THE
SPEECH ROM
FOR THE
BBC MICRO!

SPEECH ROM

Instruction Manual

Computer Concepts

CONTENTS

This manual was typeset direct from Speech Rom files by Quorum Technical Services Ltd., Cheltenham.

1. INTRODUCTION

1.1 – Overview of the System

The BBC Microcomputer is provided with a socket for the Texas Instruments TMS 5220 Speech Synthesis Processor. This device normally receives its speech data from a specially designed serial ROM (part of the Acorn speech upgrade): an arrangement which gives very good reproduction but an extremely limited vocabulary. A further, less commonly appreciated problem is that, although the individual words or short phrases contained in such ROMs do have intonation, this is fixed, and may be quite inappropriate when attempts are made to string the words together to make longer utterances.

The SPEECH ROM is designed to drive the speech processor directly. Rather than having a fixed vocabulary of words, the user may construct any English word with very little sacrifice in quality. The program employs a unique method of stress and rhythm control giving simple access to the entire range of English speech and intonation.

On a lighter note, the SPEECH ROM is able to sing, albeit without polish owing to the limitations of the speech processor which was never intended for such an application!

As with most Sideways ROMs, commands are issued to the SPEECH ROM by typing an asterisk, *, followed by the command word and its parameter(s), if required. Details of each command are given in Section 2 of the manual, and a summary of the syntax for all the available commands is also given on the help menu (see below for details). All commands may be abbreviated: the minimum form depending upon the presence of other ROMs with similarly spelt commands. Usually the first two or three characters, followed by a full stop, will suffice.

Extensive error messages are issued when commands are incorrectly entered. The error messages are explained in detail in Section 3.

The SPEECH ROM operates under interrupts. To the user this simply means that speech will carry on while a program is running. Up to about half a minute of speech can be entered to carry on speaking independently.

There is a wiring error in earlier versions of the BBC Microcomputer; those with Series 1, 2 and 3 printed circuit boards. The SPEECH ROM detects if this error is present and reconfigures itself accordingly; no hardware alterations are therefore needed. Because the 5220 Speech processor chip fits into the standard machine, any faults in this area are outside any warranty offered on the SPEECH ROM itself.

This introductory section concludes with a short explanation of the use of the ＊UTTER command so that the user can obtain an overall feel of the system before embarking on the detailed explanations contained in the ensuing sections.

1.2 Making a start

The following assumes that the SPEECH ROM and speech processor have been installed, as detailed on the enclosed sheet.

The SPEECH ROM requires two pages of memory (512 bytes) in order to operate. Because it would be inconvenient to allocate this memory permanantly, the SPEECH ROM must be *enabled* before operation. To enable the SPEECH ROM, type the command:
＊SPON **RETURN**
and press the **BREAK** key. The SPEECH ROM is now ready to accept commands. Entering commands before the above will result in a message stating that the SPEECH ROM is off. The Operating System does not allow pages of memory to be allocated without pressing the **BREAK** key. It is therefore, unfortunately, unavoidable.

Now that the SPEECH ROM is ready you may try the following short example:
＊UTTER <1> B E B E S E M ＊I K R O K UH M P Y +OO T UH
RETURN

2

Everything in the example must be typed in exactly as it is printed, including the spaces and the **RETURN**. Although the command spreads over two lines on this example, it should be entered as just one line on the computer. If the command results not in speech but in an error, check the input carefully and try again. If the command is accepted, but no sound is produced, check through the fitting instructions and ensure that the 5220 speech processor chip is correctly installed. If it is still not working then you may have to take your computer to the local qualified dealer.

Let us take each part of the input and see what it does. First ＊UTTER introduces a command instructing the SPEECH ROM to speak.

The next part of the command, <1>, tells the SPEECH ROM which intonation pattern to use; in this case it is a simple statement. This gives the 'overal sound' of the speech following. Try replacing <1> with <2> which is the simple question form, or by <1+> which is an emphatic statement, and listen to the differences. Use the **COPY** key to copy the rest of the line. The other simple intonation patterns that you can initially experiment with are:
<3>, <4>, <5> and <1->; their uses are difficult to explain without going into detail but their sounds will no doubt be familiar to you.

Finally comes a sequence of characters which specify the actual sounds to be spoken. Try the following sequence instead of the first:
＊UTTER <1> P +E T UH P +I P UH P +i K T UH P +e K UH V P +i K L D P ＊e P UH**RETURN**

This example sounds much better, by the way, with the emphatic intonation, <1+>.

Compare what you hear with what you have typed. As you can see each of the sounds of the language is represented by one or two letters. Two letter codes are necessary because there are 54 individual sounds in the English dialect spoken by the SPEECH

3

ROM. In the above examples the name *Peter* requires four separate sounds to make the complete word:

P, E, T and UH

the word *computer* requires eight:

K, UH, M, P, Y, OO, T and UH

A space is needed to separate each of these sounds so that the SPEECH ROM can distinguish between them.

Here is a full list of all the speech sounds used by the SPEECH ROM, each followed by a single word containing an example of the sound. (Bear in mind that many of the sounds have several different possible spellings of which the example given is only one; Section-4 contains a comprehensive list of spellings.)

Vowels

E - see	i - tin	e - pet
a - cat	u - rub	AR – class
o - dog	OR – fall	oo – put
OO – food	ER – hurt	A - play
I - line	O - bone	OY – boy
AW – clown	EA – steer	ea – pear
UR – tour	UH – utter	

Consonants

P - pot	B - bad	T - tea
D - dot	K - can	G - get
CH – chest	J - jam	F - fun
V - van	TH – thin	DH – then
S - sad	Z - zone	SH – shed
ZH – vision	H - ham	M - mad
N - nod	NG – sing	L - lip
R - run	Y - yet	W - wet

The special characters + and * that you can see in the examples tell the SPEECH ROM how to relate the intonation pattern to the speech sounds, we will refer to them later as *stress markers*. You may only use them in front of vowels (except UH).

The stress marker + helps to give the right rhythm to the speech. The simplest way to understand its use is to think of

each phrase as a rhyme and then put a + in front of any vowel on the beat. This is shown well in the Peter Piper example, try leaving out each + and listen to the difference – it should sound less rhythmical and flatter.

The stress marker * shows where the main emphasis comes in the phrase; which word is the most important. Repeat each phrase to yourself and you will notice how you put more effort into one of the syllables, usually one near the end. You should place the * in front of the vowel of this syllable. Listen to the difference between the two following examples (and the original) when the position of * is changed; the first alteration emphasises *BBC*, the second *computer*. If you leave * out altogether the speech becomes very flat and Dalek-like showing how important the SPEECH ROM's intonation really is!

*UTTER <1> B E B E S *E M I K R O K UH M P Y +OO T UH **RETURN**
*UTTER <1> B E B E S E M +I K R O K UH M P Y *OO T UH **RETURN**

There are five stages to composing a *UTTER command:

(a) Work out the correct speech sounds for the phrase you want to say.
(b) Decide which intonation pattern you think is appropriate.
(c) Tap out the rhythm and put a + in front of each vowel that seems to have a beat.
(d) Choose which word in the phrase needs to be emphasised and place a * in front of its most prominent vowel.
(e) Listen carefully to the result.

At this point you will probably want to refine the command by changing some of the speech sounds and moving the stress markers around. All the time keep repeating the phrase to yourself and aim to approximate the output to your own speech. Bear in mind that there is no single correct solution – most things can be said in a number of different ways.

Editing *UTTER commands using the BASIC editor, with constant use of the **COPY** key can be long-winded. For users

possessing **Wordwise-Plus** Appendix I describes a much faster and more elegant method of developing SPEECH ROM phrases.

Finally, do not attempt to take in everything at one go; it is far better to build slowly on this introduction by trying out lots of examples of your own. The SPEECH ROM is a very friendly system and you should be able to produce quite acceptable speech even if you read no further. You will find, however, that you can make the system produce even more natural and accurate speech with quite subtle intonation if you do study Sections 4 and 5. The difficulties you will encounter are not inherent in the SPEECH ROM but in the English language itself. As far as possible the SPEECH ROM tries to hide these complexities from you, but it cannot go too far without sacrificing the very flexibility it is intended to offer.

```
*UTTER <5-> a N D DH UH B *e S T UH V L +u K
RETURN
```

2. COMMANDS

2.1 The help menu

The Help menu showing the syntax of all the SPEECH ROM commands can be displayed by using the command:

```
*HELP SPEECH RETURN
```

This can usually be abbreviated to:

```
*H.S. RETURN
```

and will produce a display similar to that below:

```
SPEECH n.nn
SPON
UTTER <<tone><phoneme>...>...
SING <<pitch><note><phoneme>...>...
SYNC
VOICE <pitch>
```

There are a number of conventions associated with these commands, some of which have already been mentioned in the introduction. All commands are preceded with * and terminated with RETURN. Unabbreviated commands require a space before the first parameter. All commands may be abbreviated within the compatibility limits of your system; you may experiment to find the shortest versions possible on your machine. Abbreviated commands require a full stop, but no obligatory SPACE before the first parameter. By convention, parameters are enclosed within angled brackets for descriptive purposes. In the SPEECH ROM however, angled brackets are actually used to delimit the parameters in the input string in all cases except <*phoneme*>. Triple full stops ... indicate that the contents of the immediately preceding angled brackets may be optionally repeated. No limit to this repetition is specified, nor

does exceeding the limit cause an error message to be issued, and the system will therefore crash. No error is flagged because the exact circumstances of overflow are difficult to predict, and depend on the actual parameter values involved. However, the limit approaches the maximum BASIC line length and in practice should never be reached. Very long utterances can, in any case, always be split up over a number of BASIC lines (or the equivalent in other calling languages).

2.2 *SPON

This command, which requires no parameters, is used to initialise the SPEECH ROM. When the machine is switched on the presence of the SPEECH ROM is acknowledged by the Operating System and *HELP (with no parameter) will list it along with the other resident ROMs. At this stage, however, it is almost completely inactive and any attempt to enter its other commands will result in the error message "SPEECH is off!" being issued. In this state the SPEECH ROM in no way interferes with the operation of the machine, it is not servicing any interrupts, and no extra memory locations are being used.

*SPON alone will not activate the SPEECH ROM, but it prepares the system for activation at the next BREAK (Soft Break). The complete sequence is, therefore:-

*SPON RETURN
BREAK

At this point several important things occur which have implications for the user:

(a) The machine's memory map is re-organised, allowing the SPEECH ROM two *pages* (512 bytes) of memory for its own use (as buffers). This itself has several consequences: PAGE will be reset destroying any resident BASIC program or WORDWISE text; the SPEECH ROM must therefore be initialised **prior** to loading any application with which it is to be used. Once the SPEECH ROM is initialised the space available for applications will be smaller, and large programs which ran previously may no longer do so. This should pose no problem as existing programs not using speech can be

8

run with the SPEECH ROM inactive, and those new ones requiring speech will be written within the new memory constraint.

(b) Event Number 4 (start of vertical sync) is activated. Because the BBC Microcomputer has no method of chaining Extended Vectors properly in ROMs, this may involve clashes with other Sideways ROMs also using the Event Vector. Once again, the fact that the SPEECH ROM can be inactivated circumvents the problem. (For details of Events see the Advanced User Guide – Section 12).

(c) the SPEECH ROM claims the use of memory locations &9C0-&9FF. Some of this space is used for variables and the rest for preservation of data from Zero Page. The SPEECH ROM uses Zero Page memory locations; but because it replaces the original data, it does so completely transparently. The region of memory used by the SPEECH ROM for variables and temporary storage is normally reserved for the Acorn speech expansion. If you have this on board, attempts to use it while the SPEECH ROM is active will probably cause the SPEECH ROM to crash. The same region of memory is used as the cassette output buffer; you should not, therefore, try to save to tape while the SPEECH ROM is speaking, and for the same reason the SPEECH ROM should be reinitialised with BREAK after any cassette save operation. Some commercial games also steal this area for their own use and may not run while the SPEECH ROM is active.

Once activated the SPEECH ROM will remain so in spite of any subsequent BREAK. BREAK will also serve to reinitialise the speech system by clearing its buffers and resetting its variables. If at any time the SPEECH ROM is speaking and you wish to stop it, BREAK may be used (followed, if necessary, by OLD RETURN to retrieve a BASIC program).

Pressing CTRL BREAK (Hard Break), will de-activate the SPEECH ROM. PAGE will be reset, releasing the two pages of memory previously used by SPEECH, and all vectors will be re-initialised. Switching the machine off and then on again has the same effect.

9

2.3 *UTTER

Syntax

*UTTER <<tone><phoneme>...>...

The whole of the SPEECH system is organised around this command, which is used to input the phoneme string(s) and their associated intonation code(s). Each *UTTER command should be followed by one or more *phrases*; a phrase is the unit contained within the outer angled brackets in the above syntax summary. The term has no technical significance, but is used for reasons of simplicity, consistency and brevity. A phrase in the SPEECH ROM is equivalent to what is more properly called a *Tone Group*, more details of which will be given in Section 5. The maximum number of phrases possible with a single *UTTER command is not specified because it varies according to whatever commands have passed before, but generally about 200 characters are allowed in a command.

Each phrase must itself consist of a single <tone>, followed by one or more phonemes. Again the maximum number of phonemes is unspecified. Vowel phonemes are optionally preceded by a stress marker.

2.3.(a) Tones

The SPEECH ROM has 24 distinct tones which specify the intonation pattern to be used in pronouncing the subsequent phoneme string. Full descriptions of all the tones and their uses are given in Section V. There follows a list of valid tone codes for reference purposes.

<1>	<u1>	<1,3>	<u1,3>
<1+>	<u1+>	<1+,3>	<u1+,3>
<1->	<u1->	<1-,3>	<u1-,3>
<2>	<-2>		
<2v>	<-2v>		
<3>	<-3>		
<4>	<4->		
<5>	<5->	<5,3>	<5-,3>

10

Entry of any other value or the absence of either angled bracket will cause a "Bad tone" error. Note that no SPACE is allowed within the brackets. The rather strange looking tone codes result from the system of intonation classification which was followed during the development of the SPEECH ROM. For now the only important distinction is between *simple* tones, which contain a single numeric character (seen in the two left hand columns above), and *compound* tones, which have two numeric characters (in the two right hand columns).

2.3.(b) Phonemes

Phonemes are the fundamental units from which speech is constructed, and they will be discussed fully in Section-4. In the present section a full list is given for easy reference. Included in the list for simplicity are the silence codes, which from a syntactic point of view behave like the phonemes; that is, they may occur in the same places. Each silence code is followed by its approximate duration.

Vowels

E	i	e	a	u	AR	o	OR
oo	OO	ER	A	I	O	OY	AW
EA	ea	UR	UH				

Consonants

P	B	T	D	K	G	CH	J
F	V	TH	DH	S	Z	SH	ZH
H	M	N	NG	L	R	Y	W

Silences

↑ 0.125 sec
0.25 sec
@ 1.0 sec

In many circumstances vowels can be immediately preceded by one of the three stress markers (see below). All phonemes, except the first one of a phrase, must be preceded by a SPACE; if the phoneme contains a stress marker the SPACE must precede the stress marker. Entry of any other value or omission of the SPACE will cause a "Bad phoneme" error. The last phrase of a *UTTER command should be terminated by

11

RETURN, no final **SPACE** is needed. If a further phrase is to be entered in the same *UTTER command, the new tone must be preceded by a **SPACE**.

2.3.(c) Stress Markers

The SPEECH ROM uses stress markers to co-ordinate the intonation pattern specified by a <tone> with the phoneme string which follows. This will be explained in some detail in Section-5, where the functions and uses of stress markers are fully covered. The three symbols used are:
→ (Mode 7 right arrow)
* (Asterisk)
and + (Addition symbol)

Since the rules governing the use of these symbols are a little complicated, there follows a small section on the syntax of each one. Do not be too concerned about this, the SPEECH ROM checks syntax fully on input and issues detailed error messages if a mistake is found.

(i) The Primary Stress Marker, *, is placed directly in front of the required phoneme. It may be used with any vowel except UH. Phrases with a simple tone may contain a maximum of one primary stress marker, phrases with a compound tone, a maximum of two. If a phrase with fewer than the maximum allowed is entered, it will be accepted, but the output may not be equivalent to any standard English intonation pattern.
(ii) The Secondary Stress Marker, +, is placed directly in front of the required phoneme. It may be used with any vowel except UH. There is no maximum to the number of secondary stress markers which can be used in a given phrase.
(iii) The Accessory Stress Marker, −, is placed directly in front of the required phoneme. It may be employed with any vowel including UH but its use is restricted to phrases containing tones <2v> or <−2v>. No phrase may contain more than a single accessory stress marker, and a primary stress marker must have occurred at an earlier point in the phrase. Phrases containing tones <2v> or <−2v> but which omit both primary and accessory stress markers will be accepted, although the output will not be equivalent to any standard English intonation pattern. Similar phrases having a primary but no accssesory stress marker are not allowed.

12

2.4 *SING

Syntax
*SING <<pitch> <note> <phoneme>...>...

This command is used to control the SPEECH ROM's singing voice. Each *SING command should be followed by one or more phrases.

Each *SING phrase must itself consist of a single <pitch>. followed by a single <note> and then one or more phonemes.

2.4.(a) Pitch

The *SING command accepts a different set of values for its pitch parameter from that accepted by *VOICE. The codes are chosen to represent a major scale (no sharps or flats). Notes within each octave are denoted by the musical names of the notes, that is, the letters C to B. Successive octaves are indicated by additional numbers. There follows a list of the legal values for the *SING pitch parameter.

<E>   <F>   <G>   <A>   <B>

<C1> <D1> <E1> <F1> <G1> <A1> <B1>

<C2> <D2> <E2> <F2> <G2> <A2> <B2>

<C3>

Owing to the rather limited frequencies available in the speech processor it is impossible to produce an exact musical scale and at times the SPEECH ROM will appear somewhat sharp or flat! The frequencies are not related to a pitch standard; although some versions of the BBC Microcomputer allow fine-tuning of the speech chip by means of an on-board potentiometer (VR2 on Series 4 machines). If you are adventurous you may care to try tuning the SPEECH ROM to the sound chip for accompaniment purposes!

2.4.(b) Note

This parameter specifies the *total* duration of the phrase in

13

*SING. Each code is an abbreviation of a standard musical term, **s** meaning semiquaver, **q** meaning quaver and so on. Addition of **.** has the usual significance of increasing the duration of the note by half. The next table shows the legal values for **<note>**, with their nominal musical equivalents, and the actual SPEECH ROM durations.

<s>	semiquaver	0.2 sec
<s.>	dotted semiquaver	0.3 sec
<q>	quaver	0.4 sec
<q.>	dotted quaver	0.6 sec
<c>	crotchet	0.8 sec
<c.>	dotted crotchet	1.2 sec
<m>	minim	1.6 sec
<m.>	dotted minim	2.4 sec
<S>	semibreve	3.2 sec

You can calculate that this gives 75 crotchets to the minute. ***SING** ensures the specified duration for a whole phrase by stretching the first vowel or silence encountered in the phoneme string by the required amount. There is no way to *shorten* phrases to achieve a particular duration. Phrases must, therefore, be kept short, restricted usually to one syllable plus any of those *hurried* syllables which are squeezed into a rhythm. You must break up your song in such a way that the correct syllables *hold* the note by coming first in each phrase, the hurried notes must come later. Phrases containing a single silence can be used as musical rests. In such cases a pitch parameter is still required but its value is irrelevant. Adding silences towards the end of a phrase (after the stretched vowel) will separate one note from the next without upsetting the rhythm giving a *staccato* effect. Compare the following versions of the first four notes of the scale:

***SING <C1><q> D O <D1><q> R A <E1><q> M E
<F1><q> F AR** `RETURN`

***SING <C1><q> D O # <D1><q> R A # <E1><q> M E #
<F1><q> F AR** `RETURN`

Careful distribution of consonants between the end of one phrase and the beginning of the next can greatly improve

<div style="text-align:center">14</div>

rhythm and phrasing. Do not be afraid of breaking words in unusual places. The best way to understand these points is to take a simple song and experiment with it.

2.4.(c) Phonemes in *SING
The same phonemes are used in ***SING** as in ***UTTER**. The only difference is that stress markers are not used and will not be accepted.

2.5 *SYNC

This command, which requires no parameters, is used to synchronise the SPEECH ROM with the program issuing the speech commands. It was mentioned in the introduction that speech can be produced while the program carries on with other tasks. This is very useful, but there are also occasions when a program may need to wait until the end of speech before continuing. Whenever ***SYNC** is issued this is precisely what happens. Activate the SPEECH ROM, as shown above, and then try entering and running the following short BASIC program:

```
10 CLS RETURN
20 *UTTER <u1> W +u N @ T +OO @ TH R +E @ F +OR
@ F +I V @ S +i K S @ S +e V UH N @ +A T @ N +I N @
T *e N RETURN
30 PRINT "The first 'beeps' sound
while"'"speech is in progress" RETURN
40 SOUND 1,-10,100,40 RETURN
50 SOUND 1,-10,120,40 RETURN
60 *SYNC RETURN
70 PRINT "The second 'beeps' wait"'"until it
is finished" RETURN
80 SOUND 1,-10,100,40 RETURN
90 SOUND 1,-10,120,40 RETURN
```

Note: there should be no `RETURN` entered in the above program except where shown, but check carefully whether `SPACE` is required.

This idea can usefully be extended, for example in the following program graphics and speech are synchronised. Each image is

<div style="text-align:center">15</div>

constructed invisibly *during* the display and explanation of the previous one. ★SYNC is used to synchronise the change of display. NOTE that the following program uses the OSCLI statement present in BASIC-II; it will therefore need to be modified in order to work on BASIC-I.

```
10 MODE 1
20 RESTORE
30 ACOLOUR=RND(7)
40 VDU19,3,ACOLOUR;0;19,1,ACOLOUR;0;19,2,0;0;
50 LCOLOUR=1
60 NEWSIDES=0
70 REPEAT
80 LCOLOUR=LCOLOUR EOR 3
90 OLDSIDES=NEWSIDES
100 READ NEWSIDES,SIDES$,NAME$
110 PROCdraw(NEWSIDES,1,LCOLOUR)
120 *SYNC
130 ACOLOUR=RND(7)
140 VDU19,LCOLOUR,ACOLOUR;0;19,3-LCOLOUR,0;0;19,3,
      ACOLOUR;0;
150 PROCexplain(NEWSIDES,SIDES$,NAME$)
160 PROCdraw(OLDSIDES,2,LCOLOUR)
170 UNTIL NEWSIDES=100
180 *SYNC
190 MODE7:END
200
210 DEFPROCdraw(sides,logic,lcolour)
220 IF sides=0 ENDPROC
230 GCOL logic,lcolour
240 MOVE 600,900
250 FOR ANGLE=2*PI/sides TO 2*PI STEP 2*PI/sides
260 DRAW 600+(400*SIN(ANGLE)),500+(400*COS(ANGLE))
270 NEXT ANGLE
280 ENDPROC
290
300 DEFPROCexplain(sides,sides$,name$)
310 IF sides=100 THEN OSCLI"*UTTER <1+> UH S *ER K UH L↑
<1> H a Z UH N *i N F i N UH T N +u M B UH R o V S +I D Z @@
@":ENDPROC
320 OSCLI"*UTTER <1,3> DH i S i Z UH "+sides$+"
```

16

S +I D i D F ★i G UH # # <1+> i T i Z K +OR L D UH "+name$+" @
@ @ @ @"
```
330 ENDPROC
340
350 DATA 3,TH R ★E,T R ★I a NG G UH L
360 DATA 4,F ★OR,K W o D R UH L ★a T UH R UH L
370 DATA 5,F ★I V,P ★e N T UH G UH N
380 DATA 6,S ★i K S,H ★e K S UH G UH N
390 DATA 8,N ★A T,N ★o K T UH G UH N
400 DATA 12,T W ★e L V,D O D ★e K UH G UH N
410 DATA 100,X,X
```

Note: The RETURN characters are not shown explicitly in the above or following programs but their position is, as usual, prior to each line number. Be careful not to enter them in the middle of each true line.

2.6 ★VOICE

Syntax
★VOICE <*pitch*>

This command is used to raise or lower the overall pitch of speech produced by all subsequent ★UTTER commands. Pitch movements within each utterance will be centered on the level set by ★VOICE. The legal values accepted by ★VOICE are as follows:

<H> - high
<MH> – mid-high
<M> - mid
<ML> – mid-low
<L> - low

The entry of any other value, or the omission of either angled bracket will result in the issue of a "Bad pitch" error message. Note that no SPACE is allowed within the brackets. ★VOICE incorporates the effect of ★SYNC; pitch adjustment is delayed until the end of any speech from the latest ★UTTER command, and no separate ★SYNC is required. Note that ★VOICE *only* affects the output of ★UTTER commands, not ★SING

17

At initialisation the SPEECH ROM defaults to the *mid* voice.
This is the one which has the most natural sound. The others
are only intended for special effects.

The next example demonstrates the effect of *VOICE, and
shows how it synchronises with *UTTER.

```
10 *VOICE <L>
20 *UTTER <1+> H +OO Z B i N +E T i NG M *I P +o R i
J a a
30 *VOICE <MH>
40 *UTTER <1+> UH N D H +OO Z B i N +E T i NG M *I P
+o R i J a a
50 *VOICE <H>
60 *UTTER <1+,3> UH N D H *OO Z B i N +E T i NG M *I
P +o R i J # <1+> DH A V *E T N i T <5> *OR L +u P
70 *VOICE <M>
```

3. ERROR MESSAGES

This section documents all the error messages likely to be met
during operation of THE SPEECH ROM, whether originating
from the Operating System or from within THE SPEECH ROM
itself. A list is provided for each message indicating possible
mistakes capable of causing the error. This section is
deliberately rather repetitive in order to make it easier to see all
the potential mistakes at a glance.

SPEECH error messages first specify the error type, followed, if
appropriate, by the location. Parameter errors in *UTTER and
*SING have their location specified by phrase number and, if
necessary, phoneme number. All errors occurring within
programs report the BASIC line number.

Bad filename
Generated by the Operating System in response to an
unrecognised filename. Entering an unabbreviated *command
with no SPACE before its first parameter is one cause of this
error.
e.g. *UTTER<1> E RETURN

Bad command
This error message is issued by the Operating System on
encountering an unknown command. Omitting the full stop
after an abbreviated *command is one cause of this error.
e.g. *U<1> E RETURN

SPEECH ROM is off!
Following any attempt to enter a SPEECH ROM *command
while the system is inactive.

Bad pitch
(a) Use of an invalid value for the pitch parameter in *VOICE

or *SING. Note that the legal values are very different in each of these commands.

(b) Inclusion of SPACE within the angled brackets of the pitch parameter in *VOICE or *SING.

(c) Omission of either angled bracket enclosing the pitch parameter in *VOICE or *SING.

Bad tone

(a) Use of an invalid value for the tone parameter in *UTTER.

(b) Inclusion of SPACE within the angled brackets of the tone parameter in *UTTER.

(c) Omission of either angled bracket enclosing the tone parameter in *UTTER.

Bad phoneme

(a) Use of an invalid value for a phoneme in *UTTER or *SING. The most likely problem is either one or both characters of the phoneme code in the wrong case, or the attempted use of the invalid codes C or X.

(b) Use of any stress marker in *SING

(c) Omission of the obligatory SPACE before any except the first phoneme of a phrase.

(d) Omission of the obligatory SPACE separating the last phoneme of a phrase from the <tone> of a following phrase.

Bad note

(a) Use of an invalid value for the note parameter in *SING.

(b) Inclusion of SPACE within the angled brackets of the note parameter in *SING.

(c) Omission of either angled bracket enclosing the note parameter in *SING.

Bad '*'

Use of the primary stress marker in front of a consonant, a silence, or the vowel UH in *UTTER.

Bad '+'

Use of the secondary stress marker in front of a consonant, a silence, or the vowel UH in *UTTER.

Bad '-'

(a) Use of the accessory stress marker in front of a consonant or a silence in *UTTER.

(b) Use of the accessory stress marker in a phrase containing any tone other than <2v> or <-2v> (in *UTTER).

(c) Use of the accessory stress marker without a primary stress marker at an earlier point in the phrase (in *UTTER).

Too many '*'

(a) Use of more than one primary stress marker in a phrase containing a simple tone (in *UTTER).

(b) Use of more than two primary stress markers in a phrase containing a compound tone (in *UTTER).

Missing '-'

This error occurs if a primary stress marker has been used without a subsequent accessory stress marker in a phrase containing tones <2v> or <-2v> (in *UTTER).

System Crashes

The SPEECH ROM cannot trap two types of error which may lead to a system crash. These problems are discussed below, although neither situation will be encountered with sensible inputs.

(1) Instruction too long

As mentioned earlier an excessive number of phonemes and/or phrases within a single *UTTER or *SING command can, in theory, cause the machine to 'hang' with the SPEECH ROM's buffers each waiting for the other. In practice, however, such overly long commands are completely unnecessary.

(2) Phoneme too short

If any of the phonemes P, T, or K is used in isolation; that is as the only phoneme of its phrase, the speech processor will lock up. This is related to the extreme brevity of the versions of these particular phonemes used when they occur at the beginning of an utterance. The speech processor can be reset with BREAK (followed, if necessary, by OLD RETURN to restore a BASIC program).

Note: there are situations, although they should never be caused by the SPEECH ROM itself, when the speech processor locks up and will *not* reset with BREAK. In such cases resident programs must be saved and the machine restarted).

14. PRONOUNCIATION

To make proper use of the SPEECH ROM the user must gain some familiarity with the principles of the English language on which the system is based. For simplicity and convenience only a sketchy explanation is given which no doubt would not satisfy the academic linguist – apologies! The present section covers pronounciation: the SPEECH ROM phoneme system, its linguistic background, and its relationship to written English. Section-5 covers intonation and rhythm: tone groups and stress markers, their linguistic significance, and their relationship to grammar and meaning.

4.1 Phonetic Background

English, like other languages, can be regarded as being composed of sequences (strings) of speech units called phonemes. Each phoneme is seen as atomic, a minimal unit whose internal structure is not immediately relevant. Only certain strings are permitted by the rules of the language. Changing any phoneme in a string changes the whole utterance. Different dialects of English use somewhat different sets of phonemes, each of which is a subset of the universal set capable of transcribing the sounds of any language. The English dialect used by the SPEECH ROM is appropriately known as BBC English, or more technically, as Received Pronounciation (RP) English. RP English is usually considered to have 44 Phonemes: 20 vowels and 24 consonants. Normally these are transcribed using the symbols of the International Phonetic Alphabet (IPA) – you can see them in the phonetic spellings given for each word in some large dictionaries. Not surprisingly, many of the IPA symbols are unavailable in the standard BBC Microcomputer character set so the SPEECH ROM has to provide its own symbols for the required phonetic input. Each SPEECH ROM symbol is composed of one or two ordinary alphabetic characters, chosen to be mnemonically related to the

sound represented. The IPA version corresponding to each SPEECH ROM symbol is shown in the table below, enabling the user to refer to a dictionary for the phonetic spelling of tricky words.

From here on, to avoid confusion between letters of the alphabet and phonemes, we will adopt the usual convention of placing the latter between slashes '/', wherever they occur in the text or in tables. Slashes are not, of course, required in the input, and so do not appear in example commands or programs.

As schoolchildren know to their cost, English spelling is now in many cases far removed from the sounds it represents. Consequently, when you wish to enter a phonetic string into the SPEECH ROM it is very important to pay close attention to the way you pronounce the word or phrase. Repeat difficult words several times out loud. If the output still sounds wrong, think of what the word rhymes with to suggest to yourself alternative phonetic spellings. At first the correct answer will often be quite surprising. With a little practice, and what amounts to an unlearning of English spelling, the correct entry of most utterances will become second nature, (for example / R u F / instead of rough). In the table below an attempt is made to show the often wide range of conventional spellings which may be associated with a particular phoneme. Example words containing each spelling are given followed by the SPEECH ROM's phonetic form. The examples may be tried by entering ★UTTER <1> followed by the example string and **RETURN**. Use the stress markers but do not worry about their function for the time being. The first example is the word tree, the SPEECH ROM's phonetic representation is given as / T R ★E /, you should enter:

★UTTER <1> T R ★E **RETURN**

4.2 Phoneme Table

The table is divided into two sections, vowels and consonants. Phonemes are also categorised in many other ways, for instance

into those which are *voiced* and those which are not. Voiced phonemes are spoken with the vocal cords vibrating, all vowels are voiced (except when whispered!) but not all consonants. As the voiced/unvoiced distinction is referred to later on, the voiced phonemes are marked in the table with a hash (#).

SPEECH SYMBOL	SPEECH SYMBOL	IPA Spellings	Possible Example words with SPEECH ROM equivalents

a) Vowels

/E/ #	/i:/	ee	tree – T R *E
		e	complete – K UH M P L *E T
		ea	leaf – L *E F
		ie	piece – P *E S
		ei	ceiling – S *E L i NG
		ae	Aesop – *E S o P
/i/ #	/ɪ/	i	sit – S *i T
		y	pity – P *i T +i
		e	pretty – P R *i T +i
		ie	cities – S *i T i Z
		a	village – V *i L i J
/e/ #	/e/	e	bed – B *e D
		ea	dead – D *e D
		a	many – M *e N +i
		oe	foetid – F *e T i D
/a/ #	/æ/	a	bat – B *a T
		ai	plait – P L *a T
/u/ #	/ʌ/	u	sun – S *u N
		o	son – S *u n
		ou	country – K *u N T R +i
		oo	blood – B L *u D
		oe	does – D *u Z
/AR/ #	/ɑ:/	a	pass – P *AR S
		ar	part – P *AR T
		ea	heart – H *AR T

		er	clerk – K L *AR K
		al	calm – K *AR M
		au	aunt – *AR N T
/o/ #	/ɒ/	o	dog – D *o G
		a	was – W *o Z
		ou	cough – K *o F
		ow	knowledge – N *o L i J
		au	because – B UH K *o Z
/OR/ #	/ɔ:/	or	horse – H *OR S
		aw	saw – S *OR
		ou	bought – B *OR T
		au	fault – F *OR L T
		a	tall – T *OR L
		ore	before – B UH F *OR
		oor	door – D *OR
		oar	board – B *OR D
		our	four – F *OR
		ure	sure – SH *OR
/oo/ #	/ʊ/	u	put – P *oo T
		o	woman – W *oo M UH N
		oo	good – G *oo D
		ou	could – K *oo D
/OO/ #	/u:/	oo	food – F *OO D
		o	lose – L *OO Z
		ou	soup – S *OO P
		u	rude – R *OO D
		ew	chew – CH *OO
		ue	blue – B L *OO
		ui	juice – J *OO S
		oe	shoe – SH *OO
/ER/ #	/ɜ:/	er	her – H *ER
		ir	first – F *ER S T
		ear	heard – H *ER D
		ur	turn – T *ER N
		or	world – W *ER L D
		our	journey – J *ER N +i

	urr	purr – P *ER	
	err	err – *ER	
	olo	colonel – K *ER N UH L	
/A/ # /eɪ/	a	late – L *A T	
	ay	day – D *A	
	ai	waist – W *A S T	
	ea	great – G R *A T	
	ei	veil – V *A L	
	ey	they – DH *A	
/I/ # /aɪ/	i	time – T *I M	
	y	cry – K R *I	
	ie	die – D *I	
	ei	either – *I DH UH	
	ai	aisle – *I L	
	ye	dye – D *I	
	uy	buy – B *I	
	eye	eye – *I	
	igh	high – H *I	
	eigh	height – H *I T	
/O/ # /əʊ/	o	both – B *O TH	
	oa	road – R *O D	
	oe	toe – T *O	
	ou	soul – S *O L	
	ow	know – N *O	
	ew	sew – S *O	
	oo	brooch – B R *O CH	
	au	mauve – M *O V	
/OY/ # /ɔɪ/	oy	toy – T *OY	
	oi	voice – V *OY S	
/AW/ # /aʊ/	ou	house – H *AW S	
	ow	town – T *AW N	
/EA/ # /ɪə/	ea	idea – I D *EA	
	ear	fear – F *EA	
	eer	peer – P *EA	
	ier	pier – P *EA	

26

	ere	here – H *EA	
	ir	fakir – F *A K EA	
	ia	Ian – *EA n	
	eu	museum – M Y oo Z *EA M	
	eo	theory – TH *EA r +i	
	e	hero – H *EA R O	
/ea/ # /ɛə/	ear	pear – P *ea	
	are	care – K *ea	
	air	fair – F *ea	
	ere	there – DH *ea	
	eir	their – DH *ea	
	ae	aerobic – ea R *O B i K	
	ar	scarce – S K *ea S	
	a	vary – V *ea r +i	
/UR/ # /ʊə/	ure	lure – L *UR	
	our	tour – T *UR	
	ue	fluent – F L *UR N T	
	ua	skua – S K Y *UR	
		(Note: /Y *UR/)	

/UH/ #   /ə/

This is a very special vowel sound in English, known as the Reduced Vowel or *schwa* (from the Hebrew) and is a very neutral sound. Note that it *never* occurs in stressed syllables, but that in unstressed syllables, and particularly in fast connected speech, many other vowels tend to neutralise, becoming the schwa. You will have noticed several words in the above examples which contain /UH/. Listen carefully again to your own speech; take for example the word **again**. The most pedantic pronounciation would be /a G *A N/, but you would more often say /UH G *A N/, or if you were speaking even faster /UH G *e N/. Incidentally, the second alteration demonstrates another type of vowel change. Try entering these variations and listen to the difference. The appropriate use of vowel reduction, because it more closely simulates real speech, improves rhythm and adds greatly to naturalness. On the other hand, in some circumstances and especially with isolated words, the use of unreduced vowels may improve comprehensibility.

27

(b) Consonants

/P/	/p/	p	pin – P *i N
		pp	upper – *u P UH
		silent	psychic
/B/ #	/b/	b	bat - B *a T
		bb	rubber – R *u B UH
/T/	/t/	t	ten – T *e N
		tt	letter – L *e T UH
		th	Thames - T *e M Z
/D/ #	/d/	d	din - D *i N
		dd	ladder - L *a D UH
/K/	/k/	k	kind - K *I N D
		c	cane - K *A N
		ck	deck - D *e K
		cc	accord - UH K *OR D
		qu	conquer - K *o NG K UH
		ch	chemist - K *e M i S T
		x	tax - T *a K S
			(Note: /K S/)
		silent	knit
/G/ #	/g/	g	gap - G *a P
		gg	sluggish - S L *u G i SH
		gh	ghost - G *O S T
		gu	guild - G *i L D
		x	exact - e G Z *a K T
			(Note: /G Z/)
		silent	gnome
			sign
/CH/	/tʃ/	c h	chum - CH *u M
		tch	itch - *i CH
		t	nature - N *A CH UH
			righteous - R *I CH UH S
			question - K W *e S CH UH N
/J/ #	/dʒ/	j	just - J *u S T
		g	generous - J *e N UH R UH S

28

		gg	suggest - S UH J *e S T
		dj	adjust - UH J *u S T
		dg	midget - M *i J i T
		de	grandeur - G R *a N J UH
		di	soldier - S *O L J UH
		ch	Norwich - N *OR R i J
/F/	/f/	f	feet - F *E T
		ff	office - *o F i S
		ph	physics - F *i Z i K S
		gh	enough - i N *u F
/V/ #	/v/	v	vote - V *O T
		f	of - *o V
		ph	nephew - N *e V Y OO
/TH/	/θ/	t h	thought - TH *OR T
/DH/ #	/ð/	t h	them - DH *e M
/S/	/s/	s	soft - S *o F T
		ss	pass - P *AR S
		c	fleece - F L *E S
		sc	science - S *I UH N S
		x	axe - *a K S
			(Note: /K S/)
		silent	isle
/Z/ #	/z/	z	zoo - Z *OO
		zz	fizzy - F *i Z +i
		s	nose - N *O Z
		ss	scissors - S *i Z UH Z
		x	exact - i G Z *a K T
			(Note: /G Z/)
/SH/	/ʃ/	s h	shoe - SH *OO
		ch	machine - M UH SH *E N
		sch	schedule - SH *e D Y oo L
		s	sure - SH *OR
		ss	assure - UH SH *OR
		ti	ration - R *a SH UH N

29

	si	tension - T *e N SH UH N	
	ssi	fission - F *i SH UH N	
	sci	conscience - K *o N SH UH NS	
	ci	special - S P *e SH UH L	
	ce	ocean - *O SH UH N	
	x	luxury - L *u K SH UH R +i	
		(Note: /K SH/)	
/ZH/ # /ʒ/	z	seizure - S *E ZH UH	
	s	measure - M *e ZH UH	
	si	vision - V *i ZH UH N	
	ge	beige- B *A ZH	
/H/ /h/	h	hot - H *o T	
	wh	which - W *i CH	
	silent	rhythm	
/M/ # /m/	m	model - M *o D UH L	
	mm	summer - S *u M UH	
	mb	comb - K *O M	
	mn	autumn - *OR T UH M	
/N/ # /n/	n	neat - N *E T	
	nn	funny - F *u N +i	
	kn	knock - N *o K	
	gn	gnome - N *O M	
	pn	pneumonia - N Y oo M *O N EA	
/NG/ # /ŋ/	ng	fling - F L *i NG	
		angle - *a NG G UH L	
		(Note: /NG G/)	
	n	think - TH *i NG K	
/L/ # /l/	l	land - L *a N D	
	ll	small - S M *OR L	
	silent	talk	
/R/ # /r/	r	rent - R *e N T	
	rr	arrow - *a R O	
	wr	wrong - R *o NG	

30

	rh	rhyme - R *I M	
	silent	large	
		mover	
/Y/ # /y/	y	young - Y *u NG	
	i	spaniel - S P *a N Y UH L	
	u	muse - M Y *OO Z	
	ew	few - F Y *OO	
	eu	feud - F Y *OO D	
	eau	beauty - B Y *OO T +i	
		(Note: the last four examples /Y OO/)	
	ui	suit - S *OO T	
		(Note also: S Y *OO T)	
/W/ # /w/	w	wish - W *i SH	
	wh	when - W *e N	
	u	quick - K W *i K	
		language - L *a NG G W i J	
		suite - S W *E T	
	o	one - W *u N	
		(Note: /W u/)	
		choir - K W *I R	

4.3 Hints on Common Phonetic Problems

This section contains an assortment of unrelated pieces of information intended to help the user avoid some of the more common mistakes made in trying to create a phonetic text of the type required as input to the SPEECH ROM.

(a) There is usually a tendency to assume that all occurrences of the letter s in plurals represent the phoneme /S/, whereas when the preceding phoneme is *voiced* (see above) /Z/ should be used instead. Contrast the following examples.

M *a T S RETURN (mats)
R *E F S RETURN (reefs)
T *I M Z RETURN (times)
B *A Z RETURN (bays)

31

Note that this voicing only applies to plurals, as the next example shows.

H *AW S`RETURN` (house)
H *AW Z i Z`RETURN` (houses)

In the singular the final consonant is an unvoiced /S/ in spite of the preceding voiced phoneme, however in the plural the same consonant has a vowel on both sides and changes from /S/ to /Z/. The final consonant in the plural form is /Z/ as expected. The best way to obtain the correct phoneme is to repeat each word out loud thereby sensitising yourself to the sounds of the language.

(b) In the same way that the neutral vowel /UH/ frequently (and only) occurs in unstressed syllables, so does the vowel /i/ (but not quite exclusively so). Because it is a very short vowel, it, like /UH/, contributes markedly to rhythm. It will often be discovered that other vowels have mistakenly been used where /i/ would be correct. Note the use of /i/ in the words "village" and "enough" in the phoneme table. The letter "e" is particularly likely to cause difficulty, for example in the plural "houses" in the previous section you will have noticed it transcribed as /i/. Take also the two renderings of the next phrase; the first as spoken carefully, the second as it would more normally be pronounced at speed. The phrase is "He has":

H E H *a Z`RETURN`
H i H *a Z`RETURN`

(c) There are two distinct phonemes in English which are spoken with the tongue between the teeth, one voiced the other not. Unfortunately we spell them both the same way; that is "th" (unlike Welsh which you may have noticed has dd and th). Compare the initial sound of "this", which is voiced and represented by /DH/, with that of "thin", which is unvoiced and represented by /TH/. The next example contrasts /DH/ with /TH/ ocurring in the middle of words.

B *o DH UH`RETURN` (bother)
*e TH i K S`RETURN` (ethics)

32

(d) The letter "x" may be a little confusing as it gives rise to combinations of two phonemes, (/K S/, /G Z/ and /K SH/). Take the following three cases.

F *a L a NG K S`RETURN` (phalanx)
i G Z *a K T`RETURN` (exact)
N *o K SH UH S`RETURN` (noxious)

(e) As you will by now have realized there is an important phoneme represented by /NG/, for which English has no single letter. /NG/ is the sound that occurs at the end of "sing" and is usually written down as "ng". The sound is produced by vibrating the vocal cords, with the back of the tongue touching the roof of the mouth, and the nasal passage open. When the letter "n" occurs in front of another letter giving rise to a phoneme with this same tongue position, the "n" is actually pronounced /NG/. The letters in question are "k", "g" and "x".

K *o NG K UH`RETURN` (conker)
*a NG G UH`RETURN` (anger)
*a NG K SH UH S`RETURN` (anxious)

Notice in the second example that although the spelling is "ng" it gives rise to /NG G/.

In connected speech the effect of one phoneme on the preceding one can operate across a word boundary, as in the following example.

L *i NG K UH N`RETURN` (Lincoln)
L i NG K UH NG G R *E N`RETURN` (Lincoln Green)

Note that the phoneme /NG/ only occurs in the middle and at the end of English words, never at the beginning.

(f) You will see from the phoneme table that the possible spellings given for many vowel sounds include the letter "r". For instance under /EA/ will be found "ear" and "eer". When such written sequences occur in the last syllable of an isolated word no /R/ phoneme is to be heard in the spoken

33

version. The "r" can either be regarded as contributing to the vowel sound or as being silent. However if another syllable, beginning with a vowel is appended then the /R/ becomes manifest.

H *EA`RETURN` (hear)
and H *EA R i NG`RETURN` (hearing)

S T *EA`RETURN` (steer)
and S T *EA R i NG`RETURN` (steering)

As with the previous section this effect can occur across word boundaries.

F *AR`RETURN` (far)
and F AR R *AW T`RETURN` (far away)

P *ea`RETURN` (pair)
and P +ea R UH V T W *i N Z`RETURN` (pair of twins)

Pronouncing the /R/ in such cases is often claimed to be a sign of sloppy speaking habits. In fact it happens to at least a minor extent in everyone's speech. Worse still for the purist there is a tendency for the /R/ to intrude in similar situations even where there is no "r" in the written form, both across word boundaries and within words (the last being the worst sin of all).

L OR R a N D *OR D UH`RETURN` (Law and order)
D R *OR R i NG`RETURN` (drawing)

The naturalness of speech from SPEECH ROM can often be improved by judicious use of the /R/ phoneme as a *link*. In positions where a non-terminal "r" is followed by a vowel in the written form, inclusion of /R/ is nearly always preferable.

(g) Another link phoneme which sometimes appears between vowels, even when not represented in the written form is /Y/. Once again it seems to smooth the vowel transition. In some circumstances vowel sequences in the SPEECH ROM

such as / i a / on their own give sufficient impression of an intervening /Y/ in others it may help specifically to insert one.

P +E D i Y a T R *i SH UH N`RETURN` (paediatrician)
TH R i Y *e L UH F UH N T S`RETURN` (three elephants)

In the second example above / i Y / can be replaced by / E Y / or / E / alone, the preferred output is often a matter of taste.

(h) Sometimes, when the phoneme /OO/ follows certain consonants you will notice the need to insert the phoneme /Y/ before the vowel.

T *OO L`RETURN` (tool)
S K *OO L`RETURN` (school)
R *OO L`RETURN` (rule)
M Y *OO L`RETURN` (mule)
V Y *OO`RETURN` (view)
N Y *OO S UH N S`RETURN` (nuisance)

(i) There are slight variations of pronounciation even within RP English. One example of this is to be found in the alternative forms of words like "poor" and "sure" rhyming either with "shore" or "lure".
"poor" may be pronounced
P *OR`RETURN`
or P *UR`RETURN`

"sure" may be pronounced
SH *OR`RETURN`
or SH *UR`RETURN`

(j) If the word "the" is followed by a word beginning with a consonant the vowel is often reduced from / E / to /UH/. This happens mostly in fast speech. (In childrens' speech the same thing happens before words beginning with vowels).

DH E B UH G *i N i NG`RETURN` (the beginning)
or DH UH B UH G *i N i NG`RETURN`
DH E *e N D`RETURN`

5. INTONATION

5.1 Linguistic Background

This section is likely to contain by far the most novel information for the average user. An attempt has been made to simplify the linguistic background to an absolute minimum, but it is important to develop some familiarity with the concepts involved if the SPEECH ROM is to be used to its full potential. Each of the following headings introduces a single essential feature of English intonation. As previously, linguistic accuracy has been sacrificed for simplicity, and it is also important to bear in mind that the SPEECH ROM uses only one of a number of distinct ways of looking at the subject.

(a) Stress

During speech, some syllables seem more prominent than others; they are said to have greater *stress*. The factors contributing to the perception of stress are very complex and include vowel length, loudness, pitch change and vowel quality. Under the direction of stress markers in the input, the first three of these are handled automatically, and the user has some control over the last; for instance the substitution of other vowels by /UH/ in unstressed syllables.

(b) Wordstress

When an English word is spoken in isolation, not all of its syllables carry the same stress. One in particular will be stressed more than the rest; usually this syllable is part of the *stem* of the word and only rarely part of a prefix or an *ending*, it is sometimes called the *salient* syllable. We will be concerned with the salient syllable and occasionally, in long words, with the second most prominent syllable.

The stress pattern within a word is an abstract notion; a potential only realised when the word is spoken. In the case of words spoken in isolation the actual stress pattern is identical to the potential one.

When words are strung together their potential stress patterns are not always realised. Some words in a sequence are more prominent than the others; they achieve this prominence by realising their potential stress. The other words remain unstressed.

Monosyllabic words are always stressed when spoken singly. In connected speech monosyllabic words may be stressed if they are *content* words, but are unlikely to be stressed if they are *function* words (function words are best described as the ones you would leave out in a telegram; articles, parts of the verb to be, etc., content words are those you would leave in).

Some examples may help to clarify these ideas. Both the written and phonetic versions of each example are given, salient syllables are marked by an apostrophe before the vowel. (Ignore, for the present, the distinction between + and *). First of all some isolated words:

```
P *E T UH RETURN   (P'eter)
H *E RETURN   (H'e)
L *I K S RETURN   (l'ikes)
CH *e D UH RETURN   (Ch'eddar)
CH *E Z RETURN   (ch'eese)
DH *E RETURN   (th'e)
R *A N RETURN   (r'ain)
*i N RETURN   ('in)
S P *A N RETURN   (Sp'ain)
R UH V *E L D RETURN   (rev'ealed)
u N i M *a J UH N UH T i V RETURN   (unim'aginative)
```

and now making phrases out of some of them:

"He l'ikes the Ch'eddar ch'eese"
```
H i L +I K S DH UH CH +e D UH CH *E Z RETURN
```

"The r'ain in Sp'ain"
```
DH UH R +A N i N S P *A N RETURN
```

Several of the words clearly lose their stress when placed in a phrase, but the location within a word of retained stress does

not change. Note that the definite article not only loses its stress but undergoes vowel reduction; only rarely for contrastive purposes would the stress be retained. Observe that the salient syllable of the longer words is located in their stem.

(c) Rhythm

Some languages, such as French, have a very regular rhythm, each syllable having roughly the same length; they are said to be *syllable timed* languages. English is a *foot timed* language; each phrase being divided into a number of *feet* which are rather similar to bars in music. Each foot tends towards the same duration irrespective of the number of syllables it contains. The first syllable of a foot is stressed and therefore takes the beat. Feet with no stressed syllable have an initial silent beat (for instance this often occurs in the first foot of a phrase, just as when music begins in the middle of a bar). Nursery rhymes, with their exaggerated rhythm, provide excellent examples with which to work.

L'ittle Miss M'uffet,
she s'at on a t'uffet,
'eating her c'urds and wh'ey

This example contains seven feet of 3, 3, 3, 2, 3, 2 and 1 syllable(s) respectively. Note that the first beat of the second line falls on "sat". The full SPEECH ROM version of these lines (with tones included) would be:

```
*U.<3> L +i T UH L M i S M *u F i T↑SH i
*U.<1+> S +a T o N UH T *u F i T #
*U.<5> *E T i NG H ER <1-> K +ER D Z UH N D W *A
```

The primary and secondary stress markers indicate that their syllable is the first one of a foot; that it carries one beat of the rhythm. The secondary stress marker, +, has nothing more than this function. Speech produced when + is used correctly will exhibit a much more natural rhythm. It is well worth while reciting any phrase you want to use as if it were a line of doggerel in order to emphasise and reveal its rhythm; you can even try tapping out the beat at the same time. Techniques like these are extremely useful, particularly at first, in helping you really to hear how you say things.

38
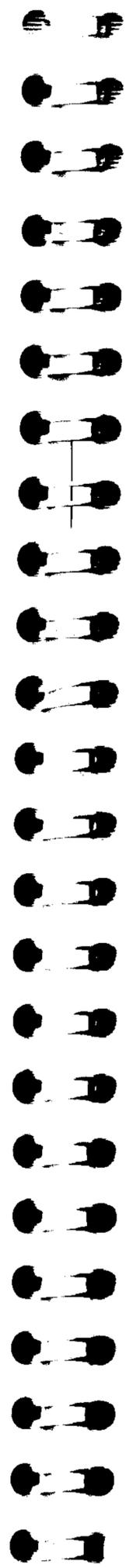
(d) Tone Groups

Up until this point we have been referring to a *UTTER phoneme string, with its associated intonation pattern, very loosely as a *phrase*, and this has not been further defined. The term phrase as it has been used in this manual corresponds, more or less, to what is technically known as a Tone Group; a Tone Group is the smallest possible utterance which possesses a complete *melody* of its own.

Normally a Tone Group is used to convey a single piece of information; often it will correspond to a clause in grammar, but not necessarily so. Single words can also carry a complete Tone Group, as in some of the above examples.

There are clearly an indefinite number of possible pitch patterns in English, but fortunately they can be reduced to a smallish number of fundamental categories. The SPEECH ROM provides 24 standard pitch contours or *tones*, each one representative of a single category. The pitch contours are imposed on the subsequent phoneme string by the <tone> parameter.

Tones serve a number of different functions in speech. Choosing one rather than another may alter the *force* of an utterance, making it sound, for example, like a question rather than a statement. In another context the choice of tone may indicate the speaker's attitude, either to the information given or to the listener. In still other circumstances the appropriate tone may largely be dictated by the underlying grammatical form.

Do not despair! some of the 24 tones are much commoner than the others; they are more neutral in the sense that they can always be used if you are in doubt. As you become familiar with the SPEECH ROM you can experiment with the less common tones, which are only appropriate in more restricted circumstances, but which can be used to create much more subtle effects.

(e) Tonic Syllable

Each Tone Group, being a unit of information, has a focus; a word which carries *new* information, the rest being *understood*

39

either from the context or because it has already been mentioned. New information most naturally falls at the end of a phrase, and will only occur earlier if special emphasis is required. Most commonly the focus is on the last content word of the phrase.

The salient syllable of the focal word is called the *tonic syllable* of the Tone Group. It is distinguished in the SPEECH ROM by the use of the primary stress marker, *, instead of the secondary stress marker, +. The following examples demonstrate the neutral case in which the focus is at the end of the phrase:

*UTTER <1> DH UH P +I L UH T S N +A M i Z Z +A F o D B *E B UH L B R o K S `RETURN`

*UTTER <1> L +i N D B ER G W o Z DH UH F +ER S T M a N T oo F L +I DH E UH T L a N T i K O SH UH N S *O L O `RETURN`

*UTTER <1> W +i CH UH V DH UH K a NG G UH R +OO Z J u M P S F *ER DH i S T `RETURN`

As you can hear, the tonic syllable is very dominant; this is partly because its vowel is louder and longer than vowels with secondary stress (which are themselves louder and longer than unstressed vowels). There is, however, a much more important reason; a large change of pitch occurs just before or within the tonic syllable.

The large pitch movement effectively divides the Tone Group into two; the first part, before the tonic syllable, is called the *pretonic segment*, the rest, including the tonic syllable itself, is called the *tonic segment*. (Note that if the tonic syllable occurs at the very beginning of the Tone Group, the pretonic segment is missing.)

Each of the SPEECH ROM's tones selects a particular pitch contour for the first segment of the Tone Group and another for the second segment. The pitch contours are then tailored to fit the phoneme sequence according to the position of the primary stress marker. By moving the primary stress from its neutral

position at the end of the utterance, different words can be made the focus for purposes of emphasis, contrast or insistence. Try the following variations of the previous examples:

*UTTER <1> DH UH P *I L UH T S N +A M i Z Z +A F o D B +E B UH L B R o K S `RETURN`
(not the Navigator's)

*UTTER <1> DH UH P +I L UH T S N *A M i Z Z +A F o D B +E B UH L B R o K S `RETURN`
(rather than his nervous condition?)

*UTTER <1> DH UH P +I L UH T S N +A M i Z Z *A F o D B +E B UH L B R o K S `RETURN`
(not Eric Beeblebrox)

In order to emphasise the word which carries the tonic syllable in the neutral form it is necessary to use a different tone:

*UTTER <1+> DH UH P +I L UH T S N +A M i Z Z +A F o D B *E B UH L B R o K S `RETURN`
(not Zaphod Entwhistle)

*UTTER <1> L *i N D B ER G W o Z DH UH F +ER S T M a N T oo F L +I DH E UH T L a N T i K O SH UH N S +O L O `RETURN`
(not Bleriot)

*UTTER <1> L +i N D B ER G W *o Z DH UH F +ER S T M a N T oo F L +I DH E UH T L a N T i K O SH UH N S +O L O `RETURN`
(I'm sure I'm right)
(Note the extra foot)

*UTTER <1> L +i N D B ER G W o Z DH UH F *ER S T M a N T oo F L +I DH E UH T L a N T i K O SH UH N S +O L O `RETURN`
(not the second)

*UTTER <1> L +i N D B ER G W o Z DH UH F +ER S T M *a N T oo F L +I DH E UH T L a N T i K O SH UH N S +O L O `RETURN`
(not the first woman)
(Note the extra foot)

*UTTER <1> L +i N D B E R G W o Z DH UH F +ER S T M a
N T oo F L *I DH E UH T L a N T i K O SH UH N S +O L O
`RETURN`
(he didn't swim it)

*UTTER <1> L +i N D B ER G W o Z DH UH F +ER S T M a
N T oo F L +I DH E UH T L *a N T i K O SH UH N S +O
L O `RETURN`
(not the Pacific)
(Note the extra foot)

*UTTER <1> L +i N D B E R G W o Z DH UH F +ER S T M a
N T oo F L +I DH E UH T L a N T i K *O SH UH N S +O
L O `RETURN`
(as opposed to the Atlantic puddle)

*UTTER <1+> L +i N D B E R G W o Z DH UH F +ER S T M
a N T oo F L +I DH E UH T L a N T i K O SH UH N S *O
L O `RETURN`
(Alcock and Brown did it together)
(Note again the use of tone <1+>)

*UTTER <1> W +i CH UH V DH UH K a NG G UH R *OO Z J
u M P S F +ER DH i S T `RETURN`
(I'm not interested in the wallabies)

*UTTER <1> W +i CH UH V DH UH K a NG G UH R +OO Z
J *u M P S F +ER DH i S T `RETURN`
(compared to throwing the Boomerang)
(Note the extra foot)

*UTTER <1+> W +i CH UH V DH UH K a NG G UH R +OO Z
J u M P S F *ER DH i S T `RETURN`
(as opposed to highest)
(Note the use of tone <1+>)

In the last series of examples, in order to emphasise the
interrogative pronoun, "which", it is really necessary to use a
different tone:

*UTTER <2v> W *i CH UH V DH UH K a NG G UH R +OO Z
J u M P S F .ER DH i S T `RETURN`
(Request to be reminded or surprised response to being told.)

42

Although the explanations of these ideas and the instructions
for their use may seem complicated at first reading, you will
find the system much more friendly in practice.

There are a few additions to the structure of Tone Groups
introduced by the use of compound tones and tones <2v> and
<-2v>, but these will be dealt with later under each individual
tone.

5.2 Meanings and Uses of Tones

(a) Major Categories

Tones are divided into five main categories (indicated by the
number in the tone code) on the basis of the pitch contour of
their tonic segment. Very generally, a falling pitch indicates
certainty and a rising pitch uncertainty; but uncertainty about
whether or not something is the case rather than details of
"who", "when", "why", "how", etc. (From here on we will make
this distinction between questions expecting a "yes/no" answer,
and WH-questions, starting with an interrogative word like those
above and expecting a specific answer).

Tone <1>

This has a pitch which begins to fall on the tonic syllable and
continues to do so to the end. Here is a selection of some of the
more common of the many uses of tone <1>:

Neutral statement:
*UTTER <1> DH +i S i Z UH S T *A T M UH N T `RETURN`
*UTTER <1> S +O i Z DH *i S `RETURN`

Neutral WH-question:
*UTTER <1> W +o T i Z i T UH B *AW T `RETURN`
*UTTER <1> W +I W O N T Y oo T *e L M +i `RETURN`

Forceful or impatient yes/no question:
*UTTER <1> W +i L Y +oo B +i K W *I UH T `RETURN`

Statement of observation or deduction used as a question:
*UTTER <1> P +E T UH Z G +O i NG H *O M DH +e N
`RETURN`

43

Neutral positive command:
*UTTER <1> T +A K o F DH a T S +i L i H *a T RETURN

Forceful negative command:
*UTTER <1> D +O N T P oo T Y UH D +OR T UH R o N DH
UH S T *A J M i S i Z W +ER DH i NG T UH N RETURN

Neutral response to a question:
*UTTER <1> Y *e S ↑ <1> I K *a N RETURN
*UTTER <1> N *O ↑ <1> I K *AR N T RETURN

Exclamation:
*UTTER <1> SH *o K i NG RETURN

Greeting:
*UTTER <1> G oo D M *OR N i NG RETURN

Tone <2>
This tone has a rising tonic segment. Like tone <1> it has a
large number of uses of which these are just some:

Neutral yes/no question:
*UTTER <2> +i Z DH +i S UH K W *e S J UH N RETURN
*UTTER <2> D +u Z i T M +A K e N i S *e N S RETURN

Tentative or deferential WH-question:
*UTTER <2> W +o T T I M i Z BR *e K F UH S T RETURN

Statement used as question seeking confirmation (cf. d) above):
*UTTER <2> P +E T UH Z G +O i NG H *O M DH +e N
RETURN

Unfavourable response; denial, contradiction or frustration of
expectation (cf. g) above):
*UTTER <2> N *O ↑ <1> I K *AR N T RETURN
*UTTER <2> DH +a T +i Z N T DH UH W *u N RETURN

Tone <3>
You should use tone <3> for statements which are dependent
or in some way incomplete. It has a slight rise in pitch at the
end which leaves an *unfinished* feeling. Uses of tone <3>
include:

Situations where one fact is incomplete without another:
*UTTER <3> H +AR F W A u P DH UH S T *ea Z ↑ <1> i
Z UH S T +ea W ea R I S *i T RETURN
*UTTER <3> L +i V UH P OO L W *u N ↑ S O <1> +e V
UH T UH N L *o S T RETURN

Acceding to a request or unexpressed expectation:
*UTTER <3> I L S +E W o T +I K UH N D *OO RETURN

Reassurance:
*UTTER <3> i T UH L S +OO N B E OR L *O V
UH RETURN

Use of a name in warning:
*UTTER <3> e L *i Z UH B UH TH RETURN

Goodbyes:
*UTTER <3> G oo D B *I RETURN
*UTTER <3> S +E Y oo T oo M *o R O RETURN

Tone <4>
The tonic of tone <4> first falls in pitch and then rises; initially
implying certainty and then immediately raising doubts again.
Tone <4> is the next most common after tone <1> and has a
number of uses mostly involving a sense of reservation or
contrast. It often implies *there is more to it than that*:

Situations where one statement makes reservations about
another:
*UTTER <4> Y oo K oo D *AR S K M E <1> ↑ B UH T I M
+I T R i F Y *OO Z RETURN
*UTTER <4> i T S +OR L R +AR DH UH K L *e V UH <1>
↑ B UH T I D +O N T TH +i N K i T UH L W *ER K RETURN

Where one fact is circumstantial to another:
*UTTER <4> W e N DH UH K *a T S UH W +A # # <1> DH
UH M *I S W i L P L +A RETURN

Personal opinion or evaluation offered for consideration:
*UTTER <4> I L +I K DH UH N Y OO *o P UH R A T i NG
S +i S T UH M RETURN

44

45

Command expressing compromise or concession:
*UTTER <4> L +i S UH N T oo i T F *ER S T RETURN
(even if you don't think you'll like it)

Tone <5>
In some ways the opposite of tone <4>, the tonic of tone <5> first rises hesitantly in pitch and then falls on a note of greater certainty. It is often used assertively or to express surprise at the listener's opinion:

Strong assertion:
*UTTER <5> DH a T S +OR L DH ea R +i Z T *OO DH UH M +a T UH RETURN

Expressing commitment:
*UTTER <5> I P R *o M i S T oo B i K W +I UH T RETURN

Expression of surprise:
*UTTER <5> R UH M *AR K UH B UH L RETURN

Reproach
*UTTER <5> DH a T W *o Z N T V e R i K +I N D RETURN

Willing agreement:
*UTTER <5> UH V K *OR S RETURN

(b) Variations of Major Categories
So far we have discussed the neutral forms of each of the major intonation categories, we now turn to their further variations. These special forms are more restricted in use and generally are intensified versions of the neutral ones. They either convey particular personal involvement on the part of the speaker or imply extra emphasis or contrast.

Tone <1+>
This is a more emphatic or excited version of tone <1>; it may also express surprise or intolerance:

Emphasis:
*UTTER <1+> DH i S i Z DH UH L +AR S T T *I M RETURN

Expressions of excitement:
*UTTER <1+> i T S S T +AR T i D S N *O i N G RETURN

Expressing the unexpected:
*UTTER <1+> SH i Z F +i N i SH D H ER D *i N UH RETURN

Agreements:
*UTTER <1+> N *a CH R UH L i RETURN
*UTTER <1+> UH V K *OR S RETURN

Expressions of irritation or intolerance:
*UTTER <1+> Y OR V +e R i S J *OO P i D RETURN

Tone <1->
Tone <1-> is the opposite of tone <1+>; giving the appearance of resignation or tolerance. It is also used to make suggestions:

Resignation:
*UTTER <1-> i T S M +u N D A M *OR N i N G RETURN

Expressing the expected:
*UTTER <1-> i T S S T +AR T i D R *A N i N G RETURN

Making suggestions
*UTTER <1-> Y oo K oo D P +A N T i T G R *E N RETURN

Tone <u1>
Placing u before the 1 has the effect of giving the pretonic segment uneven pitch, each foot receives additional emphasis. It is mostly used with 1+ in tone <u1+>, and the result is very forceful or querulous:

Forceful statements:
*UTTER <u1+> P +oo T DH UH M W +ea Y OO *OR L W A / P +oo T DH UH M RETURN

Querulousness:
*UTTER <u1> W +I UH Y OO +OR L W A Z L +A T F UH R *e V R i TH i N G RETURN

Tone <-2>

Tone <-2> is more intense than tone <2>, and can be used to express surprise, indignation, disapproval or even concern:

Exclamatory questions:
*UTTER <-2> W +I N *o T RETURN

Expressions of indignation:
*UTTER <-2> i T W +o Z N T M *I F +OR L T RETURN

Showing disapproval:
*UTTER <-2> AR DH A *OR L D +ER T i RETURN

Expressions of sympathy or concern:
*UTTER <-2> D +u Z i TH *ER T V +e R i M +u CH
RETURN

Tone <2v>

The v in tone <2v> (or <2v>) indicates that the pitch of the tonic segment falls steeply and then rises again in a very angular way. The fall occurs on the tonic syllable, marked as usual by *. The rise occurs on a subsequent word (usually the last content word) marked by . ; that is with accessory stress. Tones with v are used to ask questions where the query concerns a specific point in the phrase. Take the following series, starting with the neutral question:

*UTTER <2> D +i D L +i N D B ER G F L +I DH E UH T L
+a N T i K F *ER S T RETURN

*UTTER <2v> D *i D L +i N D B ER G F L +I DH E UH T
L +a N T i K F . ER S T RETURN
(are you sure you have it right?)

*UTTER <2v> D +i D L *i N D B ER G F L +I DH E UH T
L +a N T i K F . ER S T RETURN
(or was it Bleriot?)

*UTTER <2v> D +i D L +i N D B ER G F L *I DH E UH T
L +a N T i K F . ER S T RETURN
(or did he swim?)

48

*UTTER <2v> D +i D L +i N D B ER G F L +I DH E UH T
L *a N T i K F . ER S T RETURN
(or only after he'd done the Pacific?)

*UTTER <2v> D +i D L +i N D B ER G F L +I DH E UH T
L +a N T i K F *ER S T RETURN
(before he had the crash!)
(Note that if the last syllable is emphasised no . is required)

Tone <-3>

This version of tone <3> tends to give the impression of either unconcern or uncertainty. It is also used for cautions.

Expressing unconcern:
*UTTER <-3> N +O W u N N *O Z UH B +AW T i T
RETURN

Expressing uncertainty:
*UTTER <-3> H i M +I T UH V D *u N RETURN

Uncertain acceptance:
*UTTER <-3> I K +oo D i F Y oo W *o N T RETURN

Issuing cautions:

*UTTER <-3> K *ea F UH L RETURN
*UTTER <-3> L oo K *AW T RETURN

Tone <4->

Tone <4-> is intense; it adds a definite, contrastive or exclusive quality to the meaning of tone <4>. It often conveys a rather disgruntled or sulky frame of mind:

Expressing strong reservation:
*UTTER <4-> I M SH +OR R i T S T +OO L +A T N
*AW RETURN

Grudging agreement:
*UTTER <4-> Y oo K +a N i F Y oo L *I K RETURN

Implying "at least"; justifying:
*UTTER <4-> i T S G +o T UH N +I S B *o K S RETURN
*UTTER <4-> Y oo M +u S T UH D M *i T RETURN

49

Implying "if only"; bemoaning:
*UTTER <4-> Y oo M +I T UH V S *e D RETURN
*UTTER <4-> I W +i SH DH +A D G O H *O M RETURN

Tone <5->

An intensified form of tone <5>, this conveys very strong emotions such as awe, it can also be used sarcastically.

Expressing awe:
*UTTER <5-> +i N K R *e D UH B UH L RETURN
*UTTER <5-> H +AW D R *e D F UH L RETURN

Showing disappointment:
*UTTER <5-> i T S K UH M P L *E T L i R +OO i N D
RETURN
*UTTER <5-> W +o T UH SH *A M RETURN

Sarcasm:
*UTTER <5-> Y OR S *O K +I N D RETURN

(c) Compound Tone Groups

We have now dealt with all the simple Tone Groups; that is the ones with a single tonic syllable. There remain the compound Tone Groups; those with an extra tonic syllable. They are all composed of one of the variations of tone <1> or tone <5> followed immediately by the tonic of tone <3> (without its pretonic). The meanings are largely those of the component parts. In fact the next two examples produce identical outputs but because the paired tones occur so frequently together, they are usually treated a single compound tone with a double focus.

*UTTER <1> D *O N T B i UH <3> N Y *OO S UH N S
RETURN
*UTTER <1,3> D *O N T B i UH N Y *OO S UH N S
RETURN

The first focus carries the principle new information, the second carries subsidiary information, as is characteristic of tone <3>. Here are some more examples:

*UTTER <1,3> N *O TH *a N K Y OO RETURN
*UTTER <1,3> I L D *oo i T o N F R *I D A RETURN

*UTTER <1+,3> H *EA K u M Z DH UH S *u N RETURN
*UTTER <1+,3> +I K UH N S +E DH UH S *E F R UH M
+u P H *EA RETURN
*UTTER <1-,3> i T S +OR L *O V UH N *AW RETURN
*UTTER <u1,3> i T S DH i +O N L i W +A T UH M *A K
UH N i M P R *e SH UH N RETURN
*UTTER <5,3> DH a T P R *O G R a M Z F oo L UH V M i
S T *A K S RETURN
*UTTER <5,3> DH *a T W o Z N T UH G oo D I D
*EA RETURN
*UTTER <5-,3> i T W +o Z UH N +I S P R *e Z UH N T
DH *O RETURN

5.3 Notes on some special sequences of Tone Groups

(a) Long Sentences

A sentence can be broken up into Tone Groups in many different ways according to the needs of the speaker. Very long sentences are usually divided into several Tone groups. In statements, WH-questions and commands the sequence is usually <4>, <4>, <4>... <1> or <3>, <3>, <3>... <1>.

*UTTER <4> I W *u N D UH R i F H i <4> R *EA L I Z i
Z DH a T <4> DH UH B *e S T W +A <4> T oo R *I T UH
B +oo K i Z T oo <1> Y +OO Z W ER D W I Z P L *u S
RETURN

In yes/no questions the Tone Group sequence is most commonly <2>, <2>, <2>... <2>.

*UTTER <2> H +a V Y oo G *o T <2> DH UH F +A N T i
S T I D *EA <2> W +o T DH i S M UH SH *E N <2> K UH
N R *EA L i D +OO RETURN

(b) Added Information

Very often after using a noun we add another phrase to give extra information. These additions have a Tone Group of their own, with the same tone as the noun itself

*UTTER <1> M I F +AR DH UH N Y +OO L +OY D J *OR J
<1> ↑ DH UH P R +I M M *i N i S T UH RETURN

```
*UTTER <2> D +oo Y oo N +o P R i N S F *i L i P <2>
↑ DH UH J +oo K UH V *e D i N B R UH RETURN
```

Note that the introduction of such additions, if they come in the middle, may split up the original sentence resulting in a Tone Group sequence similar to the long sentence examples above:
```
*UTTER <1> M i S T UH J +O N Z i Z UH B *oo L
i RETURN
```

But:
```
*UTTER <4> M i S T UH J *O N Z <4> ↑ M I B *o S <1>
↑ i Z UH B *oo L i RETURN
```

Similarly:
```
*UTTER <4> T *AR Z UH N <4> ↑ DH i *A P M +a N <1>
↑ L +i V Z i N D +AR K i S T *a F R i K UH RETURN
```

In this last example "the apeman" is additional information, if, on the other hand, it had been used to distinguish "Tarzan the apeman" from "Tarzan the physicist" or if it were considered to be part of his title (as in "Edward the Confessor") then a single Tone Group is appropriate. Compare the following:

```
*UTTER <1> CH +AR L Z DH UH F +ER S T W o Z B i H
*e D i D RETURN
*UTTER <4> CH *AR L Z <4> ↑ DH UH F *ER S T <1> ↑
W o Z D i S K W *o L i F +I D RETURN
```
(and Fred, who came second, won the cup)

or this pair:
```
*UTTER <1> DH UH M +a N H OO K +u M Z F R UH M L +u
N D UH N W i L B i L *A T RETURN
```
(the other one will be on time)

```
*UTTER <4> DH UH M *a N <4> ↑ H oo K +u M Z F R UH
M L *u N D UH N <1> ↑ W i L B i L *A T RETURN
```
(who incidentally comes from London)

(c) Questions Containing "or"
When we ask questions with the word "or" there are two possible meanings distinguished only by intonation. The following examples will make this clear:

```
*UTTER <2> W +oo D Y oo L +I K K *o F i <1> ↑ OR T
*E RETURN
```
(which of the two do you prefer?)

```
*UTTER <2> W +oo D Y oo L +I K K *o F i <2> OR T
*E RETURN
```
(or are you still avoiding hot drinks?)

Notice that in the first example the sequence <2>, <1> is used to present alternatives; it expects the answer "coffee" or "tea". In the second example the sequence <2>, <2> anticipates the simple answer "yes" or "no" to the whole question. The same principle applies to cases with more than two items.

```
*UTTER <2> H a V Y oo G o T Y *e L O <2> ↑ G R *E N ↑
OR <1> B L *OO RETURN
```
(which of the three?)
```
*UTTER <2> H a V Y oo G o T Y *e L O <2> G R *E N OR
<2> B L *OO RETURN
```
(or did you get another colour altogether?)

(d) Lists
Related to the problem of multiple questions is that of lists. If a statement refers to two or more items as a list, then you should use either tone <4> alone or the sequence <3>, <3>, <3>... <1>. For example:

```
*UTTER <4> Y oo K UH N H a V +a P UH L ↑ +o R i N J ↑
OR S T R *OR B UH R i RETURN
```
(but not blackcurrant)

```
*UTTER <3> Y oo K UH N H a V +a P UH L <3> ↑ +o R i
N J ↑ OR <1> S T R *OR B UH R i RETURN
```
(which do you want?)

```
*UTTER <3> F +I V P +AW N Z UH V P UH T *A T O Z
<3> ↑ TH R +E *u N Y UH N Z ↑ UH N D <1> UH N +I S B
+i G K *a B i J RETURN
```

(e) "Too" and "Either"
When these words occur at the end of a sentence they **always**

have a Tone Group to themselves, usually tone <1> The focus of the previous Tone Group is the word that "too" or "either" is referring back to.

*UTTER <1> Y oo K UH N F R +I F *i SH <1> T *OO RETURN
(as well as eggs)

*UTTER <1> Y oo K UH N F R *I F +i SH <1> T *OO RETURN
(as well as poaching it)

*UTTER <1+> Y oo K A R N T +E T P o L i S T *I R E N <1> *I DH UH RETURN
*UTTER <1+> Y oo K A R N T *E T P o L i S T +I R E N <1> *I DH UH RETURN

(f) Tag Questions

This is the technical term for short comments such as "isn't it?" or "didn't they?", often tagged onto statements to emphasise that in the context concerned they are being used as questions. The statement and its tag may each be positive or negative. Since, additionally, each may carry a variety of tones, a great number of variations is possible. Some combinations are more common than others and have specific uses:

The tone sequence <1>, <2> used with statement and tag both positive or both negative is the neutral type:
*UTTER <1> Y oo V F *i N i SH D <2> H *a V N T Y OO RETURN
*UTTER <1> Y oo H +a V N T F *i N i SH D <2> H *a V Y OO RETURN
*UTTER <1> DH A SH *oo D <2> SH *oo D N T DH A RETURN
*UTTER <1> DH A SH *oo D N T <2> SH *oo D DH A RETURN

To express certainty or to demand an admission you should use tone <1> for both statement and tag and one of these should be positive, the other negative:
*UTTER <1> Y oo V F *i N i SH D <1> H *a V N T Y OO RETURN

54

*UTTER <1> Y oo H +a V N T F *i N i SH D <1> H *a V Y OO RETURN
*UTTER <1> DH A SH *oo D <2> SH *oo D N T DH A RETURN
*UTTER <1> DH A SH *oo D N T <1> SH *oo D DH A RETURN

The tone sequence <1>, <2> with statement and tag both positive or both negative expresses forcefulness, accusation or criticism:
*UTTER <1> Y oo V F *i N i SH D <2> H *a V Y OO RETURN
*UTTER <1> Y oo H +a V N T F *i N i SH D <2> H *a V N T Y OO RETURN
*UTTER <1> DH A SH *oo D <2> SH *oo D DH A RETURN
*UTTER <1> DH A SH *oo D N T <2> SH *oo D N T DH A RETURN

Try some of these with tone <1+> or <5> in place of tone <1> to indicate greater speaker involvement.

Examples such as the tag questions demonstrate the enormous richness of the English intonation system, and of course we can only begin to scratch its surface in this manual. The SPEECH ROM opens the door on this much neglected aspect of speech synthesis; it gives greater naturalness to even simple statements, and for users wishing to experiment further there are few limits.

5.4 Summary of Tone Group Functions

This section concludes by looking at the relationship between Tone Groups and their functions the other way round; that is with a summary of the Tone Groups appropriate for each of the main types of utterance.

(a) Statements

tone <1>   neutral statement

55

tone <1+>   emphasising
           expressing surprise
           showing excitement
           agreeing
           expressing irritation or intolerance

tone <1->   expressing lack of surprise
           implying resignation
           showing boredom

tone <u1>   stating forcefully
           showing querulousness

tone <3>   giving incomplete information
           acceding to a request
           confirming expectation
           giving reassurance

tone <3->   expressing uncertainty
           implying lack of concern
           giving grudging agreement
           cautioning

tone <4>   implying contrast
           expressing reservation
           offering personal opinion

tone <4->   expressing strong reservation
           justifying ("at least")
           bemoaning ("if only")
           giving grudging agreement
tone <5>   asserting
           expressing commitment

tone <5->   showing awe
           expressing disappointment
           being sarcastic

(b) Questions

tone <1>   neutral WH-question
           forceful yes/no question
           impatient yes/no question

56

tone <2>   neutral yes/no question
           tentative WH-question
           deferential WH-question

tone <-2>   question showing surprise
           question implying disapproval
           concerned question

tone <2v>   question specifying focus of query

(c) Statements used as questions

tone <1>   question expressing deduction

tone <2>   question seeking confirmation

(d) Commands

tone <1>   neutral positive command
           forceful negative command

tone <3>   polite "request"

tone <4>   command implying compromise or concession

tone <1,3> pleading
           persuading or dissuading

(e) Responses

tone <1>   neutral response

tone <2>   contradiction
           denial
           frustration of listener's expectation

tone <3>   confirmation
           reassurance

tone <5>   surprised response

57

APPENDIX I

SPEECH EDITOR

If you have Wordwise-Plus in your machine you can use the following short routine to develop speech very quickly. First enable the SPEECH ROM in the usual way, then select Wordwise-Plus. Next load the routine into an empty Segment and one or more complete *UTTER or *SING commands into the Text area. Position the cursor *anywhere* within a command, hold down SHIFT and, at the same time, press the appropriate red function key. The SPEECH ROM and Wordwise-Plus will co-operate to generate the speech output. The cursor will not be moved. You can use the normal editing functions of Wordwise-Plus to alter the commands and then hear the results immediately. Any incorrect commands will generate standard error messages at the top of the screen. Here is a listing of the Wordwise Plus program.

```
SELECT TEXT
N%=0
.LOOP
IF GCT$=CHR$(13) THEN GOTO SAY
N%=N%+1
CURSOR LEFT 2
GOTO LOOP
.SAY
OSCLI GLT$
REPEAT
CURSOR LEFT 2
UNTIL GCT$=CHR$(13)
CURSOR LEFT
CURSOR RIGHT N%
DISPLAY
```

APPENDIX II

DEMONSTRATION PROGRAM

We have included one program of significant length to demonstrate the flexibility and ease of use of the SPEECH ROM in action. The program is similar to the well known commercial spelling teacher but with fairly minimal facilities. It could profitably be expanded in several ways; to keep a record of the score for instance, or to focus on words causing the child particular difficulty. Verbal instructions on how to use the program could also be included. These additions have been omitted here for the sake of brevity. Note that during the running of the program the DELETE key works normally and the COPY key can be used to repeat the word being attempted.

The list of words under test can be increased or changed by altering the DATA statements from line 1130 to line 1340. Note that in each DATA statement the conventional spelling comes first, followed by the SPEECH ROM phonetic version. The value of the constant MAX at line 50 should be set to twice the number of words included.

The main point of the program is to show how text strings and string variables can be combined within OSCLI statements (available in BASIC-II only) to produce speech very flexibly. Take for example line 600, which embeds the variable containing the phonemes of the current word within an otherwise constant command. The procedure PROCLET, starting at line 780, takes this process one stage further. It is used to pronounce individual letters of the alphabet with a Tone Code specified by the parameter n$. The complete *UTTER command is put together in line 830. Notice that when the program itself gives the correct spelling of a word, the last letter is pronounced with tone <1+> in contrast to all the other letters which have tone <3>. This is made possible by calls to PROCLET with different values for n$. Speech produced using these techniques can be made to sound much more appropriate to its context. Here is the BASIC program listing:

```
10 ONERROR GOTO 1640
20 VDU23,1,0;0;0;0;
30 *FX11
40 *FX4,1
50 MAX=44: O$=""
60 PROCTITLE
70 *U.<3> P L E Z P R +e S DH UH S P *A S B AR <1> W
   e N Y oo AR R +e D i T oo S T *AR T
80 REPEAT UNTIL GET=&20
90 REPEAT
100 PROCSETUP
110 PROCASK
120 PROCRESP
130 *U.<3> # P R +e S DH UH S P *A S B AR <1> T oo
    G +O UH G *e N
140 REPEAT UNTIL GET=&20
150 UNTIL FALSE
160
170 DEFPROCASK
180 REPEAT
190 RESTORE
200 FOR X=1 TO RND(MAX) STEP 2
210 READ W$:READ P$
220 NEXTX
230 UNTIL W$<>O$
240 O$=W$
250 ON RND(3) GOTO 260,280,300
260 OSCLI"*U.<3> T R +I T oo S P *e L↑ <1> "+P$
270 ENDPROC
280 OSCLI"*U.<3> S +E i F Y oo K a N S P *e L↑
    <1> "+P$
290 ENDPROC
300 OSCLI"*U.<3> W *o T i Z DH UH S P +e L i NG UH
    V DH UH W +ER D ↑ <1> "+P$
310 ENDPROC
320
330 DEFPROCRESP
340 R=0
350 REPEAT R=R+1
360 PROCANS
370 IF M$<>W$ AND R<>3 PROCRETRY
380 UNTIL M$=W$ OR R=3
390 IF R=3 AND M$<>W$ PROCSPEL:ENDPROC
400 ON RND(3) GOTO 410,430,450
410 OSCLI"U.<1> DH +a T i Z K UH R *e K T"
420 ENDPROC
430 OSCLI"U.<1+> W +e L D *u N"
440 ENDPROC
450 IF M$=W$ OSCLI"U.<1> Y OO AR R *I T"
460 ENDPROC
470
480 DEFPROCANS
490 M$="": N=3
500 REPEAT C=GET
510 IF C=135 OSCLI"U.<3> DH UH W +ER D *i Z @
    <1+> "+P$
520 IF C=127 PROCDEL
530 IF C>64 AND C<91 THEN PROCPRIN(CHR$(C)):
    PROCLET(CHR$(C),"1"): M$=M$+CHR$(C)
540 UNTIL C=13
550 PRINT CHR$(13)
560 ENDPROC
570
580 DEFPROCSPEL
590 R=4: N=3
600 OSCLI"*U.<1> DH a T i Z +i NG K o R *e K T #
    <1,3> DH UH K o R *e K T S P +e L i NG UH V DH
    UH W E R D ↑ "+P$+" ↑ +i Z #"
610 REPEAT X$=LEFT$(W$,1): W$=MID$(W$,2)
620 PROCLET(X$,"3"): PROCPRIN(X$)
630 OSCLI"*U.<1> # #"
640 UNTIL LEN(W$)=1
650 PROCLET(W$,"1+"): PROCPRIN(W$)
660 *U.<1> @ @
670 ENDPROC
680
690 DEFPROCRETRY
700 ON RND(3) GOTO 710,730,750
710 OSCLI"U.<1> DH a T i Z +i NG K o R *e K T #
    <1+> T R +I UH G *e N"
720 ENDPROC
```

```
730 OSCLI"U.<1> DH a T i Z N +o T K W I T R *I T #
    <1+> H +a V UH N *u DH UH G +0"
740 ENDPROC
750 OSCLI"U.<1> Y OO H a V M +A D UH M i S T *A K #
    <3> T R +I W u N S M *OR"
760 ENDPROC
770
780 DEFPROCLET(c$,n$)
790 *SYNC
800 RESTORE 1360
810 REPEAT READ A$: READ B$
820 UNTIL A$=c$
830 OSCLI"*U.<"+n$+"> "+B$
840 ENDPROC
850
860 DEFPROCPRIN(L$)
870 PRINTTAB(N,R*4+2) L$
880 PRINTTAB(N,R*4+3) L$
890 N=N+1
900 ENDPROC
910
920 DEFPROCDEL
930 IF N=3 ENDPROC
940 N=N-1
950 PRINTTAB(N,R*4+2) " "
960 PRINTTAB(N,R*4+3) " "
970 M$=LEFT$(M$,LEN(M$)-1)
980 ENDPROC
990
1000 ZDEFPROCSETUP
1010 PROCTITLE
1020 FOR ROW=6 TO 18 STEP 4
1030 X=RND(6): PRINTTAB(0,ROW)
     CHR$(141);CHR$(128+X):
     PRINTTAB(0,ROW+1)
     CHR$(141);CHR$(128+X)
1040 NEXT ROW
1050 ENDPROC
1060
1070 DEFPROCTITLE
1080 CLS
1090 PRINTTAB(10,2) CHR$(141);"SPELLING
     TEST"
1100 PRINTTAB(10,3) CHR$(141);"SPELLING
     TEST"
1110 ENDPROC
1120
1130 DATA RUN,R *u N
1140 DATA CAT,K *a T
1150 DATA TOP,T *o P
1160 DATA GUN,G *u N
1170 DATA LET,L *e T
1180 DATA FAT,F *a T
1190 DATA STEW,S T Y *OO
1200 DATA THAN,DH *a N
1210 DATA TRY,T R *I
1220 DATA SHOOT,SH *OO T
1230 DATA FATHER,F *AR DH UH
1240 DATA WRONG,R *o NG
1250 DATA FAST,F *AR S T
1260 DATA LOVE,L *u V
1270 DATA TURKEY,T *ER K +i
1280 DATA KNICKERS,N *i K ER Z
1290 DATA SPLASH,S P L *a SH
1300 DATA KNIFE,N *I F
1310 DATA LAUGH,L *AR F
1320 DATA COUGH,K *o F
1330 DATA ROUGH,R *u F
1340 DATA
     SUPERCALIFRAGILISTICEXPEALIDOCIOUS
     ,S +OO P UH K +a L i F R a J i L *i S T i K <3>
     *e K S P i a L i <1> D *O SH UH S
1350
1360 DATA A,*A
1370 DATA B,B *E
1380 DATA C,S *E
1390 DATA D,D *E
1400 DATA E,*E
1410 DATA F,*e F
1420 DATA G,J *E
1430 DATA H,*A CH
1440 DATA I,*I
```

```
1450 DATA J,J *A
1460 DATA K,K *A
1470 DATA L,*e L
1480 DATA M,*e M
1490 DATA N,*e N
1500 DATA O,*O
1510 DATA P,P *E
1520 DATA Q,K Y *OO
1530 DATA R,*AR
1540 DATA S,*e S
1550 DATA T,T *E
1560 DATA U,Y *OO
1570 DATA V,V *E
1580 DATA W,D *u B UH L Y OO
1590 DATA X,*e K S
1600 DATA Y,W *I
1610 DATA Z,Z *e D
1620 DATA XX,XX
1630
1640 VDU23,1,1;0;0;0;
1650 *FX11,32
1660 *FX4,0
1670 END
```

(Note that the very long word in the **DATA** statement at line **1340** has an internal tonal structure. Because of this its intonation is not handled quite correctly when the program gives the right spelling. A more complicated routine, not really justified in this abbreviated program, would be needed to improve it further.)