Document #
Title: Encoding Pitman Shorthand scripts into Unicode Character Set
Source: Pitman scripts and its adaptation over years
Author: Ramachandran
Status: Individual contribution for consideration by UTC
Replaces: [L2/14-254]
Date: 2015-03-28

Historical overview of Pitman Shorthand and its adaptations

Pitman Shorthand scripts are widely used in English speaking countries. It is called 'Phonography' as it is a system of writing scripts based on pronunciation of English words and not using spelling of the words. Originally developed by the Englishman *Sir Isaac Pitman,* it was published in 1837. Since then, there have been several revisions to the original version for easy adaption.

There are at-least three "dialects" of Pitman's shorthand: the original Pitman's, Pitman's New Era, and Pitman's 2000. The later versions dropped certain symbols and introduced other simplifications to earlier versions. For example, strokes "rer" (heavy curved downstroke) and "kway," (hooked horizontal straight stroke) are present in Pitman's New Era, but not in Pitman's 2000. The characters that are proposed for encoding (in this proposal) are the consonants, vowels and dipthongs which are *prima facie* the base characters for Pitman Shorthand irrespective of any versions. Pitman's New Era is a fairly old version as compared to Pitman 2000 version. There is subtle difference between these two versions not with respect to base characters but some ligatures. Pitman Scripts described below is based on Pitman's New Era Shorthand which has been adapted to 30 languages all over.

Unsupported orthographies This proposal encodes Pitman Scripts used for English language. There is a roadmap to encode Pitman Scripts used for other supported languages in future, subject to availability of relevant materials or documentations which is being pursued.

Typology

Pitman Shorthand is a system of representing an English word or set of related English phrases using Consonants and Vowels (cf. Duployan Shorthand). As mentioned earlier, Pitman Shorthand (is based on simple geometric symbols viz., lines and curves) is written phonetically and not using spelling of each letters. Consonants are represented using strokes (Lines and Curves) can be easily differentiated either visually or according to their functionality and Vowels are represented using (Dot and Dash) although there are special vowels or combination of vowels which are described in the subsequent sections. Consonants and Vowels are written following a set of rules and their orientation does not change generally, except for in some cases. Pitman shorthand scripts are written Left to Right proceeding down the page. Scripts belonging to more than one word can be joined together to achieve fast writing of subsequent set of group words (e.g., the phrase, "As a matter of fact" can be combined together and written as a single script even though each word is a separate script).

This proposal for encoding Pitman Shorthand follows a similar framework and methodology as adopted in the "Proposal to include Duployan script and Shorthand Format Controls in UCS" as both the systems use almost the same writing pattern, although the way Pitman scripts rendered are different from that adopted by Duployan shorthand method.

Script Structure

In Pitman Shorthand, there is specific distinction between each Consonants in which some are written with light strokes (e.g., P, T, F, ITH, Kay, M, R) and some with heavy strokes (e.g., B, D, V, THEE, GY, ING) Vowels are represented by Light/Heavy Dot/Dash. There are 12 vowels in total (3 Light dots, 3 Heavy dots, 3 Light dashes and 3 Heavy dashes). There are three positions in which the vowels can be placed for a consonant to form meaningful phonemes.

In addition to the above vowels, there are certain special vowels called Diphthongs which are used separately or in combination with the vowels. Also, there are separate syllables called Diphones and Triphones. Apart from the original consonants, all these special vowels have separate script in Pitman Shorthand and they are also part of the allocation table.

Pitman Consonant scripts are generally written at the size of $1/6^{th}$ an inch. For easier representation of words, sometimes the size of the stroke is halved or doubled to add additional Consonant phonemes

Also, there are different variants to Consonant strokes (called dual consonants) which are added to the consonants either at the start or middle or at the end. Those variants are small and large hooks, small and large circles, ST and STR loops, Shun Hooks. These variant characters are not to be written individually but will be forming part of the main Consonant stroke to provide additional phonetic characters in a more convenient way, but following the rules in which they should be used. These variants also have separate character set and will be part of the allocation table.

Pitman Shorthand scripts can be joined together cursively for achieving the speed in writing but maintaining legibility. Wherever required, scripts belonging to more than one word can be joined together in a more flexible way to attain greater speed.

Character names

Character name of Pitman Shorthand script (that is proposed for encoding) is detailed as under.

SL #	Pitman Script[$$]	Character Name
1	\	PITMAN LETTER PEE
2	\	PITMAN LETTER BEE
3	|	PITMAN LETTER TEE
4	|	PITMAN LETTER DEE
5	/	PITMAN LETTER CHAY
6	/	PITMAN LETTER JAY
7	—	PITMAN LETTER KAY
8	—	PITMAN LETTER GAY
9	/	PITMAN LETTER RAY
10	⌣	PITMAN LETTER WAY
11	⌐	PITMAN LETTER YAY
12	⌐	PITMAN LETTER UP HAY
13	⌒	PITMAN LETTER EM
14	⌣	PITMAN LETTER EN
15	⌣	PITMAN LETTER ING
16	(	PITMAN LETTER EFF

17		PITMAN LETTER VEE
18		PITMAN LETTER ITH
19		PITMAN LETTER THEE
20		PITMAN LETTER ESS
21		PITMAN LETTER ZEE
22		PITMAN LETTER ISH
23		PITMAN LETTER ZHEE
24		PITMAN LETTER ELL
25		PITMAN LETTER AR
26		PITMAN VOWEL SIGN FIRST POSITION LIGHT DOT
27		PITMAN VOWEL SIGN SECOND POSITION LIGHT DOT
28		PITMAN VOWEL SIGN THIRD POSITION LIGHT DOT
29		PITMAN VOWEL SIGN FIRST POSITION LIGHT DASH
30		PITMAN VOWEL SIGN SECOND POSITION LIGHT DASH
31		PITMAN VOWEL SIGN THIRD POSITION LIGHT DASH
32		PITMAN VOWEL SIGN FIRST POSITION HEAVY DOT
33		PITMAN VOWEL SIGN SECOND POSITION HEAVY DOT
34		PITMAN VOWEL SIGN THIRD POSITION HEAVY DOT
35		PITMAN VOWEL SIGN FIRST POSITION HEAVY DASH

36	ー	PITMAN VOWEL SIGN SECOND POSITION HEAVY DASH
37	ー	PITMAN VOWEL SIGN THIRD POSITION HEAVY DASH
38	v	PITMAN LETTER I UP
39	ʌ	PITMAN LETTER OW DOWN
40	ʔ	PITMAN LETTER OI UP
41	ᴖ	PITMAN LETTER EW DOWN
42	?	PITMAN PUNCTUATION QUESTION MARK
43	ǀ	PITMAN PUNCTUATION EXCLAMATION MARK

Allocation

The current allocation of code blocks (**U+1BCB0…..U+1BCFF**) may suffice the accommodation of characters for other languages going by the assumption that there could be only a meager addition of characters to what is proposed now.

	1BCB	1BCC	1BCD	1BCE	1BCF
0	1BCB0	1BCC0	1BCD0		
1	1BCB1	1BCC1	1BCD1		
2	1BCB2	1BCC2	1BCD2		
3	1BCB3	1BCC3	1BCD3		
4	1BCB4	1BCC4	1BCD4		
5	1BCB5	1BCC5	1BCD5		
6	1BCB6	1BCC6	1BCD6		
7	1BCB7	1BCC7	1BCD7		
8	1BCB8	1BCC8	1BCD8		
9	1BCB9	1BCC9	1BCD9		

	1BCB	1BCC	1BCD	1BCE	1BCF
A	1BCBA	1BCCA	1BCDA		
B	1BCBB	1BCCB	1BCDB		
C		1BCCC	1BCDC		
D			1BCDD		
E			1BCDE		
F			1BCDF		

Supplemental Punctuation

	2E0	2E1	2E2	2E3	2E4
0					
1					
2					
3					? 2E43
4					| 2E44

Representations using proposed code points

Representation 1:

   1.1. Running script: This is the running text this is represented using the proposed code points. The running script is as below;

"The best time to do something worthwhile is between yesterday and tomorrow – unknown"

   1.2. Running script representation using standard Pitman Script

   1.3. In the below section, we see the code point representation of the running script. In this section, we see the Pitman script that is rendered using by the proposed characters (meant for encoding)

Apart from script, its corresponding English equivalent words are provided for better understanding of the representation.

In addition to above, the proposed code point of each script is also depicted, in the highlighted area. [Each code point word pair is delimited with //]
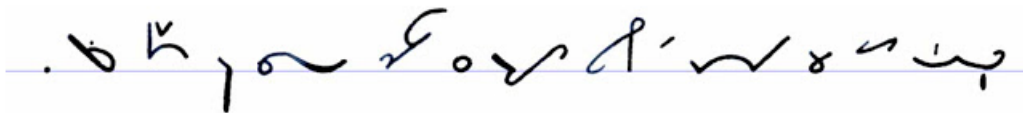

The	Best	time	to do	Something

<1BCC5><1BCD8> // <1BCB1><1BCD1><1BCC7><1BCB2> // <1BCB2><1BCDC><1BCC0> // <1BCB2><1BCD5><1BCB3>1BCD5> // <1BCC7><1BCD4><1BCC0><1BCC5><1BCC2>

Worthwhile	is	between	yesterday	and	tomorrow

<1BCB9><1BCDA><1BCB8><1BCC5><1BCB9><1BCDC><1BCCB> // <1BCD2><1BCC7> // <1BCB1><1BCD2><1BCB2><1BCB9><1BCD8><1BCC1> // <1BCBA><1BCD1><1BCC7><1BCB2><1BCD1><1BCCC><1BCB3><1BCD7> // <1BCD0><1BCC1><1BCB3> // <1BCB2><1BCD5><1BCC0><1BCD6><1BCB8><1BCDA>

Representation 2:

1.1. Running script: This is the running text this is represented using the proposed code points. The running scripts is as below;
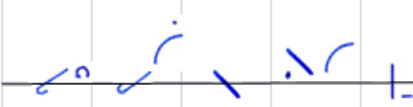
"You will be able to"

1.2. Running script representation using standard Pitman Script

1.3. In the below section, we see the code point representation of the running script. In this section, we see the Pitman script that is rendered using by the proposed characters (meant for encoding)

Apart from script, its corresponding English equivalent words are provided for better understanding of the representation.

In addition to above, the proposed code point of each script is also depicted, in the highlighted area. [Each code point word pair is delimited with //]

Character properties

1BCB0;PITMAN LETTER PEE;Lo;0;L;;;;;N;;;;;
1BCB1;PITMAN LETTER BEE;Lo;0;L;;;;;N;;;;;
1BCB2;PITMAN LETTER TEE;Lo;0;L;;;;;N;;;;;
1BCB3;PITMAN LETTER DEE;Lo;0;L;;;;;N;;;;;
1BCB4;PITMAN LETTER CHAY;Lo;0;L;;;;;N;;;;;
1BCB5;PITMAN LETTER JAY;Lo;0;L;;;;;N;;;;;
1BCB6;PITMAN LETTER KAY;Lo;0;L;;;;;N;;;;;
1BCB7;PITMAN LETTER GAY;Lo;0;L;;;;;N;;;;;
1BCB8;PITMAN LETTER RAY;Lo;0;L;;;;;N;;;;;
1BCB9;PITMAN LETTER WAY;Lo;0;L;;;;;N;;;;;
1BCBA;PITMAN LETTER YAY;Lo;0;L;;;;;N;;;;;
1BCBB;PITMAN LETTER HAY;Lo;0;L;;;;;N;;;;;
1BCC0;PITMAN LETTER EM;Lo;0;L;;;;;N;;;;;
1BCC1;PITMAN LETTER EN;Lo;0;L;;;;;N;;;;;
1BCC2;PITMAN LETTER ING;Lo;0;L;;;;;N;;;;;
1BCC3;PITMAN LETTER EFF;Lo;0;L;;;;;N;;;;;
1BCC4;PITMAN LETTER VEE;Lo;0;L;;;;;N;;;;;
1BCC5;PITMAN LETTER ITH;Lo;0;L;;;;;N;;;;;
1BCC6;PITMAN LETTER THEE;Lo;0;L;;;;;N;;;;;
1BCC7;PITMAN LETTER ESS;Lo;0;L;;;;;N;;;;;
1BCC8;PITMAN LETTER ZEE;Lo;0;L;;;;;N;;;;;
1BCC9;PITMAN LETTER ISH;Lo;0;L;;;;;N;;;;;
1BCCA;PITMAN LETTER ZHEE;Lo;0;L;;;;;N;;;;;
1BCCB;PITMAN LETTER ELL;Lo;0;L;;;;;N;;;;;
1BCCC;PITMAN LETTER AR;Lo;0;L;;;;;N;;;;;
1BCD0;PITMAN VOWEL SIGN FIRST POSITION LIGHT DOT;Lo;0;L;;;;;N;;;;;
1BCD1;PITMAN VOWEL SIGN SECOND POSITION LIGHT DOT;Lo;0;L;;;;;N;;;;;
1BCD2;PITMAN VOWEL SIGN THIRD POSITION LIGHT DOT;Lo;0;L;;;;;N;;;;;
1BCD3;PITMAN VOWEL SIGN FIRST POSITION LIGHT DASH;Lo;0;L;;;;;N;;;;;
1BCD4;PITMAN VOWEL SIGN SECOND POSITION LIGHT DASH;Lo;0;L;;;;;N;;;;;
1BCD5;PITMAN VOWEL SIGN THIRD POSITION LIGHT DASH;Lo;0;L;;;;;N;;;;;
1BCD6;PITMAN VOWEL SIGN FIRST POSITION HEAVY DOT;Lo;0;L;;;;;N;;;;;
1BCD7;PITMAN VOWEL SIGN SECOND POSITION HEAVY DOT;Lo;0;L;;;;;N;;;;;
1BCD8;PITMAN VOWEL SIGN THIRD POSITION HEAVY DOT;Lo;0;L;;;;;N;;;;;
1BCD9;PITMAN VOWEL SIGN FIRST POSITION HEAVY DASH;Lo;0;L;;;;;N;;;;;
1BCDA;PITMAN VOWEL SIGN SECOND POSITION HEAVY DASH;Lo;0;L;;;;;N;;;;;
1BCDB;PITMAN VOWEL SIGN THIRD POSITION HEAVY DASH;Lo;0;L;;;;;N;;;;;
1BCDC;PITMAN LETTER I UP;Lo;0;L;;;;;N;;;;;
1BCDD;PITMAN LETTER OW DOWN;Lo;0;L;;;;;N;;;;;
1BCDE;PITMAN LETTER OI UP;Lo;0;L;;;;;N;;;;;
1BCDF;PITMAN LETTER EW DOWN;Lo;0;L;;;;;N;;;;;
2E43;PITMAN PUNCTUATION QUESTION MARK;Lo;0;L;;;;;N;;;;;
2E44;PITMAN PUNCTUATION EXCLAMATION MARK;Lo;0;L;;;;;N;;;;;

Line breaking properties for punctuation

2E43;BA   # Po   PITMAN QUESTION MARK
2E44;BA   # Po   PITMAN EXCLAMATION MARK

Character ordering

Characters are generally ordered by straight strokes starting from slanting left and progressing clockwise towards horizontal strokes followed by slanting hook strokes are ordered in column 2 (in the allocation table). Thereafter it starts with curved horizontal strokes and then followed by curved strokes progressing from left to right direction which are ordered in column 3 (in the allocation table) Then follows the vowels, dipthongs and punctuations characters in column 4 (in the allocation table)

Each alternative strokes refers to once consonant and are differentiated by their heaviness (Refer to character names table) with exceptions for a few strokes which are light.

This allocation provides for all characters needed for Pitman English scripts. Other codes points in the block are left unallocated for addition of other languages characters to be provided for in future.

Stenographic period

Unicode has already encoded stenographic full stop U+2E3C STENOGRAPHIC FULL STOP which conveniently be used for rendering Pitman script as there are no other characters in Pitman stenography which resembles this character.

Punctuation marks

There are two punctuation marks (Exclamation and Question) proposed for encoding. They don't have any specific features and works like their counterpart in English. For the purpose of distinguishing them from English, there is a 'x' mark instead of a dot as in English.

Fonts

As Pitman scripts are rendered by straight lines and curves, there is not a need to attach a specific font to these scripts.

.

Pitman Rendering Model

Vowels in Pitman shorthand are indicated by small dot or dash along the consonant script. Mostly, the vowels are not joined to consonants but in certain cases can be joined for convenience purpose only, provided joining does not give them a different meaning.

Orientation of vowels:

All dash vowels in Pitman shorthand are written perpendicular to straight scripts with respect to their position as decided by the sound. Apart from dot and dash vowels there are special vowels called diphthongs but they take either 1st or 3rd position and there is no second position. There are 4 diphthongs in Pitman shorthand which are either joined or disjoined with consonant scripts. The diphthong EW, normally orient itself conveniently with the curved strokes. For example, when attached with words like few, mew, new, value etc.

Table: 1 → Orientation of Vowels

Compound vowels

Combining two consecutive vowels sound are referred to compounding vowels. In Pitman shorthand, there are quite a few words which require a compounding vowel when written as Pitman script. Compound vowels can be of two types i.e., combining two vowels or combining a diphthong and a vowel. There former kind of compounding is referred to as Diphone and the latter referred to a Triphone. Diphone is written in two directions (down-left or up-right) depending upon whether the first occurring vowel is a dot vowel or a dash vowel consecutively. Vowel in a Triphone is always as a dash combined with a diphthong but in the obverse direction of diphthong.

<p align="center">Table: 2 → Compounding of Vowels (Diphone and Triphone)</p>

Ligatures, Allograph, Variants and Affixes

The 'H' aspirate pronounced "HAY" is the allograph in Pitman Script. There are two forms of H in Pitman script. The different forms are used based on rules and according to the convenience of joining the stroke with other consonants.

Variants are alternative forms of scripts in Pitman shorthand. They are used in specific cases where convenience of joining strokes cursively, are required. Also they follow certain rules based on which it has to be decided whether the variant forms should be used or the full-form should be used.

Table: 3 → Ligatures, Variants, Allograph

Allograph (H Aspirate)

∽ + │ → ∽⌐		H(up) + D
＼ + ∽ + ⌣ → ＼⌐		B + H(up) + F
♀ + ⌐ → ⧵		H(down) + R

Variants

╱ + ＼ + ⌐ → ∧		RAY + B + R
＼ + │ + ⸦ → ⌐		B + T + L
│ + ╱ + ⌣ → ⌐		T + R + N
│ + ╱ + ⌣ → ⌐		T + R + F
─ + │ + ⸦ → ⌐		K + T + V
⸝ + ⌐ + ⌢ → ⌐		H(tick) + R + M
) + ⌢ + ⸦ → ⌐		ESS + M + L

Ligatures are forms which combines more than one script into a single gylph. In Pitman shorthand, there are many ligatures defined for the purpose of easy writing, combining with other stokes and also following rules under which these ligatures are to be used. Ligatures in Pitman are Discretionary ones.

Ligatures (Discretionary)

) + | + ⌣ → ℓ		**ESS + T + F**
⌢ +) + | + ⌐ → ◯		**M + ESS + T + R**
＼ +) +) → ◯		**P + ESS + ESS**
) + ✓ + ＼ → ℛ		**ESS + WAY + P**
) + ✓ → ℯ		**ESS + WAY**
✓ +) + ⌠ → ℓ		**WAY + ESS + L**

Affixes are used in Pitman shorthand for the purpose of representing W (WAY). They are joined and also disjoined.

Affixes:

✓ + — + ⌣ → ⌐		**WAY + K + NG**
| + ✓ + ⌣ + | → ⌐		**T + WAY + N + T**
╱ + ＼ + ✓ + — → ╱╲		**WAY + RAY + D + N**

References

Berly's blog @ http://www.long-live-pitmans-shorthand.org.uk/

http://pitmanshorthand.homestead.com/PitmanBooks.html and http://pitmanshorthand.homestead.com/

http://www.britannica.com/EBchecked/topic/541788/shorthand/53187/Modern-symbol-systems

http://www.jstor.org/stable/6553?seq=1#page_scan_tab_contents

Examples

Representation of Vowels and Dipthongs

light dots:

a (bat)　　　　e (bet)　　　　i (bit)

heavy dots:

ah (ball)　　　a (bait)　　　ee (beet)

light dashes:

aw (bog)　　　u (bug)　　　oo (book)

heavy dashes:

aw*(law)　　　o (boat)　　　oo (boot)

others:

I (eye)　　　oi (boy)

ow, bow　　　you, few/if you

Examples, that were extracted from the below website.

http://www.jstor.org/stable/6553?seq=3#page_scan_tab_contents

According to Isaac Pitman's analysis, there are forty sounds in the English language, twenty-four consonants, twelve simple vowels, and four diphthongs, or double vowels. Adopting (substantially) the Pitmanic symbols, we may represent these sounds as below.

CONSONANTS

\ = p as in pop
\ = b as in bob
| = t as in tat
| = d as in did
/ = ch as in church
/ = j as in judge
— = k as in kick
— = g as in gig
\ = f as in fife
\ = v as in vivid

VOWELS (SINGLE)

(The vertical line is not a part of the vowel symbol, but is used to represent any consonant stroke. A vowel symbol, as a heavy or light dot, stands for different sounds according to its position.)

|˙ = a as in pa
|. = a as in may
|. = e as in we
|▬ = a as in all

THE SHORTHAND ALPHABET   273

Below
http://www.jstor.org/stable/6553?seq=3#page_scan_tab_contents

Below is a classic example for the (**Representations using proposed code points**) provided above. The representation provided in (Page # 9 and # 10) is based on proposed characters which are only base consonants, vowels and dipthongs. The examples provided in the below screen also adhere to the representation model.
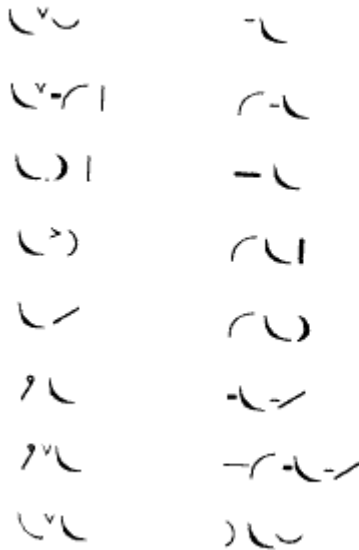
Fig. /. Three charts, reduced in size, used in the teaching of fonoline. The chart at the upper left shows the fonoline alphabet, omitting the symbol for the , sound of zh, which was not used in the first grade vocabulary. The words on the other charts are as shown below, and in the same order.

Words on Chart at Upper Right.

at	an
cat	can
fat	Dan
hat	ran
rat	fan
sat	man
that	than

Words on Chart at Left.

vine	of
violet	love
visit	give
voice	lived
very	lives
have	over
hive	clover
five	seven