

# 提高朗读软件的读音准确性

李维良 ( 江苏沛县大屯煤电公司退休管理处 221611)

**摘要:** 目前市面上出售的汉语朗读软件读音准确率不高,影响了其广泛应用。必须对要朗读的文件进行“预处理”,辨别多音字的正确读音,以及确定符号、数字和格式的正确读法,产生一个读音准确的“音稿”文件,再让发声软件读念,才能达到100%的准确性。

**关键词:** 朗读 读音 音稿

## 1 引言

我们把计算机里的文本文件能读念出声音的软件称作朗读软件。目前市面上流行的汉语朗读软件已有很多种,虽然都说自己“读音准确、流畅、自然、悦耳”,但文章的平均读音准确率都只有百分之九十几。还有百分之几的读音是错误的或者是不规范,特别是象“长、行、为、着、了、地、干、分、发、处、差”等多音字和一些符号,在一般文章中出现频度是很高的。听了这些软件的读念以后,给人的印象是有不少读错的字,有些话听不懂,有时很别扭甚至出笑话。所以这些朗读软件还不够成熟,不能用在正式场合。

产生这百分之几读错率的原因是,发声软件是按文本文件逐字地、不加分析地去读,当然错误难免。比如,把“为人民服务”和“为人民所喜爱”中的两个“为”都读成\wei4(其中小写字母代表汉语拼音,数字1-5代表声调阴平、阳平、上声、去声和轻声,下同)。对“受不了”中的两个“了”字,和“这是好酒,他好喝酒”中的两个“好”字的读音都不能区分:

把“行行出状元”读成\xing2\xing2\chu1\zhuang4\yuan2,把“护士长”读成\hu4\shi4\chang2,把“干着急”读成\gan4\zhao2\ji2,把“积极分子”读成\ji1\ji2\fen1\zi3,把“发廊”读成\fa1\lang2

把1999/12/06读成“一九九九杠一二杠零六”,把“占1-2%”读成“占一减二百分之”,把\$100读成“美元一百”;

把“共110人”和“请拨打110”中的数字都读成“一一〇”,把12亿人口读成“一二亿人口”。

新研制的《汉语说念软件HYSN1.0版》的思路是,不直接读文本文件,而是把要朗读的文件先进行“预处理”,很快生成一个读音准确的“音稿”文件,再交给发声软件逐字地读念。

## 2 确保读音准确的方法

### 2.1 对数字或者符号读法进行规范化处理

比如上面提到的“占1-2%”,在音稿文件中换成“占百分之一到二”、把\$100换成“一百美元”就可以照读无误了。这种对要读念的符号进行换字、增减或颠倒的过程只有通过另外生成一个“音稿”文件的办法解决。

问题是有的符号又有两种以上含义和读法,如-、.: /等符号。这要经过进一步判断。比如/符号,当程序检验出其前后只有两个数字,而且前面数字的值小于后面数字时(一般文章中都是真分数),要改成“几分之几”的读法,否则读成\gang4。

对日期读法的处理也是一样,如年、月、日的写法有1999.12.06、1999/12/06和1999-12-06等。当遇到三个数字连续出现,并对三个数字值进行检验,而且其间只有两个./或-符号时,换成“一九九九年十二月六日”的读音。

有些数字需要按传统读法,即增加“位数名称”去读。如上例的“共110人”应换成“共一百一十人”,把“12亿”换成“十二亿”。遇到数字是否这样读,一般要看其说的是“号码”还是“数量”,即后面有没有量词而定。

### 2.2 判别多音字的正确读法

该软件首先把汉语中的常用词汇(约8万多条)编入一到四字词的词库(这里把单个汉字有独立含义、以词的身份出现时,称一字词)。并对每个词汇的正确读音按照《新华字典》和《现代汉语词典》等权威性辞书存入正确发音。每个发音后面还要存入可能的词性。本软件把汉语中的12种词性进行适当的合并和简化,以便于多音字词性的鉴别,本文不再赘述。

三四字词和大多数一二字词每个词汇只有一个读音。这种以词确定读音的方法比“逐字”确定读音的方法解决了很多问题,如“不了了之、护士长、干打垒、行走、喜

好”等词汇中虽有多音字,但在词库中找到整个词汇的正确读音,就不会出错了。

当一字词(包括某些二字词)具有两个以上的读音时比较难于判断。但是,既然是“多音字”,必然含意不同,出现场合也不同。究竟采用哪种读音,在语句中肯定是有规律可循的。

在GB\_2312汉字系统中有6763个汉字,其中有700个左右的多音字,把它分成三类:

①有200多个属于生僻字,在一般文章中很少见到,如“白术”在中药里读作\ bai2 \ zhu2;“读”这个字在古代当“停顿”讲时读成\ dou4(相当于逗)。程序中对这些字按一般读法去读,需要采用特殊读音时,要靠用户使用下文中所说的自注音功能解决。

②还有200多个多音字不单独出现,也就是只在组成二三四字词时才出现,如“恶”只在“憎恶、可恶”等词中才读作\ wu4;文章中以一字词身份出现的“恶”只能读作\ e4。还有象“阿、便、扁、别”等多音字都只有在组成“阿胶、便宜、一叶扁舟、别扭”等词汇时才读音特殊,上面已经解决;它们在以一字词身份出现时读音都是唯一的。这一类多音字也不需要判断。

③还有160多个多音字常常以一字词的身份出现,对这些一字词(包括某些二字词)的两个以上读音和对应的词性也都存入词库,如“好”要存成“\ hao3x \ hao4f”,“发”要存成“\ fa1d \ fa4m \ fa1l”(其中x f d m l分别代表形容词、副词、动词、名词和量词)。然后按照以下途径进行判断和取舍。

该程序运行时,每遇到一句话,从四字词开始向一字词逐一检索,确定这句话由哪几个词汇组成。再把每个词到词库中去查它的正确发音和词性。当只有一个音时,照搬入音稿文件;当它有两个以上发音时,主要参照其前后的词序、词性,也就是根据语法结构进行判断。

在汉语中相邻词汇往往组成一个短语,构成一个相对独立的语义单元,如,数词+量词+名词(象“十斤苹果”)、形容词+名词(象“美丽的村庄”)、副词+动词(象“积极贯彻”)……等等。程序中罗列了汉语中可能发生各种短语结构。

当含有多音字的一二字词到词库中检索时,同时查到了其可能的几种发音和相应的词性,然后结合相邻的词汇查验其可能的语法结构,从而找到正确读音。

比如,“这是好酒”这句话,“好”字虽然有两种读音和词性(形容词hao3和副词hao4),由于后面是名词“酒”,

属于“形容词+名词”结构,应采用对应形容词的读音,应读作\ hao3。而“他好喝酒”中“好”字后面是动词“喝”,属于“副词+动词”结构,采用对应副词的发音,应读作\ hao4。再如,“她穿着时装,她讲究穿着”句中的两个“穿着”也是根据语法结构判断出前一句“穿着”是动词,应读作\ chuan1 \ zhe5;后一句属于名词,读作\ chuan1 \ zhuo2。

又比如,遇到“为”这个词时,当发现其后面为“名词+动词”结构时,读作\ wei4,如“为大家办了实事”;属于“为+名词+所+动词”结构应读作\ wei2,如“为专家们所认同”。同样,属于“以+名词+为+...”结构(如以大局为重),也应读作\ wei2。

这里说的语法结构中的那些单元,有时候是一个词汇,有时候是一个短语。如“她穿着时装”和“她今天穿着鲜艳的时装”都按同一语法结构模式进行判断。

有时候还要靠其前后的词汇来决定其正确读音。如“干”,除了使用上面提到的使用语法结构进行判断外(属于形容词或副词时应读作\ gan1;属于动词或名词时应读作\ gan4),还要参照其前后词汇中有“衣服、油漆、溶液”一类的词汇时读\ gan1;有“工作、事情”一类的词汇时读作\ gan4。

读念长句子时,中间也应该有适当的停顿。而这种停顿必须放在其中的两个短语之间效果才好。如“我们的物质文明方面和精神文明方面都需要加紧建设。”这句话中应在“和、都”两字前面稍作停顿,其他地方停顿反而可能弄巧成拙。而要做到停顿恰当,只有在经过上述的语法结构分析,确定这个句子有哪几个短语组成之后才能办到。

有些字的音调应该由程序作自动的鉴别,如“一”就有三种声调、“不”有两种声调。比如,始终如一\ shi3 \ zhong1 \ ru2 \ yi1、一天一夜\ yi4 \ tian1 \ yi2 \ ye4、不吃不睡\ bu4 \ chi1 \ bu2 \ shui4。

上面不过举了几个简单的例子,实际上在程序中有时为了一个多音字要经过几道手续和十几个if或case语句的判断。判断得越全面越严谨,准确率就越高。这样,发声软件再按音稿逐字朗读时,就不会出现明显的错误,让人容易听懂,准确率可以达到99.5-99.9%以上,足以满足日常收听的需要。

### 2.3 自注音功能

尽管如此,还不能保证100%的准确性。原因有二,一是计算机“没有语文和社会实践知识,目前还不可能真正

懂得文章的本意”，计算机程序只能从上下文的“字面上”意思去判断；二是因为汉语的字里行间内涵十分丰富，象一些人名、小地名、简称、译音、口语或古文中间包含多音字时，或者因句子太短、书写不规范等原因，仍然可能出现个别读音错误。该程序增加了自注音功能，假如用户要求读念准确率必须是100%时，应允许在这个读错汉字的后面，加注理想的注音。如“战国七雄指的是齐楚燕\yan1 韩赵魏秦”，“北京前门外有个大柵\shi5 柵\lan5 大街”，以及绕口令中“喇嘛提了五斤鳎目\ma5”等。

程序中遇到自注音时将其照搬入音稿文件，不另判断。也可以让用户将其常用的特殊词和音存入自定义的词典音库，文章中遇到这种字词时，首先在这里选取读音。

#### 2.4 正确辨析欲朗读文件的格式

对文本文件中的标题、注释、空格、换行、回车、各种标点、段落缩进及表格等也都要认真分析处理，近年来的网上文章书写格式更是五花八门，该程序力图让计算机理解这些格式的含义，反应在音稿里，并舍弃不必要的内容，才能让人不会产生歧义、听得明白和舒服。程序生成的音稿文件是一个虚拟文件，不和用户见面。可以是逐

句、逐段、逐页或全文形成一个音稿文件，分别去读念，而且文件名是相同的，占内存很小。音稿文件根据发声软件的要求，可以由每字的拼音组成，也可以由不是多音字的汉字组成（如“好喝酒”改成“耗喝酒”；“酒好喝”改成“酒郝喝”）。形成音稿文件耗用的时间也很少，可以让用户感觉不到。

对欲朗读的文件采用首先运行本软件形成音稿的办法，解决了读音准确性的问题。若再配合发声程序，对语句的轻重音、语调、停顿和柔和性进行恰当处理，就可以组成完美的朗读系统软件。只有这样，才能让计算机真正代替人的“嘴”讲话，才会有广泛的应用前途。比如在汉语教学、推广普通话、外国人学汉语、培训播音员、文章校对、咨询讲解、为盲人服务、读念网上文章，以及用来制作玩具、模特和机器人等方面都有其“用武之地”。■

#### 参考文献

- 1 李维良，汉语的随想随打与句输入，计算机系统应用，1993.9。
- 2 李维良、高正祥，随想随打写文章，电子工业出版社，1996.2。
- 3 陶志新等，快捷编程 Visual Basic 5.0，人民邮电出版社，1999.2。