

中国語 HV-Script フォーマット仕様書

第 1.0.0 版
2004 年 2 月 13 日

ヤマハ株式会社
半導体事業部 ソフトウェア開発部

[注意事項]

本書は HV-Script に関するドキュメントであり、再生システムの動作を保証するものではありません。

本書の内容は予告なく変更される場合があります。

本書の著作権は、ヤマハ株式会社に帰属しています。

本書の内容の転載・一部複製には、ヤマハ株式会社半導体事業部の許諾が必要です。



Copyright (C) 2004 YAMAHA CORPORATION

All rights reserved

目次

1	はじめに	3
2	基本構成	3
3	フォーマット	4
3.1	Header.....	4
3.2	Body.....	4
3.2.1	音声文字列.....	5
3.2.2	韻律記号	9
3.2.3	制御文字列.....	12
3.2.4	イベント	14
3.2.5	コメントアウト.....	15
4	HV-Script 例.....	16
5	文字・記号コード対応表	17
6	デフォルト音色マップ	23

更新履歴

版	日付	内容
1.0.0.0	2004 年 2 月 13 日	初版

1 はじめに

HV-Script は、韻律記号を含んだ合成文字列、発音する音の設定、再生アプリケーション等へのメッセージから成る音声合成を行うためのフォーマットであり、ユーザによる入力を容易にするためテキスト入力となっている。

2 基本構成

HV-Script は **Header** と **Body** から構成される。

Header は HV-Script 開始と言語及び **Body** で使われる文字コードを表す 4Byte の ASCII 文字で構成される。

Body は **Header** で指定される文字コードで書かれた文字列で、発声情報や再生情報で構成される。

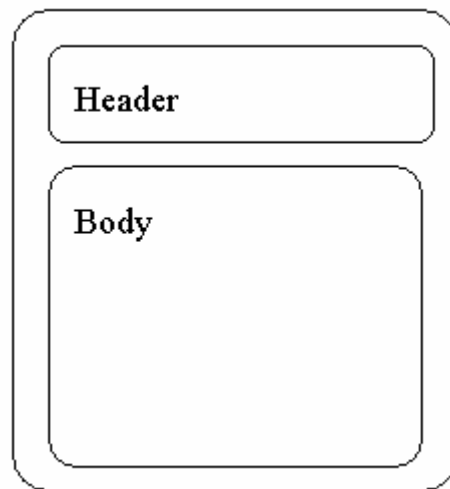


図 1 HV-Script 基本構成

3 フォーマット

3.1 Header

Header は HV-Script の先頭に記述される。4Byte のアルファベット大文字(ASCII)で表される。

HV#C

HV# (0x48,0x56,0x23) HV-Script 開始

C (0x43) 国・言語・文字コードを表す。この場合、中国語を意味する。

3.2 Body

Body は ASCII の文字コードで記述されており、以下で構成される。

- ・音声文字列 : 発声する文字を表す。
- ・韻律記号 : 音声文字列に対し、文節(*1)単位または音節(*2)単位で韻律(*3)を与える記号
- ・制御文字列 : 音量や音の高さ、話速などを設定する。
- ・イベント : 再生アプリケーション等へのメッセージ送信。
- ・コメントアウト : HV-Script 中にコメントを書き込むことができる。コメントは再生処理に影響しない。

コメントアウト内では、GB2312、Big5、ASCII の文字コードの使用が可能。

Body で使われるアルファベットは読み記号には小文字を、その他の記号には大文字を使う。

記号・文字のコード表を表 9～表 15 に示す。

- *1 文節 : 文章の先頭から「文節の区切り記号」までで区切られる文章の単位のこと。
(「文節の区切り記号」については後述を参照。)

(例) HV#CW5ni2hao3-. W3huanying2. >2huanying2.

文節 文節 文節

- *2 音節 : 1 つの「読み記号」が 1 音節に当たる。

- *3 韻律 : 音の強弱、長短、高低などによって作り出される言葉のリズム。

3.2.1 音声文字列

発声する文字及び無音、文節の区切りなどを示した文字列で、以下で構成される。

- ・読み記号
- ・長音
- ・無音
- ・文節の区切り記号

3.2.1.1 読み記号 [アルファベット列(PinYin)][数値[0-4](声調符号)]

発声する音節を PinYin で入力する。入力可能な音節の PinYin を表 9～表 12、その ASCII コードを表 13 に示す。「読み記号」は、音節を表現するアルファベット列と声調を表現する数値で構成される。アルファベットは小文字で入力すること。声調符号は数値で音節毎に必ず入力すること。声調符号は実際に発音する声調で使用する。 (入力例: 你好吗 : ni2hao3ma0)

声調符号のピッチ変化を図 2 に示す。声調を失う轻声は声調符号を 0 と記述し、直前の音節の声調によって図 3 のようなピッチになる。文節内で第 3 声の後に第 3 声以外の音節がある時は図 2 にある第 3 声の変化の前半部分のみの変化となり、後半は、最下ピッチで一定の半 3 声になる。声調符号によるピッチの値は、文節内の直前の音節までに高低アクセント記号でシフトされた高さの分だけピッチ変化の高さがシフトする。

また、高低アクセント記号が読み記号または長音に付加されたときは、その読み記号の発話には声調符号によるピッチの変化は行わない。

長音を用いて発音の長さを伸ばしたときは、その長さに応じたピッチの変化を行う。

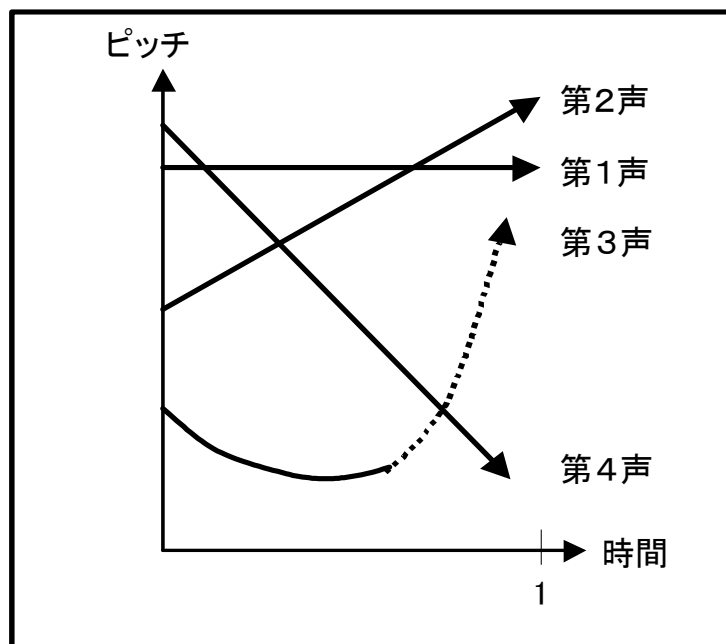


図 2 声調によるピッチ変化

横軸の"1"は文字ごとの発話長を表す。発話長は文字ごとに異なる。(以下同様)

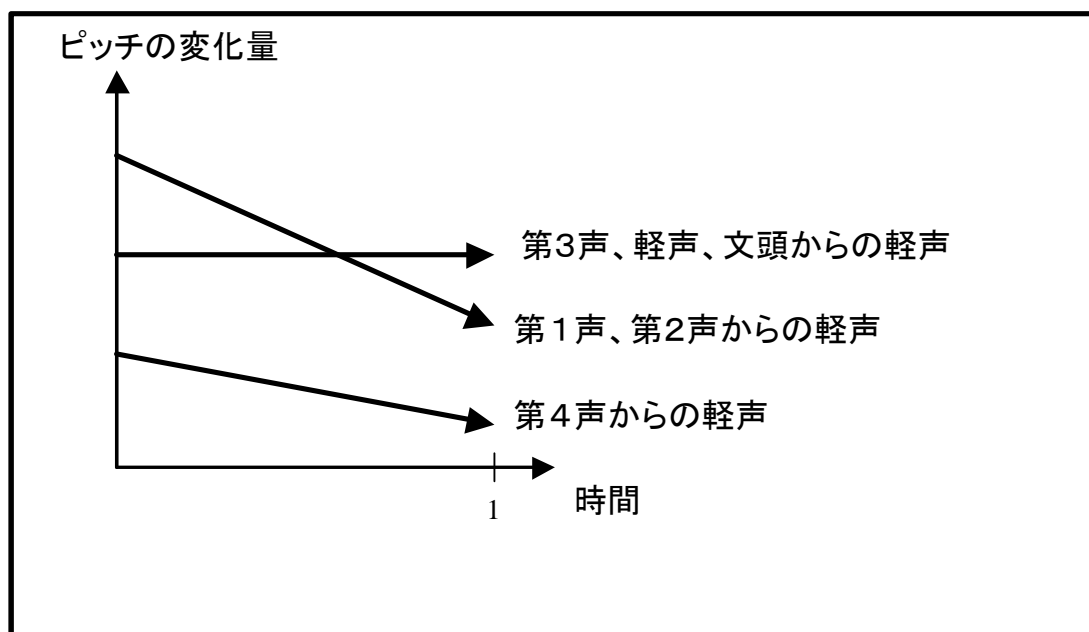


図 3 直前の音節の声調にともなう轻声のピッチ変化

3.2.1.2 長音 [" - "," ~ "]

直前の「読み記号」をのばして発声する。連続して入力した数だけ長くなる。

「-」はピッチが一定の長音であるのに対し、「~」はピッチの揺れをもった長音になる。

例 : 「ni2----hao3~~~」

3.2.1.3 無音 [" "]

一定の無音を挿入する。連続して入力した数だけ無音時間が長くなる。

3.2.1.4 文節の区切り記号 [",", ".", "?", "*"]

文節の終わりを示す。

", "では「無音」 1 つ、"."では 2 つ分の長さの「無音」が強制的に挿入される。

"?"は直前の「読み記号」または「長音」に対してピッチと音量を上げ、疑問の韻律を表現する。読み記号の声調符号によるピッチの変化は行わない。「?99」のように後ろに 1 から 99 の数値を付けることでピッチ及び音量変化量が大きくなる。また 2 つ分の長さの「無音」が強制的に挿入される。「????」のように"?"を連続して入力した場合、それぞれが文節の終わりを示しているとみなす。この場合、ピッチ及び音量変化量は"?1"と変わらないが、無音は"? "の数だけ増加する。

"*"は直前の「読み記号」または「長音」に対してピッチと音量を下げて韻律を表現する。読み記号の声調符号によるピッチの変化は行わない。「*99」のように後ろに 1 から 99 の数値を付けることでピッチ及び音量変化量が大きくなる。また 2 つ分の長さの「無音」が強制的に挿入される。「*****」のように"*"を連続して入力した場合、それぞれが文節の終わりを示しているとみなす。この場合、ピッチ及び音量変化量は"*1"と変わらないが、無音は"* "の数だけ増加する。

3.2.2 韻律記号

音声文字列に対し、文節単位または音節単位で韻律を与える。

3.2.2.1 アクセント記号 ["'","^","/","_","\$","<",">","&","="]

「読み記号」または「長音」の直前に記述し、ピッチを変化させる高低アクセント(表 1,図 4)、音量を変化させる強弱アクセント(表 2,図 5)で韻律の設定を行う。「文節の区切り記号」または表 3 に示す変化を戻す記号までは変化量を累積していく。

1つの「読み記号」または「長音」に対し、「高低アクセント記号」と「強弱アクセント記号」を同時に記述することを許すが、複数の異なる「高低アクセント記号」または「強弱アクセント記号」を記述した場合は後着優先となる。ただし、「/」「=」「&」の変化を戻す記号との併用は可能であり、「/」のように入力された場合は、変化量を0に戻してから変化を行う。

また、連続して同じ記号を入力すると変化量が大きくなり(図 6)、「99」のように後ろに1から99の数値をつけることにより連続入力と同じ意味を持つ。

「アクセント記号」の直後の「読み記号」または「長音」に対して"?"または"*"が有効な場合、「アクセント記号」は読み飛ばしされる。

表 1 高低アクセント記号

記号	内容
'	語頭でピッチを上げる
^	発音中ピッチを上げる
_	語頭でピッチを下げる
\$	発音中ピッチを下げる

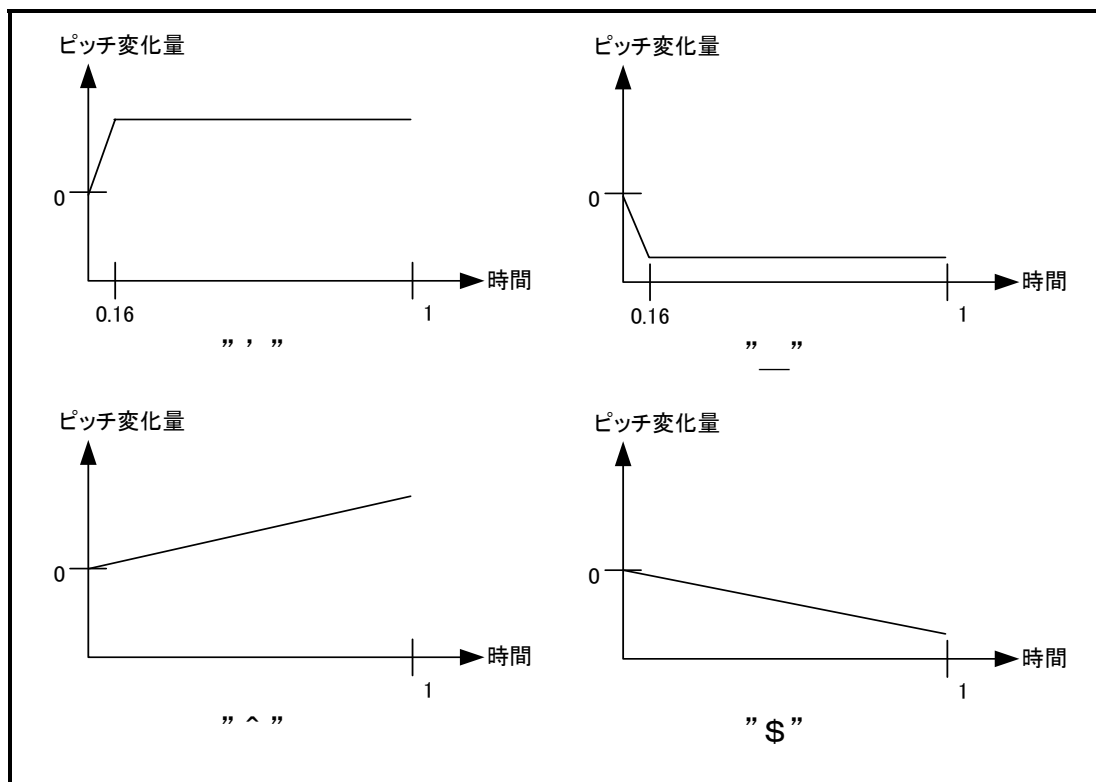


図 4 高低アクセント記号

表 2 強弱アクセント記号

記号	内容
<	発音中に音量を上げる
>	発音中に音量を下げる

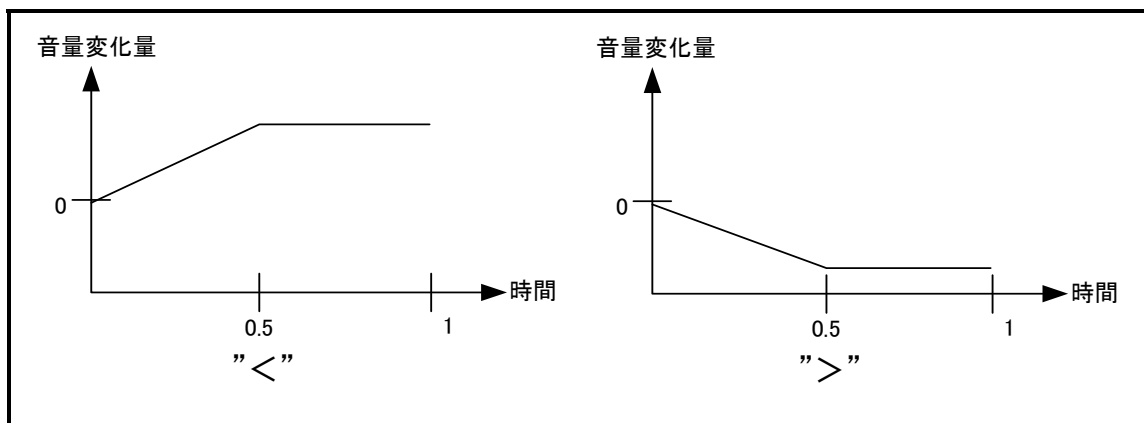


図 5 強弱アクセント記号

表 3 アクセントシフトクリア

記号	内容
/	アクセント記号で変化したピッチをクリアする。
=	アクセント記号で変化した音量をクリアする
&	アクセント記号で変化したピッチと音量をクリアする。

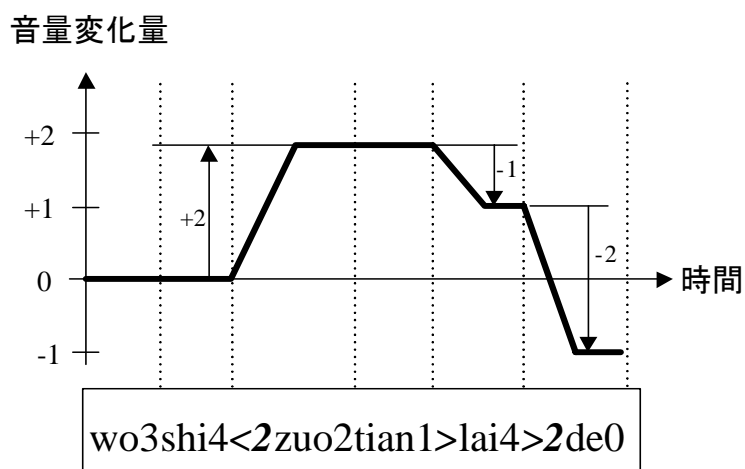


図 6 アクセント記号 (例)

3.2.2.2 文節の韻律 ["@", ":", ";", "!", "+"]

文節の先頭に記述し、文節区間に一様にかかる韻律(調子)を与える。「文節の韻律記号」と作用を表 4, 図 7 に示す。

設定した文節にのみ有効であり、記述しない場合は変化なし。

“+” は文節の各「読み記号」、「長音」に対してランダムに「高低アクセント記号」を設定するため、文節に書かれた「高低アクセント記号」は無視される。

表 4 文節の韻律記号

記号	内容
@	2 つ目の読み記号で高くなり最後で下がる
!	2 つ目の読み記号で低くなり最後で上がる
;	2 つ目の読み記号で高くなる
:	2 つ目の読み記号で低くなる
+	ランダムピッチ

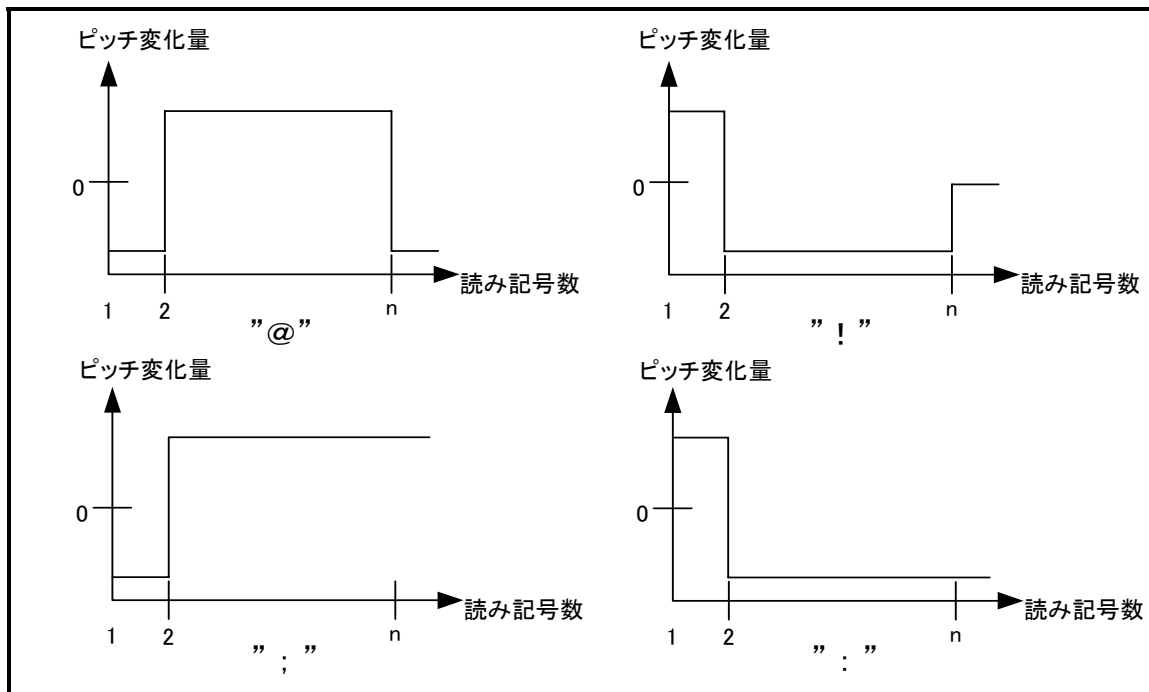


図 7 文節の韻律(読み記号数 n の場合)

3.2.3 制御文字列

以下の記号を用いて、音量やピッチなどの変化を与える。変化させた変化量は次に変更するまで同じ値が維持される。各記号のアルファベットは大文字であること。

なお、以下の設定値の定義で"[]"で囲まれている部分は必須入力、"0"で囲まれている部分は任意入力であることを示す。

3.2.3.1 音量

V[数値]

- 数値設定範囲 : 1-5（初期値：4）
- 説明 : 音量を変えるメッセージ。
各設定数値に対する音量変化量は表 5 のとおり。
入力位置の次の「読み記号」から、次に「音量」が指定されるか再生が終了するまで有効。
「読み記号」と「長音」、「長音」と「長音」の間の「音量」は、「長音」以降の次の「読み記号」から反映される。

表 5 音量変化パラメータ表

メッセージ	変化量 (dB)
V1	-24
V2	-18
V3	-12
V4	-6
V5	0

3.2.3.2 ピッチ

[音階]（オクターブ数値）

- 音階設定範囲 : C, C#, D, D#, E, F, F#, G, G#, A, A#, B
- オクターブ指定範囲 : 1-3（初期値：2）
- 説明 : 発音する声の高さを設定するメッセージ。
音階とオクターブ（表 6）の指定。オクターブの高さは、音色に合った高さに自動的に設定される。
オクターブ数値を省略した場合はそれまでの値で設定される。
ピッチの指定がない場合は、音色のデフォルト値で再生される。
入力位置の次の「読み記号」から、次に「ピッチ」が指定されるか再生が終了するまで有効。
「読み記号」と「長音」、「長音」と「長音」の間の「ピッチ」は、「長音」以降の次の「読み記号」から反映される。

表 6 オクターブ指定範囲変化量

オクターブ指定範囲	変化量
1	デフォルト-1 オクターブ
2	デフォルト
3	デフォルト+1 オクターブ

3.2.3.3 速度

S [数値]

-
- 数値指定範囲 : 0-99 (初期値 : 50)
- 説明 : 発話速度を変えるメッセージ。
 指定数値が 1 増えるごとに発話長が一定量短くなり、減るごとに長くなる。
 入力位置の次の「読み記号」、「無音」、「区切り記号」から、次に「速度」が
 指定されるか再生が終了するまで有効。
 以下の例のような、長音中に記述した速度は次の「読み記号」から有効となる。
 (例) HV#C S40ni2---S50--hao3ma0.

3.2.3.4 発話長統一

L [数値]

-
- 数値指定範囲 : 0-1 (初期値 : 0)
- 説明 : 記述された位置以降全ての発話長を同じにするか、それぞれ固有の長さにするかのスイッチ。スイッチの ON/OFF は表 7 の通り。
 入力位置の次の「読み記号」、「無音」から、次に「発話長統一」が指定されるか再生が終了するまで有効。
 以下の例のような、長音中に記述した発話長統一は次の「読み記号」から有効となる。
 (例) HV#C L1ni2---L0--hao3ma0.

表 7 発話長統一設定値

数値指定	読み記号の発話長
0	固有の長さ
1	同じ長さ

3.2.4 イベント

HV-Script の再生時に再生アプリケーション等にするメッセージ。

3.2.4.1 声質変更

K[数値]、X[数値]

数値設定範囲	: 0-15（初期値：K0）
説明	: 合成音声の声質を数値設定の声質に変える。 K0-K15 はデフォルト音色。X0-X15 は拡張音色であり、あらかじめ拡張音色が設定されていないといけない。 入力位置の次の「読み記号」、「長音」から、次に「声質変更」が指定されるか再生が終了するまで有効。 デフォルト音色の内容については、6 章「デフォルト音色マップ」を参照。

3.2.4.2 ユーザイベント

U[数値]

数値設定範囲	: 0-9
説明	: HV-Script 再生時に、ユーザ指定のイベントを実行させるためのメッセージ。

3.2.4.3 高低アクセント・文節の韻律変化度

W[数値]

数値設定範囲	: 1-5（初期値：3）
説明	: 「高低アクセント記号」、「文節の韻律記号」の韻律変化度を変えるメッセージ。 指定数値が大きいほど大きく変化する。各設定数値における変化量は表 8 のとおり。 入力位置の次の「読み記号」、から、次に値が指定されるか再生が終了するまで有効。

表 8 高低アクセント・文節の韻律変化パラメータ表

メッセージ	変化量(倍)
W1	1/6
W2	1/2
W3	1
W4	2
W5	5

3.2.5 コメントアウト

HV-Script 中にコメントを書き込むことを可能にする

[(文字列)]

説明	:	"[]" で囲まれた領域はコメントとして扱う。 入れ子や各命令の間に記述する(例: V[yin1liang2]5)ことは不可。 GB2312、Big5、ASCII の使用が可能。
----	---	---

4 HV-Script 例

HV-Script の入力例を以下に示す。

```
HV#CW5ni2hao3-.W3huan1ying2,>2huan1ying2,  
K1ni2hao3-.hao2jiu3S53bu2jian4le0.  
K0S50qing3zuo4.qing3he1<cha2.  
K1W5S40<2^3xie4&xie0.  
K0S50bie2ke4qi0.
```


5 文字・記号コード対応表

表 9 音節表(1/4)

声母 韻母	介音なし												
	a	o	e	-i[zh]	-i[z]	er	ai	ei	ao	ou	an	en	ang
b	ba*	bo					bai*	bei*	bao*		ban*	ben*	bang
p	pa*	po					pai*	pei*	pao*	pou	pan*	pen*	pang
m	ma*	mo	me				mai*	mei*	mao*	mou	man*	men*	mang
f	fa*	fo						fei*		fou	fan*	fen*	fang
d	da*		de*				dai*	dei*	dao*	dou*	dan*	den*	dang
t	ta*		te*				tai*		tao*	tou*	tan*		tang
n	na*		ne*				nai*	nei	nao*	nou	nan*	nen	nang
l	la*	lo	le*				lai*	lei	lao	lou	lan*		lang
g	ga*		ge*				gai*	gei	gao*	gou*	gan*	gen	gang
k	ka*		ke*				kai*	kei	kao*	kou*	kan*	ken	kang
h	ha*		he*				hai*	hei	hao*	hou*	han*	hen	hang
j													
q													
x													
zh	zha*		zhe*	zhi			zhai*	zhei	zhao*	zhou*	zhan*	zhen	zhang
ch	cha*		che*	chi*			chai*		chao*	chou*	chan*	chen*	chang
sh	sha*		she*	shi*			shai*	shei*	shao*	shou*	shan*	shen*	shang
r			re*	ri*					rao	rou*	ran	ren*	rang
z	za*		ze		zi*		zai*	zei*	zao*	zou	zan*	zen*	zang
c	ca*		ce		ci*		cai*		cao*	cou*	can*	cen*	cang
s	sa*		se		si*		sai*		sao	sou	san*	sen*	sang
ゼロ	a*	o*	e			er	ai*	ei	ao	ou	an*	en	ang

*付の音節は儿化入力可能

表 10 音節表(2/4)

声母 韻母	介音なし		介音 i									
	eng	ong	i	ia	ie	iao	iou -iu	ian	in	iang	ing	iong
b	beng		bi*		bie	biao*		bian*	bin*		bing	
p	peng		pi*		pie	piao*		pian*	pin*		ping	
m	meng		mi*		mie	miao*	miu	mian*	min*		ming	
f	feng*											
d	deng	dong	di	dia*	die*	diao	diu	dian*			ding	
t	teng	tong	ti		tie*	tiao*		tian*			ting	
n	neng	nong	ni*		nie	niao*	niu*	nian*	nin*	niang	ning	
l	leng	long	li*	lia*	lie	liao*	liu*	lian*	lin*	liang	ling	
g	geng	gong										
k	keng	kong*										
h	heng	hong										
j			ji*	jia*	jie	jiao*	jiu*	jian*	jin*	jiang	jing	jiong
q			qi*	qia*	qie	qiao*	qiu*	qian*	qin*	qiang	qing	qiong
x			xi*	xia*	xie	xiao*	xiu	xian*	xin*	xiang	xing	xiong
zh	zheng	zhong*										
ch	cheng	chong										
sh	sheng											
r	reng	rong										
z	zeng	zong										
c	ceng	cong										
s	seng	song										
ゼロ	eng		yi*	ya*	ye	yao	you	yan*	yin*	yang	ying*	yong*

**'付の音節は儿化入力可能

表 11 音節表(3/4)

声母 韻母	介音 u									介音 ü			
	u	ua	uo	uai	uei -ui	uan	uen -un	uang	ueng	ü	üe	üan	ün
b	bu												
p	pu												
m	mu*												
f	fu*												
d	du		duo		dui*	duan*	dun*						
t	tu		tuo*		tui*	tuan*	tun*						
n	nu		nuo			nuan				nv*	nve		
l	lu		luo			luan	lun*			lv*	lve		
g	gu	gua*	guo*	guai*	gui*	guan*	gun*	guang					
k	ku	kua*	kuo	kuai*	kui*	kuan*	kun*	kuang					
h	hu	hua*	huo*	huai*	hui*	huan*	hun*	huang					
j										ju*	jue	juan*	jun*
q										qu*	que	quan*	qun*
x										xu*	xue	xuan*	xun*
zh	zhu	zhua*	zhuo	zhuai*	zhui*	zhuan*	zhun*	zhuang					
ch	chu	chua	chuo*	chuai	chui	chuan	chun	chuang					
sh	shu	shua*	shuo*	shuai*	shui*	shuan*	shun*	shuang					
r	ru		ruo		rui	ruan*	run						
z	zu		zuo*		zui*	zuan*	zun*						
c	cu		cuo*		cui	cuan*	cun						
s	su		suo*		sui	suan*	sun						
ゼロ	wu	wa*	wo*	wai*	wei*	wan*	wen*	wang*	weng	yu*	yue	yuan*	yun*

*'付の音節は儿化入力可能

表 12 音節表(4/4) (話言葉で使用される音節)

yo
ng

表 13 アルファベット

記号	ASCII	メッセージ	記号	ASCII	メッセージ
A	0x41	ピッチ	g	0x67	PinYin
B	0x42	ピッチ	h	0x68	PinYin
C	0x43	ピッチ	i	0x69	PinYin
D	0x44	ピッチ	j	0x6A	PinYin
E	0x45	ピッチ	k	0x6B	PinYin
F	0x46	ピッチ	l	0x6C	PinYin
G	0x47	ピッチ	m	0x6D	PinYin
K	0x4B	声質変更	n	0x6E	PinYin
L	0x4C	発話長統一	o	0x6F	PinYin
S	0x53	速度	p	0x70	PinYin
U	0x55	ユーザイベント	q	0x71	PinYin
V	0x56	音量	r	0x72	PinYin
W	0x57	高低アクセント・ 文節の韻律変化度	s	0x73	PinYin
X	0x58	声質変更	t	0x74	PinYin
a	0x61	PinYin	u	0x75	PinYin
b	0x62	PinYin	v	0x76	PinYin
c	0x63	PinYin	w	0x77	PinYin
d	0x64	PinYin	x	0x78	PinYin
e	0x65	PinYin	y	0x79	PinYin
f	0x66	PinYin	z	0x7A	PinYin

表 14 記号

記号	ASCII	メッセージ
‘(スペース)	0x20	無音
!	0x21	文節の韻律
#	0x23	ピッチ
\$	0x24	高低アクセント
%	0x25	文節の韻律
&	0x26	高低アクセント強弱アクセントシフトクリア
'	0x27	高低アクセント
*	0x2A	文節区切り
+	0x2B	文節の韻律
,	0x2C	文節区切り
-	0x2D	長音
.	0x2E	文節区切り
/	0x2F	高低アクセントシフトクリア
:	0x3A	文節の韻律
;	0x3B	文節の韻律
<	0x3C	強弱アクセント
=	0x3D	強弱アクセントシフトクリア
>	0x3E	強弱アクセント
?	0x3F	文節区切り
@	0x40	文節の韻律
[0x5B	コメントアウト開始
]	0x5D	コメントアウト終了
^	0x5E	高低アクセント
_	0x5F	高低アクセント
~	0x7E	長音

表 15 数値

記号	ASCII	メッセージ
0	0x30	数値入力
1	0x31	数値入力
2	0x32	数値入力
3	0x33	数値入力
4	0x34	数値入力
5	0x35	数値入力
6	0x36	数値入力
7	0x37	数値入力
8	0x38	数値入力
9	0x39	数値入力

6 デフォルト音色マップ

以下に、デフォルト音色の特徴について記載する。

表 16 デフォルト音色マップ

No.	Name	Basic Voice	Pitch Shift	Fixed Pitch	Prosodic Volume	特徴
0	Normal Man	m	0	—	E	Default Man
1	Normal Woman	w	0	—	E	Default Woman
2	Young Man	m	0	—	E	Fine Young Man
3	Young Lady	w	300	—	E	Gentle Young Lady
4	Boy	m	1200	—	E	Boy
5	Girl	w	1200	—	E	Girl
6	Radio Voice Man	m	100	—	E	Man Voice From AM-radio
7	Radio Voice Woman	w	300	—	E	Woman Voice From AM-radio
8	Hard-boiled	m	-1200	—	E	Man, Hard-boiled
9	Witch	m	0	—	E	Old Witch Woman
10	Nasal Voice	m	0	—	E	Nasal Voice
11	Rat Voice	m	200	—	E	Rat Voice
12	Water	m	0	—	E	Voice Of Underwater
13	Martian	m	0	—	E	Like Alian, Vibrated Voice
14	Robot	m	0	80	E	Robot Voice Without Intonation
15	Synth	m	1200	—	E	Synth Sound

※No の 0～15 は、声質変更（HV-Script のイベント記号の 1 つ）の K0～K15 に対応する。

※Basic Voice は、音色の基本となる性別で、m：男性 w：女性 を意味する。

※Pitch Shift は、男女それぞれ、標準の声に対するピッチの変化量を表す。単位は cent。

※Fixed Pitch は、再生するピッチを固定する場合に使用する。単位は Hz。

ピッチを固定した音色の場合、HV-Script 中のピッチ変化は無視され、指定されたピッチで再生される。

※Prosodic Volume は、HV-Script 中の音量変化指定の有効／無効に使用し、

E：Enable（有効） U：Unable（無効）を意味する。

音量変化指定が無効の音色の場合、HV-Script 中の強弱アクセントや音量指定は無視される。