Chapter 4

Final Consonant and Tone Classification

As we know, the different final consonants have different phonetic characteristics and the final consonants is tone critical part. So the effects of final consonants on tone give more variations to the tone features and make the tone classification more difficult. In this chapter, we are trying to observe the tone classification for different final consonants in order to find the further way for improving the tone classification performance.

Here first the phonetics knowledge of final consonant for Thai language is introduced. Then the experiment framework for proposed method is presented. After that it's the experiment's results and discussions. Finally the summary for this chapter is concluded.

4.1   Final Consonant of Thai language

In Thai language the final consonants include: plosive final consonants that are p,t,k; Nasal final consonants that are m,n,ng; Fricative final consonants that are f,s; Affricate final consonant that is ch; Approximant consonants that are j,w; lateral approximant final consonant that is l. In our speech data, we have 5726 tones totally. Among these tones, the tones that have no final consonant are 2084. The tones that have nasal consonants are 1913. The tones that have plosive consonants are 863. The tones that have approximant consonants are 836. The tones that have fricative consonants are 24. The tones that have affricate consonant and lateral appoximant consonant are 4. It's shown in table 4-1:

Table 4-1 Distribution of final consonants in speech database

Vowel	2084	36.3954%
Vowel+m,n,ng	1913	33.4090%
Vowel+p,t,k	863	15.0716%
Vowel+j,w	836	14.6%
Vowel+f,s	26	0.4541%
Vowel+ch,l	4	0.0699%
	5726	100.00%

From Table 4-1, we found that speech data mainly lie in 4 goups, i.e. no final consonant, nasal final consonant, plosive final consonant, approximant final consonant. About 70% of them are the tone with no final consonant and the tone with nasal consonant. Among these four groups, nasal and approximant final consonants are voiced and plosive final consonants are unvoiced. Then it can be further grouped into 3 groups that are no final consonant, voiced final consonant, voiceless final consonant. Also as we know, the tone information mainly exists in voiced sound. So for voiceless final consonant, it doesn't have big effects on tone features. Then we can group final consonant into 2 groups, i.e. no final consonant and voiceless final consonant, voiced final consonants.

4.2   Final Consonant and Tone classification

Here we use the tone classification system is post-system of alphabet recognition. So alphabet information of the syllable is already known before tone classification. Then the tone feature is chosen as the input for different NN according to the alphabet information. We believe that the different phonetic characteristics among alphabet make the tone different in a way. The framework we proposed here is shown in Figure 4-1:
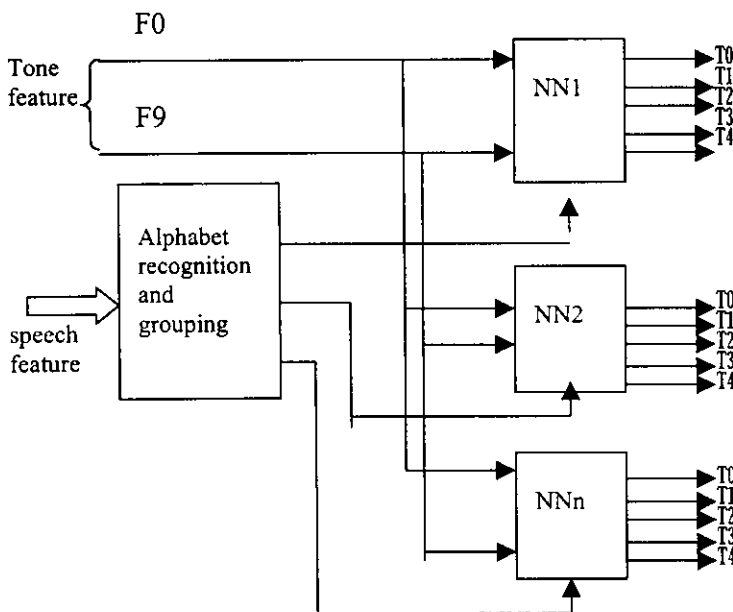
Figure 4-1   Framework of Tone classification using alphabet information

First the speech feature will be extracted from speech file. Then the speech recognition system will recognize the exact alphabet of speech. After that the specific NN is chosen for tone classification according to the recognized alphabet information.

NN1, NN2, NNn are NN tone classifier trained with specific tone data of grouped final consonants.

4.3 Experiments and Discussions

According to the proposed framework, we have done three experiments depending on the different grouping of final consonant. First we group final consonant according to the voiced and voiceless property of final consonant. Then we group them according to the distribution of final consonant in Thai language.

Expr. 1: Voiced and voiceless final consonant

As we discussed in section 4.1, based on the distribution and phonetics knowledge of the final consonant, first we group the speech data into 2 groups: one is the tone with no final consonant, plosive final consonant, other final consonants. Another is the tone with nasal final consonant and approximant final consonant. This grouping is based on the voiced and voiceless property of final consonants. As we know, the tone information is mainly lying on the vowel part and exists in voiced sound. So here the first group mainly includes the vowel with no final consonant and vowel with voiceless final consonant. The second group includes the vowel with voiced final consonant. The testing results are:

Table 4- 2 Tone Classification with Two sub-classifier
based on voiced and voiceless of final consonants

Group	Classificaiton
Vowel & Vowel+p,t,k,f,s,ch,l	69%(443/642)
Vowel+m,n,ng,j,w	82.01%(474/578)
Total	75.16%(917/1220)
	72.21%(881/1220)

Expr. 2: Grouping according to the distribution of final consonant.

From the Table 4-1, we found that about 70% tone data lie in those with no final consonants and those with nasal final consonant. We grouped the speech data into 3 groups with different final consonants: first is the tone with no final consonant. Second is the tone with nasal consonants. Third is the tone with following final consonant. The testing results are:

Table 4-3 Tone Classification with Three sub-classifier
based on distributions of final consonants

Group	Classificaiton
Vowel	67.37%(286/429)
Vowel+m,n,ng	83.02%(313/377)
Vowel+p,t,k,j,w,f,s,ch,l	77.29%(320/414)
Total	75.57% (919/1220)
	75.16%(917/1220)

Expr. 3: Further grouping according to distribution of final consonant.

In this experiment, we further group the final consonant of last group in experiment 2 into another 2 groups that is plosive final consonant and the following final consonant. Thus first two groups are the same as the last one. The third group is the tone with the plosive final consonant. And the following are the final group. The testing results are:

Table 4-4 Tone Classification with four sub-classifier
based on final consonants

Group	Classificaiton
Vowel	67.37%(286/429)
Vowel+m,n,ng	83.02%(313/377)
Vowel+p,t,k	79.7%(161/202)
Vowel+j,w,f,s,ch,l	83.96%(178/212)
Total	77.13%(938/1220)
	72.21%(881/1220)

From all of above experiments, the best performance is 77.13% that is 938 from 1220. We can found that the performance for tone with no final consonant always give us the poorest performance comparing with those that have final consonants. So we conclude that the final consonant is helpful for the stable of pitch contour. For the tones that have final consonants, the different final consonant give us different performance, such as: the tones with nasal final consonant and with approximant final consonants give the performance as high as 83%, the tones with plosive final

consonants give 79% performance. Then we conclude that the different final consonants have the different effects on pitch contour.

4.4 Summary

In this Chapter, the experiments that using final consonant for tone classification are analyzed and discussed. Through the experiments, we conclude that the tones that have final consonant are more stable than those that only have vowel part. Different final consonants have different effects on pitch contour. But the specific effects for each group of final consonants still need the further observation and research. In this method, we use the final consonant knowledge that depends on the output of former recognition system. That said even though we got higher performance here, the total performance still needs to multiply the performance of alphabet recognition. This made the improved performance is not validated for the whole system. But what we think is the synthesis system can use this information for synthesize the tone differently according to the different effects of different final consonant. Of course, in order to using this phonetics knowledge in real systems, the statistics data need to be studied and provided.