

# YFX计算机音标处理系统简介

熊 正 辉

微型计算机在我国正日益普及,但在汉语方言调查研究工作中应用计算机却有两个技术上的障碍。一个是现有的微型机操作系统不能处理音标,而方言调查研究时时刻刻都离不开音标。另一个是国家标准编码中的汉字只有 6763 个,对汉语方言调查研究来说不够用。本文作者长期从事汉语方言调查研究工作,在使用计算机做一些工作的时候,总是碰到这两个问题。为此,从工作的实际需要出发,利用手头上使用的计算机和汉字操作系统,研制了一套能处理音标并能随时补充新造汉字的软件系统。为称述方便,暂命名为 YFX 系统。这个系统有以下几个特点:①音标齐全。②打印用的音标字模,仿照铅印音标字模,不同的音标有不同的宽度,因而打印出来比较接近铅字印刷的样子,比较美观。③可以根据自己的需要,随时补充新造汉字。④自动排版。⑤不侵占国家标准中规定的汉字字符编码。⑥对汉字操作系统的原有程序文件不做改动。下面对 YFX 系统做一个简要的介绍。

## 一 音标

方言季刊 1979 年第 2 期刊登了《本刊使用的音标》。这份音标表中一共有 600 个音标。除调号外,每个音标都编有一个供编辑排版用的编号。从计算机输入的角度看,音标表中的音标可以分成四类:①字母音标,编号为 1—398, 484—488, 共 403 个。②符号音标,编号为 399—483, 489—494, 共 91 个。③本调调号 53 个。④变调调号 53 个。

YFX 系统的音标字模完全按照上述音标表设计。打印用的字模,宽度分四类: 24×6 点阵, 24×12 点阵, 24×18 点阵, 24×24 点阵。按汉字的宽度说,也就是有的音标等于 1/4 个汉字,有的等于半个汉字,有的等于 3/4 个汉字,有的等于一个汉字。字模采用不同的宽度,造型比较美观,但增加了技术处理上的难度。

YFX 系统的 600 个音标放在区位码的 88—94 区。这七个区在国家标准编码中空着未用。所以,虽然音标多达 600 个,对原有的汉字字符编码不会侵害。区位码每个区有 94 个码, 7×94 = 658。现在音标只用了 600 个,尚有 58 个码可用来补充一些新的音标。从实际经验看,现有的铅字调号还不是很够用。利用 58 个空位子就可以弥补这一点。

音标的输入,不同类型采用不同的输入法。本调调号最简单,只要用数字键打入调值再加一次空格键。例如按“214 □”(□表示空格键),就输入了“ㄥ”这个调号。变调调号也比较简单,只要在前面加一个数字“0”。例如要输入“ㄨ”这个调号,只需按“0 214 □”。字母音标和符号音标有两种输入法。一种是利用音标表中的编号,只是在编号前加一个加号“+”。例如按“+ 394 □”就可以输入“ㄐ”这个音标,按“+ 432 □”就可以输入送气符号“‘”。做过文字编辑的人会有体会,记住某些常用音标的编号并不困难。另一种输入法是把字母音标按英文字母归大类,符号音标归一大类。像汉字的拼音输入法一样,同类的音标都显示在屏幕的

最底下一行,用数字键做一次选择。为了减少重码的音标数目,还可以在大类下再分小类,用一些字符代表不同的小类。这里就不详述。

音标的输入码对照表是一个独立的文件,用户可以在规定的原则下修改某些音标的输入码,以适应自己的需要。

## 二 新字

YFX 系统把区位码 10—15 区作为新造字的编码。这六个区在国家标准编码中也是空着未用的。六个区共有 564 个码,也就是说,用户可以根据自己的需要补充 564 个新字。当然,如果要处理《广韵》《集韵》《说文解字》等书中的全部汉字,补充五百多个新字是远远不够的,必须另找途径。但是处理一般的方言资料,有五百多新造汉字就可以解决许多问题。

新字的输入,采用四角号码输入法。因为四角号码有重码,所以像拼音输入法一样,还要用数字键做一次选择。

## 三 造字

YFX 系统有三个造字程序: 16×16 点阵汉字和音标, 24×24 点阵汉字, 24×24 点阵音标。16×16 点阵的字模供屏幕显示用, 24×24 点阵字模供打印机打印用。

造字程序除了通常的功能之外,根据实际的需要还增加了几个特殊的功能。例如可以把一个字的一部分笔画暂存起来,然后跟另一个字的一部分笔画合成;可以把字模做 90 度和 180 度旋转;可以把字模做左右对称翻转。这些功能可以提高造字的效率和保持音标的一致性。例如造音标  $v$ , 只要把已经造好的  $a$  调出来做 180 度旋转就成了。又如,把调号 214 $\lambda$  做 180 度旋转就造出了变调号 254  $r$ ,做左右对称翻转就造出了变调号 412 $\lambda$ 。

## 四 自动排版

屏幕显示的时候,音标占一个汉字的宽度,而打印出来则有宽有窄。因此汉字跟音标混排的一段文字,不能像通常那样通过屏幕观察来对齐。解决的办法是在输入原文时,加上一些格式命令,用一个自动排版的程序按照命令要求的格式打印出来。YFX 系统的自动排版程序有以下一些命令。其中  $N$  是数目字,单位是一个西文字符。一个汉字等于两个西文字符,换句话说, $N$  的单位是半个汉字。 $n$  也是数目字,单位是一个点。一个汉字的宽或高都是 24 个点。一英寸为 180 个点。

{行宽  $NNN, nn$ } 用来设定一行打印多少个汉字。 $NNN$  是一行的字符数,  $nn$  是汉字跟汉字间的间距。按点数算,一行的宽度 =  $(NNN \div 2) \times (24 + nn)$ 。

{字距  $nn$ } 在行宽不变的情况下,对一部分字的间距做局部调整。

{行间  $nn$ } 设定两行之间的间距。如空一字为  $nn=24$ , 空半个字为  $nn=12$ 。

{居中} 使标题、小标题等排在一行的中央。

{回行前空  $NN$ } 用来设定一段文字回行时是顶格还是前空一字或几字。如果回行顶格,  $NN=00$ 。

{分栏  $N\bar{N}, NN, \dots, NN$ } 用来设定分几栏排,每栏各占多少个字符。

{加空  $nn$ } 加宽两个字或标点符号之间的间距。如果把  $nn$  换成英文字母“X”或“x”,程序在自动回行时会把一行剩余的点数在此命令后加空,这样可使每行的行尾完全对齐。

{减空  $nn$ } 缩小两个字或标点符号之间的间距。

{尾齐} 分栏排的时候每栏都是头齐。如某栏要尾齐,可用此命令。

因为自动排版程序要按照行宽的要求重新安排一行的字数,所以原文输入时的回车进行

符(编码为 0DH 和 0AH)一律无效。如果用户要在文中某处回行(例如一段的末了),要用字符“#”。在分栏排时,两栏的文字之间用字符“|”隔开。“#{”这三个字符作为命令识别符,就不能当普通字符使用。如果在文中要用到这三个字符时,在原文中要分别写成“\ #”,“\ |”,“\ {”。而要在文中用到“\”这个字符时,则要写成“\\”。

### 五 硬件和软件

YFX 系统研制时用的主机是 IBM-80386,内存 1M,硬盘 40M,VGA 显示卡,显示器分辨率 640×480。打印机是 LQ-1600K。汉字操作系统是 ST-CDOS 4.0B。

### 六 样品

虚线前面是输入的原文,后面是自动排版后打印出来的结果。

{行宽080,02}{行间12}

1234567890123456789012345678901234567890

1234567890123456789012345678901234567890#

{居中}前一字阳上#

{分栏07,07,17,17,17}

4{加空06}1{加空06}[√]{尾齐}-|是非 z1 fi|雉鸡 dz1 tci|痔

疮 dz1 ts'5|坐输 zo cy#

4{加空06}2{加空06}[√]{尾齐}-|苳麻 dz1 mo|肚头 du dx|

尽牙 zin ɲo|在行 ze h3#

(方言1979.5)#

'盪,{加空X}盪'两字顺昌都是阳上。{减空06}'盪暗仔'[t'5-ɬ 5-ɲ ti-ɲ]{加空06}是'黑下了,傍晚'的意思。{减空06}'盪冤枉缸'[t'5-ɬ ɲɔŋ ɲ5-ɬ t5ŋ]是说'落入冤枉缸,受冤'。顺昌'投衣服'也说'盪'[t'5-ɬ]。{减空08}{加空X}'盪盆'[t'5-ɬ p'ue-ɲ]是说婴儿洗三。由顺昌阳上可以推定这两字都是来自古定母上声。#

(方言1992.113)#

{回行前空04}438 娘仔嘢!姻缘都是天注定,月老定会安排。[nioŋ kīāŋ aŋ!

ɲ-ɲ uaŋ tou-ɲ siŋ t'ī-ɲtsu-ɲ tīāŋ, gueŋ-ɲ

lauŋ tīāŋ oiŋ ũā-ɲ paiŋ] (方言1991.28)#

-----  
1234567890123456789012345678901234567890123456789012345678901234567890

前一字阳上

41[√ -] 是非 z1 fi 雉鸡 dz1 tci 痔疮 dz1 ts'5 坐输 zo cy

42[√ -] 苳麻 dz1 mo 肚头 du dx 尽牙 zin ɲo 在行 ze h3

(方言1979.5)

'盪,盪'两字顺昌都是阳上。'盪暗仔'[t'5-ɬ 5-ɲ ti-ɲ]是'黑下了,傍晚'的意思。'盪冤枉缸'[t'5-ɬ ɲɔŋ ɲ5-ɬ t5ŋ]是说'落入冤枉缸,受冤'。顺昌'投衣服'也说'盪'[t'5-ɬ]。'盪盆'[t'5-ɬ p'ue-ɲ]是说婴儿洗三。由顺昌阳上可以推定这两字都是来自古定母上声。

(方言1992.113)

438 娘仔嘢!姻缘都是天注定,月老定会安排。[nioŋ kīāŋ aŋ!

ɲ-ɲ uaŋ tou-ɲ siŋ t'ī-ɲtsu-ɲ tīāŋ, gueŋ-ɲ lauŋ tīāŋ oiŋ ũā-ɲ paiŋ] (方言1991.28)

作者要感谢翟英谊同志,他在制作和修改音标字模的过程中协助做了大量的工作。