DESIGNED & MANUFACTURED
BY FULLER MICRO LIVERPOOL
71 DALE STREET, LIVERPOOL.
Telephone 051-236 6109

DESCRIPTION

The Fuller Orator utilizes the general instrument SP025A speech processor. This is a single chip N-channel MOS/LS1 circuit that is able through a stored program, to synthesize speech. The ORATOR is a complete speech system, containing an allophone library which can create any phrase. The unit simply plugs on to the back of the Sinclair ZX spectrum, it is completely self cointained with its own amplifier andloudspeaker. The ZX power supply plugs into the Orator powering it and the Spectrum.
In addition to synthesizing speech the ORATOR performs two other functions:
a) Cassette interface
b) Spectrum beep amplifier
The cassette interface allows you to leave both loading and saving leads plugged into the Spectrum whilst loading or saving. The unit also enhances the saving signal, so that cheaper cassette recorders can be used. The beep amplifier performs just that function. The volume being controlled by a small pre-set variable resistor on the pcb.
Using the Orator thro buss, the ZX printer or microdrive can be plugged onto the Orator without fear or interference. The Fuller Orator is a port mapped device and is enabled by the command OUT 159,a. Where 'a' is the numeric value assigned to an allophone (see table) in order to learn all the sounds within the Orator program the following sample program will step through the allophone list with a slight pause in between each.

```
10 FOR a = 1 to 255
20 OUT 159,a
30 pause 5
40 next a
50 GOTO 10
```

LANGUAGE

In order to successfully use a set of allophone sounds to synthesize words, there are a few preliminary points which should be made about speech and language. First, there is no one-to-one correspondence between written letters and speech sounds; second, speech sounds are no discrete units such as beads on a string; and last, speech sounds are acoustically different, dependent upon position within a word.
The first point compares to what a child encounters when learning to read. Eah sound in a language may be represented by more than one letter and, conversely, each letter may represent more than one sound. (See the examples in Table 1). Because of these spelling

irregularities, it is necessary to think in terms of sounds, not letters, when dealing with speech allophones.
The second point concerns segmentation of the speech signal. An adult who has learned to read usually thinks of the acoustic stream of speech as a string of discrete sounds which he calls by their letter names. In fact, speech is a contuously varying signal which cannot easily be broken into distinct sound-size units. For example, if an attempt is made by extract the **b** sound from the word **bat** by taking successively larger chunks of the acoustic signal from the beginning of the word, either a non-speech noise or the syllabe **ba** is heard. In other words, there is no point at which the **b** sound can be heard in isolation.
The final point, and the most important for users of an allophone set, is that the acoustic signal of a speech sound may differ depending upon word positions. For example, the initial **p** in **pop** will be acoustically different from the **p** in **psy,** and may be different from the final **p** in **pop**. The ear will perceive the same acoustic signal differently depending upon which sounds precede or follow it. The word **cot** can be made to sound like **cod** by lengthening the duration of the **o** and, conversely, the word **cod** can be made to sound like **cot** by shortening the duration of the **o.**

PHONEMES OF ENGLISH
The sounds of a language are called phonemes, and each language has a set which is slightly different from that of other languages. The table contains a chart of all the consonant phonemes of English. It also shows all the vowel phonemens of English.
Consonants are produced by creating a constriction or occlusion in the vocal tract which produces an aperiodic sound source. If the vocal chords are vibrating all at the same time, as in the case of the voiced fricatives VV, DH, ZZ, and ZH, (see Table) there are two sound sources: one which is aperiodic and one which is periodic.
Vowels are produced with a relatively open vocal tract and a periodic sound source (unless they are whispered) provided by the vibrating vocal chords. Vowels are classified according to whether the front or back of the tongue is high or low (see table), whether they are long or short, and whether the lips are rounded or unrounded. In English, all rounded vowels are produced in or near the back of the mouth (UW, UH, OW, AO, OR, AW).
Speech sounds which have features in common behave in similar ways. For example, the voiceless stop consonants PP, TT, and KK (see Table) should be preceded by 50-80 msec of silence, and the voiced stop consonants BB, DD, and GG by 10-30 msec of silence.

ALLOPHONES
Phoneme is the name given to a group of similar sounds in a language. Recall that a phoneme is acoustically different depending upon word position. Each of these positional variants is an allophone of the same phoneme. An allophone, therefore is the manifestation of a phoneme in the speech signal. It is for this reason that the inventory of English speech sounds is called an allophone set.

HOW TO USE THE ALLOPHONE SET
(See Tabe for instructions on how to create all the sample words mentioned in this section).
The allophone set (see Table) contains two or three versions of some phonemes. It may be necessary to use one allophone of a particular phoneme for word-or syllable-initial position and another for word-or syllable-final position. A detailed set of guidelines for using the allophones is given in the Table. Note that these are suggestions, not rules.
For example, DD2 sounds good in initial position and DD1 sounds good in final position, as in "daughter" and "collide." One of the differences between the initial and final versions of a consonant may be that an initial version is longer. Therefore, to create an initial SS, two SSs

can be used instead of the usual single SS at the end of a word or syllable, as in "sister." Note that this can be done with TH, and FF, and the inherently short vowels (to be discussed below), but no other consonants. Experiment with some consonant clusters (strings of consonants such as **str, cl**) to discover which version works best in the cluster. For example KK1 sounds good before LL as in "clown," and KK2 sounds good before WW as in "square." One allophone of a particular phoneme may sound better before or after back vowels and another before or after front vowels. KK3 sounds good before UH and KK1 sounds good before IY, as in "cookie." Soume sounds (PP, BB, TT, DD, KK, GG, CH, and JH) require a brief duration of silence before them. For most of these, the silence is included in the allophone, but more may be added as desired. Therefore, there are several pauses in the allophone set varying from 10-200 msec. To create the final sounds in the words "letter" and "little" use the allophones ER and EL. Think about how a word **sounds,** not how it is spelled. For example, the NG allophone obviously belongs at the end of the words "sing" and "long," but notice that the NG sound is represented by the letter N in "uncle." Some sounds may not be represented in words by any letters, such as the YY sound in "computer."
As mentioned earlier, there are some vowels which can be doubled to make longer versions for stressed syllables. These are the inherently short vowels IH, EH, AE, AX, AA, and UH. For example, in the word "extent" on EH is used in the first syllable, which is unstressed, and two EHs are used in the second syllable, which is stressed. Of the inherently long vowels there is one, UW, which has a long and short version. The short one, UW1, sounds good after YY in computer. The long version, UW2, sounds good in monosyllabic words like "two." Included in the vowel set is a group called R-couloured vowels. These are vowel + R combinations. For example, the AR in "alarm" and the OR in "score." Of the R-coloured vowels there is one, ER, which has a long and short version. The short version is good for polysyllabic words with final ER sounds like "letter," and the long version is good for monosyllabic words like "fir." One final suggestion when creating sentences is to add a pause of 30-50 msecs between words and a pause of 100-200 msec between clauses.

TABLE

LENGTH OF PAUSE

		PA1 (10ms)	before BB, DD, GG, and JH
		PA2 (30ms)	before BB, DD, GG, and JH
		PA3 (50ms)	before PP, TT, KK, and Ch, and between words
		PA4 (100ms)	between clauses and sentences
		PA5 (200ms)	between clauses and sentences
Letter	Allophones	Data	Examples
A	*AE	26	extract, acting
	AO	23	talking, song
	AX	15	label, instruct
	R-	Coloured	
	AR	59	farm, alarm, garment
	XR	47	hair, declare, stare
B	BB1	28	final position - rib, between vowels - fibber in clusters - brown
	BB2	63	initial position before a vowel - beast
C	CH	50	church, feature
D	DD1	21	final position - played, end
	DD2	33	initial position - down, clusters, drain
E	EH	7	extent, gentlemen
	IY	19	treat, people, penny
	EY	20	great, statement, tray
	R-	Coloured	vowels
	ER1	51	letter, furniture, interrupt
	ER2	52	monosyllables - bird, fern, burn
	FF	40	may be doubled for initial position and used singly for final position

G	GG1	36	before high vowels YR IY IH EY EH XR
	GG2	61	before high back vowels UW UH OW OY AX and clusters green and glue
	GG	34	before low vowels AE AW AY AR AA AO OR ER in medicine clusters - anger and final position peg
H	HH1	27	before front vowels YA IY IH EY EH XR AE
	HH2	57	before back vowels UW UH OW OY AO OR AR
I	IH	12	sitting, stranded
	AY	6	kite, mighty
J	JH	10	judge, injure
K	KK1	42	before front vowels YR IY IH EY EX XR AY AE ER AX
	KK2	41	final position - speak - final clusters, task
	KK3	8	before back vowels UW UH OW OY OR AR AO
L	LL	45	like hello - steel
	EL	62	little, angle, gentlemen
M	MM	16	milk, alarm, ample
N	NN1	11	before front and central vowels YR IY IH EY EH XR AE ER AX AW AY UW final clusters - earn
	NN2	56	before back vowels UH OW OY OR AR
	NG	44	string, anger
O	OY	5	noise, toy, voice
	OW	53	zone, close, snow
	AW	32	sound, mouse, down
	AO	23	song
	AA	24	pottery, cotton
	UH	30	cookie
	R-	Coloured	
	UW2	31	in monosyllabic words - too, food, for, tune, adorne, store
	OR	58	
P	PP	9	pleasure, ample, top
Q	KK3	8	quick
R	RR1	14	initial position - read, write, x-ray
	RR2	39	initial clusters - brown, crane, grease
S	*SS	55	this may be doubled for initial position and singly in final position
	SH	37	shirt, leash, nation
T	TT1	17	final clusters - before ss, test, its
	TT2	13	all other positions - test, street
	TH	29	may be doubled for initial position and used singly in final position
	DH1	18	word - int, pos, this, the, they
	DH2	54	word final and between vowels - bathe, bathing
U	UW1	22	after clusters - wity, yy - computer
	UH	30	full
	AX	15	instruct
	YY1	49	clusters, cute, computer
	ER2	52	burn
V	VV	35	vest, prove, even
W	WW	46	we, warrant, linquist
	WH	48	white, whim, twenty
Y	YY1	49	clusters - cute, beauty, computer
	YY2	25	initial position - yes, yam, yo-yo
Z	ZZ	43	zoo, phase
	ZH	38	beige, pleasure

*These allophones can be doubled.