

面向方言语音合成的文本分析研究

郭威彤¹ 杨鸿武¹ 宋继华² 顾 香¹ 甘振业¹

(1. 西北师范大学物理与电子工程学院, 兰州 730070; 2. 北京师范大学信息科学与技术学院, 北京 100875)

摘 要: 为实现方言的统计参数语音合成, 提出一种从文字到方言读音的文本分析方法。通过对比普通话和方言在声韵母方面的发音异同, 设计方言的语音评估方法音标字母(SAMPA), 用来标注方言声韵母的读音, 得到从普通话读音到方言读音的转换规则。对输入的汉语文本进行分析, 获得语法词、声母、韵母信息, 使用基于转换的错误驱动学习算法获得语句的韵律词和韵律短语边界, 利用普通话读音到方言读音的转换规则, 获得方言发音的SAMPA音标, 从而将输入的文本转换为统计参数语音合成所需的上下文相关标注。测试结果表明, 该方法能较为准确地生成上下文相关标注。

关键词: 文本分析; 字音转换; 语音评估方法音标字母; 语音合成; 语法分析

中文引用格式: 郭威彤, 杨鸿武, 宋继华, 等. 面向方言语音合成的文本分析研究[J]. 计算机工程, 2015, 41(9): 184-189.
英文引用格式: Guo Weitong, Yang Hongwu, Song Jihua, et al. Research on Text Analysis for Dialect Speech Synthesis[J]. Computer Engineering, 2015, 41(9): 184-189.

Research on Text Analysis for Dialect Speech Synthesis

GUO Weitong¹, YANG Hongwu¹, SONG Jihua², GU Xiang¹, GAN Zhenye¹

(1. College of Physics and Electronic Engineering, Northwest Normal University, Lanzhou 730070, China;

2. College of Information Science and Technology, Beijing Normal University, Beijing 100875, China)

【Abstract】 A text analysis method for converting grapheme to dialect phoneme is proposed for statistical parametric dialect speech synthesis. A set of Speech Assessment Methods Phonetic Alphabet(SAMPA)-based symbols are designed for labeling pronunciation of dialect by comparing the differences between Mandarin and dialect. A set of conversion rules is also designed that can transform Mandarin pronunciation to dialect pronunciation. The text analysis is conducted for Chinese sentences to obtain lexicon words and their initials and finals. A transformation-based error driven learning algorithm is used to obtain the prosodic words and prosodic phrases boundaries. The conversion rules are employed to obtain the SAMPA of dialect initials and dialect finals. The input sentences are converted into context-dependent labels. Test result shows that the proposed method can generate correct context-dependent labels.

【Key words】 text analysis; grapheme-to-phoneme conversion; Speech Assessment Methods Phonetic Alphabet(SAMPA); speech synthesis; syntactic analysis

DOI: 10.3969/j.issn.1000-3428.2015.09.034

1 概述

语音合成^[1]作为一种新的信息传递技术, 已被应用在人机交互中。然而, 不同民族、不同国家有自己不同的语言, 因此, 多语种的语音合成成为了人机语音交互领域的研究热点^[2-3]。

目前, 汉语普通话的语音合成系统已经比较成熟, 可以合成出自然度、可懂度较高的语音, 这很大

程度上是因为有比较完善的汉语文本分析系统, 可以正确地将输入的汉语文本转换为拼音。文本分析首先将输入的文本进行规范化处理, 获得特殊符号的读音。进而对文本进行分词, 获得文本的词边界和词性。在此基础上, 通过韵律预测, 获得文本的韵律边界信息。最后, 利用字音转换获得文本的正确读音。文献[4]利用数据驱动的方法, 采用二元文法对文本进行分词, 并训练韵律结构预测模型。文

基金项目: 国家自然科学基金资助项目(61263036, 61262055); 甘肃省杰出青年基金资助项目(1210RJDA007); 甘肃省青年科技研究计划基金资助项目(1208RJYA078); 西北师范大学青年教师科研能力提升计划基金资助项目(NWNU-LKQN-12-27)。

作者简介: 郭威彤(1982-), 女, 硕士研究生, 主研方向: 自然语言处理、模式识别; 杨鸿武(通讯作者), 宋继华, 教授、博士; 顾 香, 硕士研究生; 甘振业, 副教授、博士。

收稿日期: 2014-09-11 修回日期: 2014-10-21 E-mail: guowt@nwnu.edu.cn

献[5]对输入文本进行分词及词性标注,利用句法分析树进行实体识别及搭配词识别。文献[6]通过在文本分析结果的基础上,引入韵律节奏的预测机制,实现了文本处理和韵律预测的融合。文献[7]提出一种基于支持向量机(Support Vector Machine, SVM)的多音字规则自动调整体系,有效地处理了汉语多音字的问题。以上的研究工作提高了汉语普通话文本分析的准确性。

中国是一个地域辽阔,方言和民族语言丰富的国家,大多数人说的都是方言或者是带方言口音的普通话。因此,研究普通话到方言的跨语言语音合成,实现同一个说话人的普通话、方言和带方言口音的普通话语音的语音合成对于自然友好的人机交互有重要的意义。近年来,对于面向语音合成的民族语言的文本分析,已经展开了研究。如文献[8]实现了正向和逆向的最大匹配藏文分词算法,文献[9]采用最大匹配法和分词词库相结合的方法实现藏语文本的自动分词,文献[10]总结了维吾尔语的音节划分规则、词根词缀的划分规则以及韵律变化规则。但是现有研究缺少对方言的标准字音转换的分析,缺乏对方言特殊发音的考虑,无法获得方言准确的韵律上下文的信息,从而不能合成出自然的方言语音。目前已实现了中英文混合语音合成^[11]以及普通话到闽方言中台湾话的语音转换^[12],在方言的语音合成中,虽然实现了华北方言中聊城话^[13]、沈阳话^[14]、天津话^[15]、兰州话^[16]的语音合成,但只是利用语音修改技术,将普通话的韵律修改为方言的韵律,不能合成出方言特有的读音。

普通话和兰银官话是甘肃地区的主要语言,而兰州方言又是兰银官话的一个重要代表。本文在汉语文本分析的基础上,进行普通话和兰州方言混合语言的文本分析,为面向基于隐马尔可夫模型(Hidden Markov Model, HMM)^[17]的语音合成系统提供上下文相关的标注,从而实现方言/普通话的混合语言语音合成。

2 语音评估方法音标字母标注设计

语音评估方法音标字母(Speech Assessment Methods Phonetic Alphabet, SAMPA),即机读音标,是欧洲的 ESPRIT 开发的一种计算机可读的音标系统,用 ASCII 字符表示国际音标的所有符号,表示世界上各种语言。本文利用张家骅修订完善的汉语拼音的机读音标(SAMPA-SC)^[18]的设计思想,设计兰州方言的机读音标(SAMPA-LZ)。通过对照普通话和兰州方言的声韵母的国际音标,发现两者大部分的国际音标是相同的。因此,本文以国际音标为参考,

将两者国际音标一致的部分,直接利用汉语拼音标记兰州方言的读音。对于两者国际音标不一致的部分,则采用简单化原则,利用定义的简单符号进行标记。设计流程如图 1 所示。

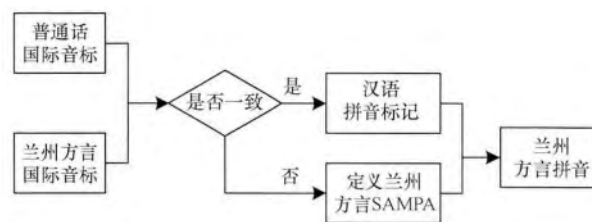


图 1 兰州方言的机读音标设计流程

2.1 声母

普通话共有 22 个辅音,其中 21 个都可作声母,另一个辅音/ng/只作韵尾。而兰州方言共有 26 个声母,其中 21 个声母的国际音标和普通话的国际音标一致。对这 21 个兰州方言的声母,直接用对应的汉语拼音来标音。如汉语拼音声母/p/和兰州方言声母/p/的国际音标都是/pʰ/,则将兰州方言声母/p/的读音标记为汉语拼音声母/p/。

兰州方言剩余的 5 个声母是兰州方言特有的声母,无法用汉语拼音表示出来,但是都有相应的国际音标,用国际音标表示这 5 个声母,分别是 pf, pfʰ, v, z, n₀。对这 5 个声母,以国际音标为基础,根据简单化原则,按如下方法设计 SAMPA-LZ:

(1) 如果国际音标(International Phonetic Alphabet, IPA)可以直接用 ASCII 字符表示,并且从未使用过该国际音标,则直接采用国际音标来定义兰州方言的 SAMPA-LZ。

(2) 对于难以用键盘输入的国际音标,使用和和国际音标相近的未使用过的键盘符号来定义。例如兰州方言中的声母/n₀/,无法直接从键盘键入,为了和已经使用过的/n/区别,就用符号 n'来表示声母/n₀/。表 1 给出兰州方言中特有的 5 个声母的国际音标和定义的机读音标 SAMPA-LZ。

表 1 兰州方言独有声母的机读音标

国际音标	兰州方言的机读音标	实例	对应汉字
pf	pf	Pfu	主
pfʰ	pfʰ	pfʰu	处
v	v	ṽə	嗯
z	z	zan	软
n₀	n'	n'i	泥

表 1 列举汉字的兰州方言读音和普通话完全不同。对于这样一些汉字,需要根据声韵母转换规则进行转换。

2.2 韵母

普通话共有 38 个韵母,除了 3 个特殊的单元音

韵母外,常用的韵母有 35 个。兰州方言共有 32 个韵母,其中,有一部分韵母的发音和普通话的韵母一致,还有一些韵母是普通话没有的。针对这些韵母,采用和声母相似的简单化原则,利用易于从键盘输入的符号进行标音。表 2 列出了兰州方言特有韵母的国际音标、对应的汉语拼音、汉语拼音的机读音标以及定义的兰州方言的机读音标。

表 2 兰州方言特有韵母的机读音标

汉语拼音	汉语拼音的机读音标	兰州方言特有韵母的国际音标	兰州方言的机读音标
an	an	$\tilde{\varepsilon}$	An
en	@ n	$\tilde{\partial}$	eN
in	in	$\tilde{i}\partial$	ieN
uen	u@ n	u	ueN
uan	uan	u	uAn
vn	yn	y	veN
ang	aN	\tilde{a}	aA
uai	uai	u	uE
ao	au	ɔ	auo
iao	iau	iɔ	iauo
ou	ou	əu	oiu
iou	iou	iəu	ieu
ian	ian	i	iAn
van	yan	y	vAn

在表 2 中,汉语拼音中的韵母/an/,在兰州方言里变为韵母/ $\tilde{\varepsilon}$ /,而这个韵母的国际音标与普通话的不同,因此,定义了与汉语拼音/an/的机读音标/an/不一致的标音符号/A $\tilde{\varepsilon}$ /来代表兰州方言韵母/ $\tilde{\varepsilon}$ /。其他韵母都采用同样的方法处理。

2.3 声调

声调是语音必不可少的特征,具有区别意义的作用。普通话孤立音节有阴、阳、上、去和轻声 5 种不同的声调,而兰州方言只有阴、阳、上、去 4 种声调,但是两者的调值完全不同。因此,本文利用声调的调值作为兰州方言声调的机读音标兰州方言的机读音标,如表 3 所示。

表 3 兰州方言声调的机读音标

兰州方言声调	调值	兰州方言的机读音标
阴平	53	53
阳平	51	51
上声	33	33
去声	13	13

2.4 基于 SAMPA-LZ 的拼音转写

在跨语言的语音合成中,需要利用文本分析将输入的文本转换为普通话或者方言的读音。在本文中,首先利用普通话的文本分析,将汉字转换成汉语拼音。然后通过对比普通话读音和兰州方言读音,

一致的用汉语拼音表示,对于不一致的,在基于普通话和兰州方言声韵母转换规则的基础上,利用本文定义的 SAMPA-LZ 标音规则,最终将汉语拼音转换为兰州方言的读音。

2.4.1 普通话和兰州方言声韵母的转换规则

普通话和兰州方言在声、韵、调上有一致的地方,但存在许多不同的发音。本文通过分析普通话和兰州方言在声、韵、调上的异同,总结出了 41 条普通话的声韵母转为为兰州方言的声韵母的转换规则。转换规则用统一的格式表示“→”左边是普通话的声母+韵母,“→”右边是对应的兰州方言的声母+韵母。普通话的声韵母用汉语拼音表示,兰州方言的声韵母如果和普通话的一致,就用汉语拼音表示,不一样的用定义的机读音标 SAMPA-LZ 表示。例如,规则 b p m s + o ao ai ei → b p m s + e 的意思是,如果普通话的声母/b/ /p/ /m/ /s/和韵母/o/ /ao/ /ai/ /ei/组合,则在兰州方言中,声母不变,但韵母都发/e/的音。如“白”在普通话中读/bai/,在兰州方言中读/be/。

2.4.2 普通话和兰州方言声调的转换规则

虽然普通话和兰州方言都有阴、阳、上、去 4 种声调,但是由于两者在每一种声调上的调值完全不同,使得这 2 种语言在听感上大相径庭,这就是所谓的变调。普通话和兰州方言变调规律如表 4 所示。

表 4 普通话和兰州方言调值变化规律

调形	普通话	兰州方言
阴平	55	53
阳平	35	51
上声	214	33
去声	51	13

同时,本文还考虑了兰州方言在连续语流中的变调规律:

(1) 2 个上声相连时,前上变阳平 51。

(2) 去声在阴平、阳平、上声、及去声前一律变成中平 33。

(3) 2 个阳平字相连,前字变成中平 33。

3 兰州方言的文本分析

语音合成系统分为前端的自然语言处理和后端的语音信号生成 2 个部分。自然语言处理主要包括文本分析模块和韵律生成模块。在整个语音合成系统中,文本分析起着关键性的作用。文本分析的结果直接影响韵律预测的准确性和合成语音的自然度。本文以声韵母作为合成基元,利用普通话的文本分析,在语法规则知识库和语法词典的指导下,通过文本规范化、语法分析和韵律边界预测,获得输入

文本的普通话拼音、词边界信息、韵律边界信息和语句信息。然后利用兰州方言的标音系统,修改与普通话发音不一致的拼音。进而根据兰州方言声韵母组合的变音规律,实现兰州方言的变音处理,得到输入

入文本的兰州方言读音。在此基础上,利用文本分析获得的上下文信息,生成语音合成后端所需要的上下文相关标注。图2给出了普通话和兰州方言的文本分析流程。

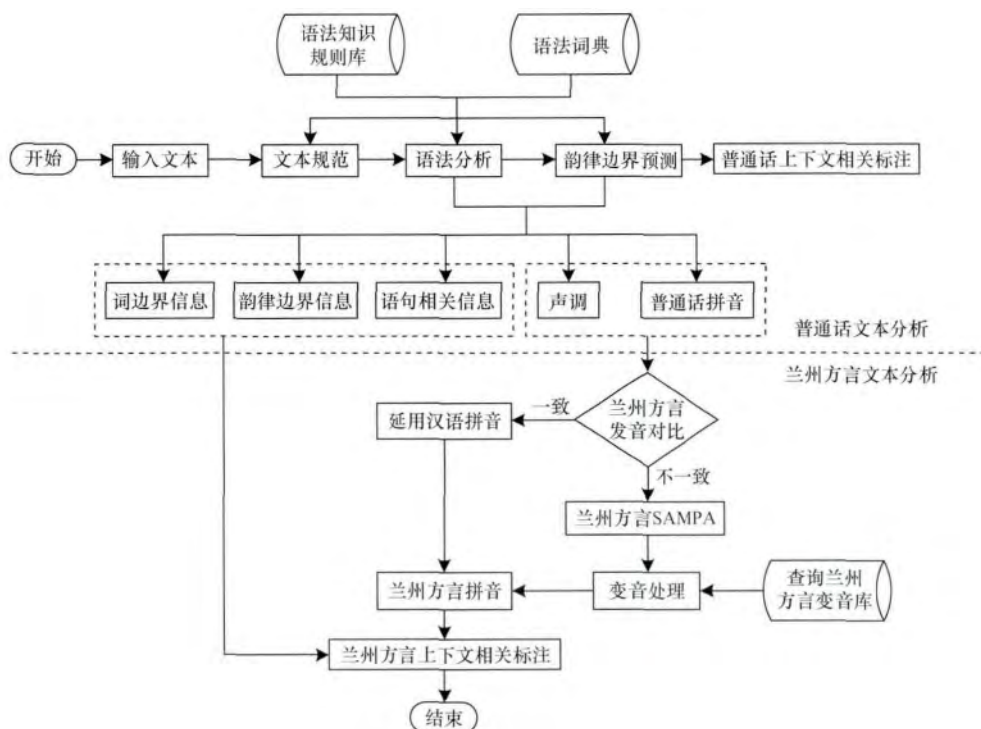


图2 普通话和兰州方言文本分析流程

语音合成后端利用上下文相关的标注生成声韵母的语音信号,而上下文相关的标注中最关键的信息是声韵母的读音,以及声韵母的上下文相关信息。普通话和兰州方言有相同的上下文信息,因此,利用普通话的文本分析,经过文本规范化、语法分析和韵律边界预测,即可获得兰州方言的上下文相关信息。但是,因为兰州方言有部分发音与普通话不同,无法用汉语拼音来表示这些特殊发音。对于这样一些特殊发音的兰州方言的声韵母,利用兰州方言的机读音标进行标音。对于输入的汉语文本,通过遍历语法词典将文字序列转换成普通话的拼音序列,从而获得普通话的声韵母读音,然后查找机读音标SAMPA-LZ,修改特殊的兰州方言读音,结合文本分析得到的词信息、韵律边界信息和声调信息,最终得到兰州方言的上下文相关标注。

3.1 文本规范

文本规范就是将非汉字字符串转换成汉字串以确定读音的过程^[19]。对输入文本进行分析,将文本中除中文字符以外的非标准词,如英文字符、数字字符以及符号字符转换成对应的汉字。

文本规范的处理效果直接影响着文本拼音信息的正确性。一个非标准词在不同的上下文可能对应不同的标准发音,所以,必须从特殊符号出发,提取

有用的上下文信息,归纳出在特定环境下的不同处理策略。如“985 高校”需要按照字符串规范记为“九八五”,而“985 名教师”则需按照数字规范为“九百八十五”。本文利用有限状态自动机方法,采用最长匹配策略,利用词典从真实文本中将最长串识别为非标准词。然后,采用最大熵算法的统计模型,选取适当的特征模板训练建模,同时,设定一定规则,对部分非标准词消歧,最后,再次通过遍历词典,产生非标准词的标准拼音。

有时文本规范对符号的理解不能简单通过上下文确定。可将不同文本规范化的结果全部保存,在后续的处理过程中获得了足够的信息后,根据一定的准则再做判断。

3.2 语法分析

目前语法分析多以句子为划分对象而不是以整篇文本作为划分对象,所以,首先要对文本进行句子划分。本文确定句子边界的基本思路是利用标点符号。对于纯汉语文本来说,能确定句子边界的符号有“;”“。”“、”“?”“!”“”“”“!”等。在句子边界确定以后,利用N元文法模型,结合动态规划(Dynamic Programming, DP)算法来完成分词。假定一个单词出现的概率分布只与这个词前面的 $n-1$ 个单词有关,与更早出现的无关,即:

$$W^{\text{opt}} = \operatorname{argmax}_p (w_1) p(w_2 | w_1) \cdots p(w_n | w_{n-1})$$

其中 w_1, w_2, \dots, w_n 表示词串; W^{opt} 表示最优词语序列。然后根据每个词与相邻词的结合概率,从各种可能的组合中找出概率最大的词串作为划分结果。本文使用来自人民日报语料库的 56 382 个词条的词典和一个多音字字典,将划分出的词查词典来确定拼音。

3.3 韵律边界预测

准确的韵律特征是合成高质量语音的保证,获取准确的韵律结构是获得准确韵律特征的保证。汉语的韵律层级分别是韵律词、韵律短语和语调短语。一般来说,语调短语的边界容易判断,基本上可以认为标点符号(逗号、句号、问号、感叹号等断句标点符号)就是语调短语的边界。对于韵律词边界和韵律短语边界,本文基于邻接度(Adjacent Degree, AD)特征描述语法结构和韵律结构的关系,利用基于转换的错误驱动学习算法(Transformation-based Error Driven Learning Algorithm, TBL)实现韵律词和韵律短语的预测^[20]。

本文选取前一个语法词、当前语法词和后一个语法词的词长、词性,以及均匀分布模型产生的韵律词预分类信息作为韵律边界预测的输入特征。定义 300 多条韵律规则模板,利用贪婪搜索算法对样本的输入特征及样本上下文的输入特征进行搜索,将搜索到的实例特征填入相应的模板规则中。在遍历完模板中的所有模版规则后,就得到了该样本的实例化规则。针对文本中的每一个样本重复以上的搜索过程后,就得到所有样本的完整实例化规则集,利用 TBL 算法实现对未标注文本韵律边界的预测。

4 测试结果与分析

为了验证本文提出的文本分析方法能够正确产生方言语音合成所需的上下文相关的信息,选取普通话和兰州方言的所有声韵母以及静音和停顿作为上下文相关的语音合成基元,设计了一套上下文标注格式,用来标注每个合成基元的声韵母层、音节层、词层、韵律词层、韵律短语层和语句层的上下文相关特征,如表 5 所示。

表 5 上下文相关的标注格式

标注层	标注格式
声韵母层	$p1 \sim p2 - p3 + p4 = p5 @ p6 \sim p7$ /A: a1_a2 - a3_a4 = a5
音节层	/B: b1_b2! b3_b4#b5@ b6! b7 + b8@ b9#b10_b11 /C: c1 + c2 - c3 = c4#c5
词层	/D: d1 - d2/E: e1&e2^e3_e4/F: f1 - f2/
韵律词层	/G: g1 - g2/H: h1 - h2@ h3 + h4/I: i1 - i2
韵律短语层	/J: j1&j2 = j3 + j4/K: k1 = k2_k3^k4&k5_k6/L: l1!2 = l3 - l4
语句层	/M: m1#m2 + m3 + m4! m5

表中小写字母是变量(指代包括音段信息、声调信息、位置信息、词性信息、句型信息等);大写的字母是不同标注层的标示符号。“+”,“-”等符号用于区分不同模式串。

普通话和兰州方言采用相同的上下文相关的标注格式。 $p1 \sim p7$ 表示在基元层每一个声韵母的相关信息; $p3$ 表示当前的声韵母; $p2$ 是 $p3$ 的前一个声韵母; $p4$ 是 $p3$ 的后一个声韵母; $p1$ 表示 $p2$ 的前一个声韵母; $p5$ 表示 $p4$ 的后一个声韵母; $p6, p7$ 表示 $p3$ 所处的音节中的前后位置;在音节层到词语层分别用大写字母 A ~ M 表示。在音节层, A 表示前一个音节; B 表示当前音节; C 表示后一个音节,在所表示的音节内不同符号分别表示音节的声韵母信息和声调信息。在词层, D 表示前一个词; E 表示当前词; F 表示后一个词,分别给出了词的位置信息和所包含的音节数信息。同样, G, H, I 表示韵律词层,包含韵律词的位置信息与韵律词中包含的字信息。用 J, K, L 表示韵律短语层,包含韵律短语的位置信息,以及韵律短语中包含的词和韵律词的信息。M 表示语句层,包含语句中所含的音节、词、韵律词、韵律短语的数量信息以及语句的句型信息。陈述句用 d 表示,疑问句用 q 表示,祈使句用 i 表示,感叹句用 e 表示。

在上下文相关的标注格式里,不仅考虑合成基元各个层级的上下文相关特征,还涉及静音和停顿部分。静音段和停顿段的表示如表 6 所示。

表 6 静音与停顿的符号

类型	符号
起始或结尾的静音	sil
句子中的短停	pau

对于输入的汉语语句,首先利用文本分析获得每个音节的兰州方言声韵母信息,组成训练声学模型所需的单音素标注文件。同时,利用 TBL 算法获得语句的韵律词和韵律边界信息。在此基础上,利用单音素标注、韵律边界信息和文本分析获得的词信息,生成上下文相关的标注文件,用于声学模型的训练和语音合成。图 3 给出了汉语语句“春天像个小姑娘花枝招展的笑着”的部分上下文相关标注,包括单音素标注中的所有声、韵母和音节层的上下文相关信息,省略了其他各层信息。如图 3 的第 3 行中的“sil^ch - un + t = iaN”表示当前韵母 /un/ 的前一个音素为 /ch/; 前音素为 sil; 后一个音素为 /t/; 后音素为 /iaN/。可以看出,文本分析程序正确地获得了汉语语句的方言发音信息。根据表 5 可知,“/A”,“/B”,“/C”分别表示当前韵母 /un/ 的前一个音节、当前音节和后一个音节的上下文信息,这也与

“春”字字母的上下文相关信息是一致的,表明文本分析正确地获得了方言的声韵母信息以及每个声韵母所在的上下文相关信息。

	1	2	3	4	5	6	7
1	xx^xx-sil+ch=un@x	x/A:xx	xx-xx	xx=xx/B:xx	xx!xx	xx#xx@xx!xx+xx@xx#	
2	xx^sil-ch+un=t@1	2/A:xx	xx-xx	xx=xx/B:ch	un!11	11#2@1!2+1@2#1	7/C:
3	sil^ch-un+t=ian@2	1/A:xx	xx-xx	xx=xx/B:ch	un!11	11#2@1!2+1@2#1	7/C:
4	ch^un-t+ian=h@1	2/A:ch	un-11	11=2/B:t	ian!11	11#2@1!2+1@2#2	6/C:x+
5	un^t-ian=h=A@2	1/A:ch	un-11	11=2/B:t	ian!11	11#2@1!2+1@2#2	6/C:x+
6	t^ian-h+A=q@1	2/A:t	ian-11	11=2/B:x	iang!14	14#2@1!1+1@2#3	5/C:q+e
7	ian^h-A+q=e@2	1/A:t	ian-11	11=2/B:x	iang!14	14#2@1!1+1@2#3	5/C:q+e
8	h^A-q+e=x@1	2/A:x	iang-14	14=2/B:q	e!14	14#2@1!1+1@2#4	4/C:x+iao-1
9	A^q-e+x=iao@2	1/A:x	iang-14	14=2/B:q	e!14	14#2@1!1+1@2#4	4/C:x+iao
10	q^e-x+iao=q@1	2/A:q	e-14	14=2/B:x	iao!13	13#2@1!3+1@3#5	3/C:q+u-1
11	e^x-iao=q+u@2	1/A:q	e-14	14=2/B:x	iao!13	13#2@1!3+1@3#5	3/C:q+u-1
12	x^iao-q+u=n@1	2/A:x	iao-13	13=2/B:q	u!11	11#2@1!2+1@2#6	2/C:n+ian
13	iao^q-u+n=iang@2	1/A:x	iao-13	13=2/B:q	u!11	11#2@1!2+1@2#6	2/C:n+
14	q^u-n+iang=pau@1	2/A:q	u-11	11=2/B:n	iang!15	15#2@1!1+1@3#1	7/C:h+
15	u^n-iang=pau=h@2	1/A:q	u-11	11=2/B:n	iang!15	15#2@1!1+1@3#1	7/C:h+
16	n^iang-pau+h=ua@x	x/A:n	iang-15	15=2/B:xx	xx!xx	xx#xx@xx!xx+xx@xx#	
17	iang^pau-h+ua=zh@1	2/A:n	iang-15	15=2/B:h	ua!11	11#2@1!4+1@5#1	7/C:

图3 部分上下文相关标注的测试结果

5 结束语

本文对输入的汉语语句,先后经过文本规范、语法分析、韵律边界预测3个过程的处理,实现了普通话的字音转换,并得到普通话合成基元的声韵母层、音节层、词层、韵律词层、韵律短语层以及语句层的上下文相关信息。通过对比普通话和兰州方言的声韵母,定义一套标记兰州方言特有发音的标音符号 SAMPA-LZ,总结了普通话到兰州方言的声韵母转换规则,并利用转换规则将普通话的声韵母转换成兰州方言的声韵母,从而实现了方言的字音转换。在此基础上,利用兰州方言的声韵母和文本分析、韵律预测获得的上下文信息,产生了用于语音合成的上下文相关标注。同时,利用上下文相关的信息设计一套用于声学模型聚类的决策树问题集。

参考文献

- [1] Chu Min, Lu Shinan. A Text-to-Speech System with High Intelligibility and High Naturalness for Chinese[J]. Chinese Journal of Acoustics, 1996, 15(1): 81-90.
- [2] Boulard H, Dines J, Majimai-Doss M, et al. Current Trends in Multilingual Speech Processing[J]. Sadhana, 2011, 36(5): 885-915.
- [3] Yang Hongwu, Keiichiro O, Gan Zhenye, et al. Realizing Tibetan Speech Synthesis by Speaker Adaptive Training[C]//

Proceedings of Signal and Information Processing Association Annual Summit and Conference. Washington D. C., USA: IEEE Press, 2013: 1-4.

- [4] 李晓红. 面向语音合成的文本处理技术的改进[D]. 北京: 北京交通大学, 2010.
- [5] 姚金国, 代志龙. 基于文本分析的知识获取系统设计与实现[J]. 计算机工程, 2011, 37(2): 157-159.
- [6] 陶建华, 蔡莲红, 赵晟. 汉语语音合成中的文本分析和韵律处理[C]//中国中文信息学会20周年学术会议论文集. 北京: 清华大学出版社, 2001: 272-279.
- [7] 陈志刚. 中文语音合成系统中文本分析的若干关键技术[D]. 合肥: 中国科学技术大学, 2003.
- [8] 索南扎西. 藏语语音合成关键技术研究[D]. 拉萨: 西藏大学, 2011.
- [9] 高璐, 陈琪, 李永宏, 等. 藏语语音合成中文本分析的若干问题的研究[J]. 西北民族大学学报: 自然科学版, 2010, 31(2): 27-33.
- [10] 马欢, 吾守尔·斯拉木. 维吾尔语语音转换系统文本分析模块初探[J]. 计算机工程, 2006, 32(16): 267-268.
- [11] 姚金国, 代志龙. 基于 HCSIPA 的中英文混合语音合成[J]. 计算机工程, 2013, 39(4): 14-17.
- [12] Pan Nenghuang, Yu Mingshi, Tsai Z. A Chinese to Taiwanese Text-to-Speech System[J]. Communications of Institute of Information and Computing Machinery, 2008, 11(4): 27-38.
- [13] 李明, 蔡莲红, 李勇, 等. 普通话与聊城话的声学特征对比及转换[C]//第7届中国语音学学术会议暨语音学前沿问题国际论坛论文集. 北京: 北京大学出版社, 2006: 1-4.
- [14] 贾珈, 蔡莲红, 李明, 等. 汉语普通话与沈阳方言转换的研究[J]. 清华大学学报: 自然科学版, 2009, 49(S1): 1309-1315.
- [15] 王兵, 苏恩泽. 天津话语音合成系统[J]. 计算技术与自动化, 1995, 14(4): 37-39.
- [16] Guo Weitong, Yang Hongwu, Pei Dong, et al. Prosody Conversion of Chinese Northwest Mandarin Dialect Based on Five Degree Tone Model[J]. JDCTA: International Journal of Digital Content Technology and Its Applications, 2012, 6(17): 323-332.
- [17] Zen Hega, Tokuda K, Black A. Statistical Parametric Speech Synthesis[J]. Speech Communication, 2009, 51(11): 1039-1064.
- [18] 张家骥. 汉语普通话机读音标 SAMPA-SC[J]. 声学学报, 2009, 34(1): 81-86.
- [19] 贾玉祥, 黄德智, 刘武. 中文语音合成中的文本正则化研究[J]. 中文信息学报, 2008, 22(5): 45-51.
- [20] 杨鸿武, 朱玲. 基于句法特征的汉语韵律边界预测[J]. 西北师范大学学报: 自然科学版, 2013, 49(1): 41-45.

编辑 刘冰