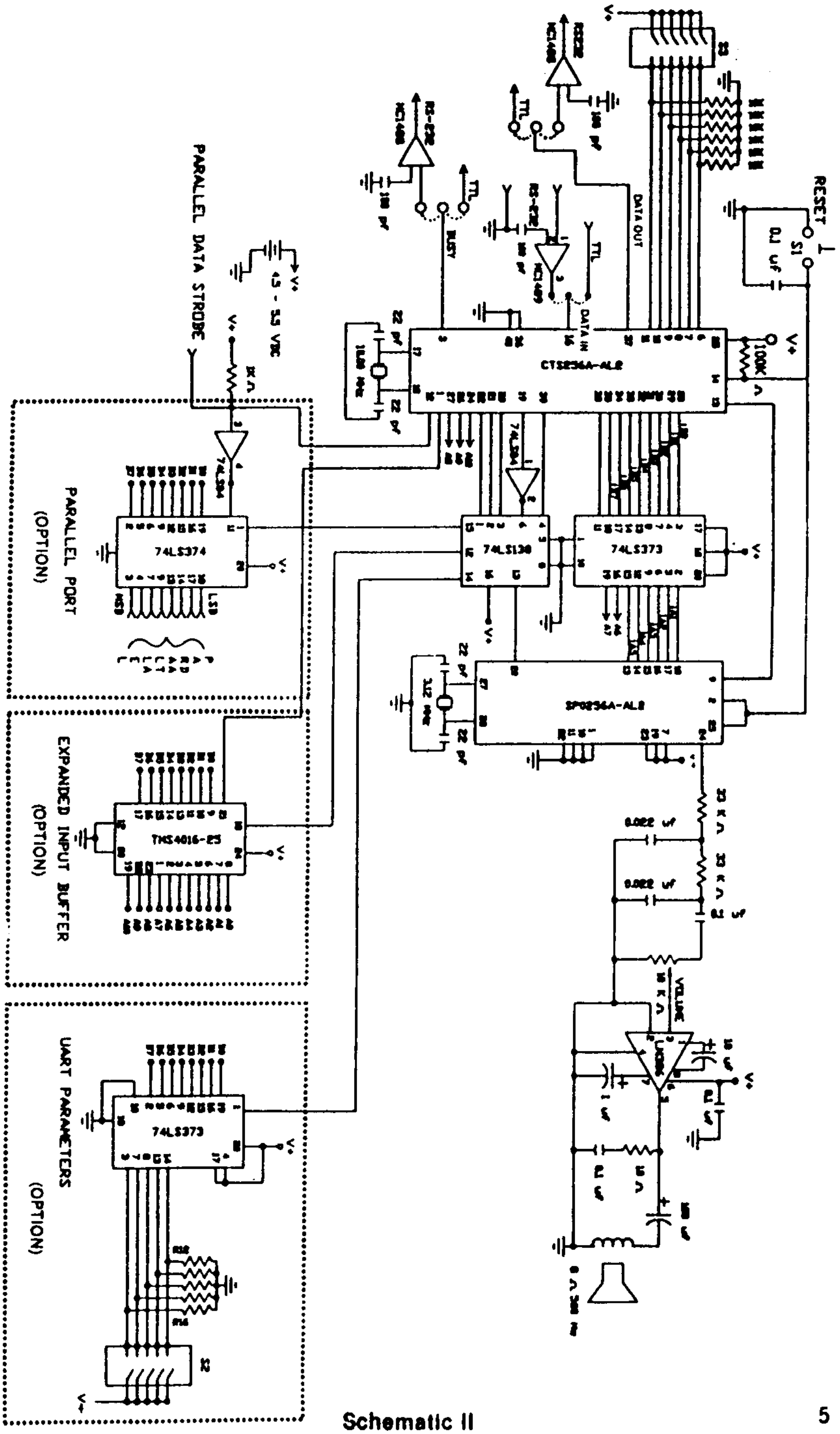
INPUT MODE SELECTION				
PIN	6	7	8	MODE
	0	0	0	PARALLEL INPUT MODE
	0	0	1	SERIAL MODE - 50 BAUD
	0	1	0	SERIAL MODE - 110 BAUD
	0	1	0	SERIAL MODE - 300 BAUD
	1	0	0	SERIAL MODE - 1200 BAUD
	1	0	1	SERIAL MODE - 2400 BAUD
	1	1	0	SERIAL MODE - 4800 BAUD
	1	1	1	SERIAL MODE - 9600 BAUD
PIN	9			
	0			DEFAULT SERIAL PARAMETERS (7 BITS, 2 STOP BITS, NO PARITY)
	1			SELECTABLE SERIAL PARAMETERS (SEE TABLE 2)
PIN	10			
	0			INTERNAL RAM BUFFERS (20 BYTE INPUT/26 BYTE OUTPUT)
	1			EXTERNAL RAM BUFFERS (1972 BYTE INPUT/256 BYTE OUTPUT)
PIN	11			
	0			CARRIAGE-RETURN ONLY DELIMETER
	1			ANY DELIMETER
'N	3			(OUTPUT ONLY)
	0			INPUT BUFFER IS "BUSY"
	1			INPUT BUFFER IS "NOT BUSY"

TABLE 2
U10 (74LS373) UART CONFIGURATION

PIN	3	4	
	0	1	6 BITS/CHARACTER
	1	0	7 BITS/CHARACTER
	1	1	8 BITS/CHARACTER
PIN	8		
	1		PARITY ENABLED
	0		PARITY ENABLED
PIN	7		
	1		EVEN PARITY
	0		ODD PARITY
PIN	3		
	0		ONE STOP BIT
	1		TWO STOP BITS


V+
S3
10 K
RESET
S1
0.1 uF
100K
V+
CTS256A-AL2
DATA IN
TTL
RS-232
MC1489
100 pf
DATA OUT
TTL
RS232
MC1488
100 pf
RS-232
MC1488
100 pf
BUSY
TTL
22 pf
10.00 MHz
22 pf
74LS00
74LS00
74LS373
V+
SP0256A-AL2
22 pf
3.12 MHz
22 pf
V+
33 K Ω
0.022 uF
33 K Ω
0.022 uF
0.1 uF
VOLUME
10 K Ω
LM386
10 uF
V+
0.1 uF
1 uF
10 Ω
0.1 uF
100 uF
8 Ω
V+
4.5 - 5.5 VDC

Schematic I



Schematic II

FOR THE TECHNICALLY CURIOUS

External EPROM(s)

External EPROM(s) can be added to the CTS25A-AL2 to improve the pronunciation of proper names, acronyms and technical words, and can also be added to store user-programs. These are called Exception and User EPROMs,respectively. The coding and use of these EPROMs requires special equipment and experience in assembly-level programming. If you need more information on using external EPROMs, send a letter including your name and address, referencing the CTS256A-AL2 Code-to-Speech Processor to:

Radio Shack Technical Publications
900 Two Tandy Center
Fort Worth, Texas 76012

Attention: Group 4749

Code-To-Speech Algorithm

Upon power-up (or hardware reset) the CTS256A-AL2 determines the system configuration with respect to external ROM/RAM/Switch selectable UART parameters.

ROM

A search is made in 1K increments (from 200H to E000H) for the 5-byte code which identifies an external EPROM. If only an Exception EPROM is found, or if both an Exception EPROM and a User EPROM are found, the mode control information is to be resident in the Exception EPROM. If only a User EPROM is found, the mode control information is to be resident within it. An Exception and/or User EPROM can reside anywhere within the decodeable address space of the CTS256A-AL2, providing its start address falls on a 1K boundary.

RAM

If the Expansion Buffer option is selected and no external EPROMs are present, the start address of the additional RAM is 3000H. This buffer may be expanded in 256 byte contiguous block increments. The algorithm will find the end of the Expansion Buffer by searching for the first non-RAM location at 256 byte intervals. If the Expansion Buffer option is selected and either one of the Exception/User EPROMs are present, then the user-defined start and end addresses of the additional RAM shall be resident in four reserved locations within the EPROM.

Switch Selectable UART Parameters

Total flexibility of the UART parameters is possible utilizing the Full-Parameter UART mode. If no external EPROMs are present, the algorithm will expect these parameters to be switch selected at 1000H. If external EPROM(s) is present and this mode is selected, then the address of the switch selectable parameters can be user-defined.

After completion of the initialization, one of two paths is taken dependent upon the system configuration. (1) In a dedicated code-to-speech system (which implies no User EPROM present), the algorithm idles so long as the input buffer remains empty. (2) If User EPROM is present, the algorithm relinquishes control to the User program. This is the time-share code-to-speech mode, in which a User program can run time-shared with code-to-speech processing.

In the dedicated code-to-speech mode, standard ASCII characters are processed as they are input. Processing starts with an alphabetical search of the Exception EPROM(s), if present. If it is present and no exact match for the character string is found, or if it is not present, the algorithm employs a letter-to-sound rule table against which main, right, and left context matches are performed. This results in the translation of a particular word into the proper string of allophone addresses necessary for its pronunciation. This list of allophone addresses is sent to the SP0256A-AL2 after a carriage return, of after any delimiter; depending on the selection of Input Delimeter switch.

In the time-share code-to-speech mode, the User program can control the stepping of the code-to-speech algorithm through its main routines as it processes a word or phrase. This is accomplished through calls to the main code-to-speech routines. After each routine is completed, control returns to the User program. In this way the User can incorporate speech in real time to their program running in the CTS256A-AL2. For this mode to function properly, the User must maintain a housekeeping routine to save the status of the processor between User/code-to-speech time-sharing.

The actual conversion of text to speech is accomplished within the CTS256A-AL2 using a conversion algorithm based on work done by the Naval Research Laboratories, and expanded to cover the features of the SP0256A-AL2 Speech Processor. As characters are passed to the routine from the input buffer, the algorithm applies left and right context rules to the character until a match is found. The resulting conversion produces a series of alloenadddresses which are buffered up to the end of a word, then passed on to the SP0256A-AL2 Narrator® Speech Processor. The set of rules used produces speech which is approximately 85% accurate in terms of correct pronunciation. Text which contains many technical words or proper names may fall below this accuracy rating, while text which consists of simple text may have a substantially higher level of accuracy.

For example, the following is a portion of the rules which are used to determine the pronunciation of the letter "e" when it occurs in a word:

MATCHING STRING	ALLOPHONES	EXAMPLE
<[E]<	= IY	"E"
#:[E]<	= ""	FADE
' *[E]<	= ""	I'VE, WE'RE, YOU'RE
*[E]<	= IY	SHE, ME, WE
#[ED]<	= PA2, DD1	TRIED, GREED, SPEED
#:[E]D <	= ""	DRAGGED
[EV]ER	= EH,VV	EVERYTHING
#*[EL]	= EL	GRAVEL, TRAVELER
[ERI]#	= YR, IY	EXPERIENCE
#:[ER]#	= ER1	GENERAL, OBLITERATE
[E] %	= IY	DELETING
[ERI]	= EH, EH, RR1, IH	PERILOUS, PERIMETER
[ER]#	= EH, XR	
[ER]	= ER1	OBSERVE, UNDERSTAND
<[EVEN]<	= IY, VV, IH, NN1	EVEN
<[EVEN]	= IY, VV, EH, EH, NN1	EVENT, EVENTUAL
#:[EW]	= YY1, UW2	SINEW
@[EW]	= UW2	DEW CHEW, NEW
[EW]	= YY1, UW2	FEW, SPEW

= One or more Vowels (A, E, I, O, U, or Y)
. = One or B. D, G, J, L, M, N, R, V, W, X
% = One of ER, E, ES, ED, ING, ELY (A suffix)
& = One of S, C, G, Z, X, J, CH, SH (A sibilant)
@ = One of T, S, R, D, L, Z, N, J, TH, CH, SH preceeding a "u" sound. (Ex. rule, mule)
= One consonant (B, C, D, F, G, H, J, K, L, M, N, P, Q, R, S, T, V, W, X, Z)
+ = One of E, I, Y (A front vowel)
: = Zero or more consonants
* = One or more consonants
> = One of O or U
< = Anything other than a letter
? = Zero or more vowels
Brackets [ ] surround the pattern to be substituted.

Each rule is applied by first matching the characters enclosed in brackets, then working outward from these characters, first to the left and then to the right, to determine if the character pattern matches the rule being tested.

Looking at a specific rule:
#:[E]<   = ""
would be read as:
The letter "E" which is preceeded by zero or more consonants preceeded by one or more vowels and followed by any non-alphabetic character (space, punctuation, etc...) is to be replaced by a null string (the "E" will be silent). In the example word "FADE", the "E" is preceeded by one consonant ("D"), which is preceeded by one vowel ("A"), and it is followed by a non-alphabetic character (end of the word), therefore the "E" is silent.

One other example:
#:[ER]# = ER1
The letters "ER" which are preceeded by zero or more consonants preceeded by one or more vowels and which are followed by one or more vowels is replaced by the allophone "ER1". In the example word "GENERAL", the "ER" is preceeded by one consonant ("N") which is preceeded by one vowel ("E") and is followed by one vowel "A", therefore the "ER" is pronounced as the allophone "ER1".

Note that the sequence of the rules is not arbitrary. It is important to realize that a very general rule may be a subset of a more specific rule, and therefore care was taken in its placement so that a more specific rule preceeded a more general rule, since the rules are searched sequentially.