

Chapter 2

Methodology

2.1 Overview of the Experiments

Three experiments were conducted in this study. All of them examined the contour type of T2. Experiment 1 explored the contour type of T2 with different syllable vowels in isolated words. Experiment 2 investigated the final T2 in disyllabic expressions with different tones in the initial. Because tonal variations not only occur within phrase level, two more experiments were held. Experiment 3 examined the final T2 of a disyllabic word in sentence-initial and sentence-final positions.⁷ All experiments are compared between two groups: TM monolinguals and Tw-TM bilinguals.

2.2 Subjects

Twenty-four subjects aged 17 -27 participated in this study, with twelve representing TM monolinguals and the other twelve Tw-TM bilinguals. Each group contains six males and six females. All of the twenty-four subjects are students mostly studying at universities. Only one student is in high school.

Within the first group, **the TM monolinguals**, almost all were born and raised

⁷ Originally, there was another experiment (Exp.4) which investigated the effect of phrase boundaries. Unfortunately, it was not analyzed due to the time limitation of the researcher.

in Taipei or the northern counties of Taiwan. Taiwan Mandarin⁸(TM), which is the participants' everyday language, is also their first language and the primary language spoken at home between them and their parents. Besides Mandarin, they can all speak English; some of them even know a little Japanese. Furthermore, most of them cannot speak Taiwanese⁹. Even though some can understand Taiwanese, they seldom speak it or use it to communicate.¹⁰

These participants are defined as TM monolingual speakers in this study by the following criteria (from the most severe one to the most loosed respectively).

1. The subjects hardly use Taiwanese in daily life, no matter whether they understand Taiwanese or not. The percentage of Taiwanese in their daily language use is restricted within 10%.
2. Either one or both of the subject's parents are TM monolinguals.

In the second group, **the Tw-TM bilinguals**, the subjects' mother tongue is Taiwanese (Tw). At home, they usually speak Tw to their parents; in daily life, i.e. at school, half of them speak 50% TM and 50% Tw while the other half speak TM more than Tw. Due to this bilingual usage, these participants are considered as the Tw-TM bilinguals in this study. All of the twelve subjects can speak Taiwanese, Mandarin

⁸ Since these subjects were all born in Taiwan, many of them do not have the Neutral tone contrast spoken mainly by the people from the northern region of Mainland China.

⁹ Those cannot speak Tw mostly because their parents do not speak Tw. Some of their parents do speak Tw but seldom talk with them in Tw. These subjects can perceive and understand TW but can not produce it well.

¹⁰ Among the 12 monolinguals, 3 cannot understand Taiwanese, 2 understand it but cannot speak it, the remaining 7 can understand Tw and speak limited phrases, but they seldom use it to communicate.

Chinese and English. Some of them can speak a little Japanese as well. Their parents are all of Taiwanese origin and are able to speak both Tw and TM but use the former most often. The following criteria were used to define the category of “TM-Tw bilinguals” from the strictest one to the loosest one respectively.

1. The subject’s mother tongue is Taiwanese or both TM and Tw.
2. The mother tongue of the subject’s parents is Taiwanese. No other dialects of Mandarin are used.
3. Their usage percentage of Taiwanese in the daily life is higher than 10% and at home is higher than 40%. Some of them speak 90% Taiwanese in their daily life.

Within the Tw-TM group, regional differences are designed. Four are from the northern part of Taiwan (Taipei), four from the central (Taichung) and four from the south¹¹ (Chiayi, Kaohsiung and Pingtung). Each four has 2 males and 2 females. The criteria for the definition of which region they represent are as follows.

1. The subject was raised from the northern, the central, or the southern region of Taiwan, and has not moved to other region of Taiwan for more than two years. Take for example the subject B_M_Feng in Appendix 2. He was born and raised within the southern region (born in Pingtung and educated in Pingtung and Kaohsiung). Though living in Taichung now, he has lived there shorter than two years, so he is still considered a southern subject.

¹¹ The northern refers to the regions above Miaoli; the central refers to Taichung, Changhua and Nantou; the southern ranged from Yunling to Pingtung.

2. The subject might be born at other regions before, but moved to the current region at his or her earlier ages, i.e preschool ages¹². For example, B_F_Landy was considered a subject from the central region. She was born in Taipei, but has moved to Taichung before she was 5, and has lived in Taichung for over 13 years.

2.3 Variables

There may be many reasons causing tonal variations. Previous studies on tonal variations and the researcher's pilot study on TM T2 provide different aspects of possible reasons. Hence in this study two kinds of variables were explored: sociolinguistic variables and linguistic variables.

2.3.1 Sociolinguistic variables

Considering the hypothesis that more tonal variations produced by bilinguals (Lo, 2004; H. J. Hsu, 2004), the regional difference of TM T2 (H. J. Hsu, 2004) and the general view towards the gender difference, three sociolinguistic variables were explored: language background, gender and region.

2.3.1.1 Language background

There are two levels of language backgrounds: TM monolinguals and Tw-TM bilinguals. Lo (2004) and H. J. Hsu (2004) all found a sentence-final T2 variation. All the subjects they investigated are Tw-TM bilinguals. It shows a tendency that those bilinguals may produce tonal variation in certain situations. But no study covered the

¹² The detailed information of the bilingual subjects can be seen in Appendix 2.

production of the TM monolinguals in Taiwan. Hence a comparison between two language backgrounds was made in this study.

2.3.1.2 Gender

Females are always viewed as the better speaker than males with regard to pronunciation. They may produce more correct form or speak more standard language than males. Previous studies on Mandarin tones (Fon, 1997; Lo 2004) only took one gender as subjects, which neglect the difference between genders. With regard to pronunciation, we would like to know if tonal variation correlates with gender difference, or the variation has become an overall tendency in spite of gender difference.

2.3.1.3 Regions

H. J. Hsu (2004) has explored T2 and T3 from two regions of Taiwan: Taipei (the capital of Taiwan) and Taichung (a city in central Taiwan). She found that Taichung bilinguals produce a falling contour of T2 in the sentence-final position. While in the same position, Taipei speakers perform the normal contour of T2. H. J. Hsu's finding outlines the effect of region on tonal variation. Fu (1999) also found a special T3 variation produced by some of the Taichung speakers. These findings suggest that there may be regional difference in tonal variation. In this study, we divide the region of Taiwan into three parts: the northern, the central and the southern

parts of Taiwan¹³ and from each sampled four residents as subjects.

2.3.2 Linguistic variables

Three linguistic variables were explored: vowel qualities, tonal environments (the preceding tones) and sentence positions.

2.3.2.1 Vowel qualities

Three levels of vowel quality are used as an independent variable in this study--the three extreme positions in the oral cavity: [i], [u] and [a]. In an acoustic study, Ho (1976) has investigated the relation between Mandarin tones and sentence positions, syllable vowels and preceding consonants. She found the influence of syllable vowels is comparatively lower than sentence positions. With different syllable vowels,¹⁴ the four tones remain the basic tonal shapes, but differ in duration and the absolute F0 position. That is, the syllable vowels do not affect the contour, but to some extent the register. Moreover, high vowels, such as [i] and [u] usually have higher fundamental frequency than mid and low vowels. Considering the effect of the syllable vowels, we like to confirm if syllable vowels systematically affect T2 variation in this study. Or, they only influence to a minor degree as the previous study found.

2.3.2.2 Tonal environments

Han & Kim (1974) investigated Vietnamese disyllabic utterances and found

¹³ The definitions of the northern, central and southern parts of Taiwan were described in 2.2 subjects.

¹⁴ Mandarin vowels: [i], [y], [u], [e], [ɤ], [o] and [a].

that tones in different tonal environments vary phonetically both in the contour and overall pitch height. Potisuk et al. (1997) also found that the height and shape of a tone may be altered by adjacent tones in trisyllabic sequences of Thai. From those studies, we know that tonal contexts play an important role on the tonal variation. Therefore, by examining the tonal production in different tonal contexts, we can get a further understanding the reason of tonal variations.

We have investigated four levels of tonal environments: the four lexical tones in Mandarin Chinese. The tonal environment is set only on the preceding position of the target tone in this study, because TM T2 was mostly found to have variations at the phrase-final (including sentence-final) position.

2.3.2.3 Sentence positions

Positional effects on tonal variations have been discussed by many researchers (Ho, 1976; Peng, 1997; Fon & H. J. Hsu, 2004; Lo, 2004; H. J. Hsu, 2004). Ho (1976) also found that on the acoustic variations of Mandarin tones, sentence environment affects more than syllable vowels and preceding consonants. Peng (1997) has investigated Taiwanese tones in sentence-final, phrase-final and non-final positions. She found that the “final lowering” of a tone’s pitch occurs in both sentence-final and phrase-final positions and has the strongest effect in sentence-final positions¹⁵. Fon & H. J. Hsu (2004) have also found the positional effect on TM tones. They found that the overall pitch height of T2 and T3 show the highest in sentence-initial position, but

¹⁵ In Peng’s study, she uses “utterance-final” position instead of “sentence-final” position. But the former can be replaced by the later in this study.

are lowered in sentence-final positions and are also lowered to a smaller extent in sentence-medial positions. Although the pitch height is influenced by sentence positions, the shape of a tone is intact in the above studies. In addition to the sentence effect on the pitch height, the sentence effects on pitch contours of TM T2 were also found by Lo (2004) and H. J. Hsu (2004). Lo (2004) has proposed that certain southern-Min speakers¹⁶ perform a falling T2 in the sentence-final or utterance-final positions. H. J. Hsu (2004) also found a falling T2 in sentence-final positions performed by certain TM-Tw bilinguals in Taiwan.

In sum, Lo (2004) and H. J. Hsu (2004) found a falling pattern of T2 at sentence-final and phrase-final position. They also found that the initial position seldom has tonal variations, but neglect the medial position. In these regards, the present study mainly focuses on the variation pattern at the final position (phrasal and sentential), and compare the production at sentence-medial and sentence-final positions. Hence, two levels: medial and final sentence positions were investigated as an independent variable.

2.4 Materials

Three experiments were designed and presented to subjects in two PowerPoint files, with E1 and E2 in one file and E3 in separate files. Each file started with a brief introduction by the reminding slide and followed by the stimuli. All experiments contained practice items, test items and controlled items in semi-random order on the

¹⁶ In this study, these speakers are defined as TM-Tw bilinguals.

PowerPoint slides¹⁷. Each slide contained only one item and lasts for 2- 5 seconds depending on the number of syllables in the test item. Subjects were asked to read out the materials on the PowerPoint slides in normal speed. A 2-5-second break took place after every 4-7 slides. An alarm bell rang in every interval. The detailed design of each experiment is described below.

2.4.1 Experiment 1: T2 in isolated words

In experiment 1, 18 test items, 17 fillers and 3 practice items were used. The test items are 18 monosyllabic T2 words. All of them are of CV structure composed of a consonantal onset and a vowel. The onsets are all obstruent consonants [t], [t^h], [tʂ^h] and [tʂ] and the composing vowels are the high front vowel [i], the high back [u] and the low vowel [a]. They are combined into 6 syllables [tu], [t^hu], [tʂi], [t^hi], [ta] and [tʂ^ha]. Each syllable is repeated 3 times in 3 different characters. Therefore, there are a total of 18 test words. The seventeen fillers are segmentally identical to the 6 test syllables but different in tones, i.e. T1, T3, and T4. The practice items are randomly sampled monosyllabic words, which are distinct from the test items and the fillers. For the actual tonal stimulus pairs used, please consult Appendix 4.

Unlike other structure with a coda, CV structure is the one in which it is easier to observe the F0 contour because it can avoid the pitch perturbation caused by the voiced codas. Obstruent onsets were chosen because they can combine with more vowels to form CV syllables in Mandarin. Phonotactic effects are ignored in the

¹⁷ The examples of the PowerPoint slide can be seen in Appendix 3.

experiment because consonant onsets were proved to have little influence on the realization of tones (Fon & H. J. Hsu, 2004).

2.4.2 Experiment 2: Final T2 in disyllabic expressions

The second experiment aimed to examine a group of disyllabic phrases with target T2 syllables in different tonal contexts. The target T2 syllables are the same as in Experiment 1, but are now following different tones as they are combined with another syllable to form meaningful disyllabic expressions. There were altogether 32 [4 (different adjacent tones) x 8 (phrases)] stimuli and 16 fillers recorded. The fillers were disyllabic phrases composed of the other possible combinations of some other possible combinations of tones. The material is presented in Appendix 5.

2.4.3 Experiment 3: Final T2 of disyllabic expressions in different sentence positions

Experiment 3 examined the production of T2 in sentence-medial and sentence-final positions. Two carrier sentences with different blank positions were designed. They are (1) the initial position “___ jiu shi zhe yang” (___ __ is like this) and (2) the final position “zhe jiu shi ___” (This is ___ __). The same thirty-two test items used in Experiment 2 were placed in two carriers. Since the target word (T2 word) occurs at the final position of each test phrase, it now appears in either sentence-medial or sentence-final positions.

Fillers for this experiment, same as those used in E2, appear in different kind of carriers. The carrier sentences for fillers are similar with those for test items. They are

(1) “__ __ jiu shi zhe ge yi si” (__ __ is what this means), (2) “__ __ jiu shi zhe ge” (__ __ is this one) and (3) “__ __ jiu shi na yang” (__ __ is like that). These fillers were designed in this way to make the filler sentences sound more natural. Furthermore, there were also practice sentences. They were arranged at the beginning of the experiment, with the same carriers as the test sentences, but with different T2-final phrases in the blank. All sentences besides practice items were semi-randomized.

All sentences were shown in PowerPoint slides. Each slide was shown in 5 seconds. Every seven slides comes a 5-second recess for subjects to rest. The whole length of the experiments was 8.35 minutes. The subjects were allowed to drink, move quietly. They could stop the procedure if words were wrongly pronounced. In total, there were 32(stimuli) x 2 (positions) = 64 target sentences and 16 filler sentences recorded. Details can be seen in Appendix 6.

2.5 Procedure

The recording took place in classroom situations. All classrooms were quiet rooms on university campuses¹⁸. There were a notebook computer on a table, a microphone in front of the computer screen, and a chair at the table. Subjects were shown the PowerPoint files once and then were asked to read out the materials on the screen, using natural intonation and speed. Two test sessions and one short interval were conducted. Session 1 contains two experiments and each of the other two

¹⁸ Four of the subjects were recorded in a quiet room at home environment due to an inevitable limitation.

contains one. Each session starts with a brief instruction, followed by some practice items. Test items and fillers were ordered semi-randomly. The recordings will be done at a sampling rate of 48 kHz. After the recording sessions, a questionnaire was given to the subject in order to investigate their language background¹⁹.

2.6 Equipment

2.6.1 Hardware

The items of the hardware used in the experiments are:

1. MSI digital recorder and Labtec desk mic524 unidirectional microphone was used for recording.
2. IBM notebook was used to play the materials, the PowerPoint slides to the subjects.

2.6.2 Software

The items of software used for analyzing the data are:

1. Goldwave sound-editing software for cutting long speech data into short ones.
2. CSL 4100 analyzing package
3. Praat program
4. Speech Station 2
5. SPSS V12.0 statistical software for analyzing the results.

¹⁹ The sample questionnaire is in Appendix 7.

2.7 Measurement and Categorization

After collecting the speech data, the way we measure the data closely related to the way we analyze the data. The following sections describe the measurement as well as the criteria we rely on to judge a normal T2 and its variations.

2.7.1 Auditory judgment

The speech data were twice listened to and judged by the researcher. When there is an indecisive tone, it was left to a second listener, who also had phonetics training before, to judge. All the target words in the experiments were assigned to three categories: normal rising, level, and low-rising²⁰. The three types were perceptually judged by considering each subject's relative tonal range and contour. The criteria of the three categories are listed below.

Normal Rising (NR): The normal rising has a clear and sharp rising contour. It sharply rises from the middle to a higher register, and sounds like the one with the traditional value of 35. In some cases, it may naturally start from the middle and slightly go down toward the low register and rise sharply to a higher register. The one starts from the middle register and then rises smoothly higher than the beginning point is also considered normal rising. The tonal values contained in this category are 324, 314, 35 and 25. Typical examples are shown in Fig. 4 and Fig. 5.

²⁰ The falling tone proposed by previous studies was found not common in this study. In addition, it always appears in a low and narrow pitch range. We have thus treated it as a level tone.

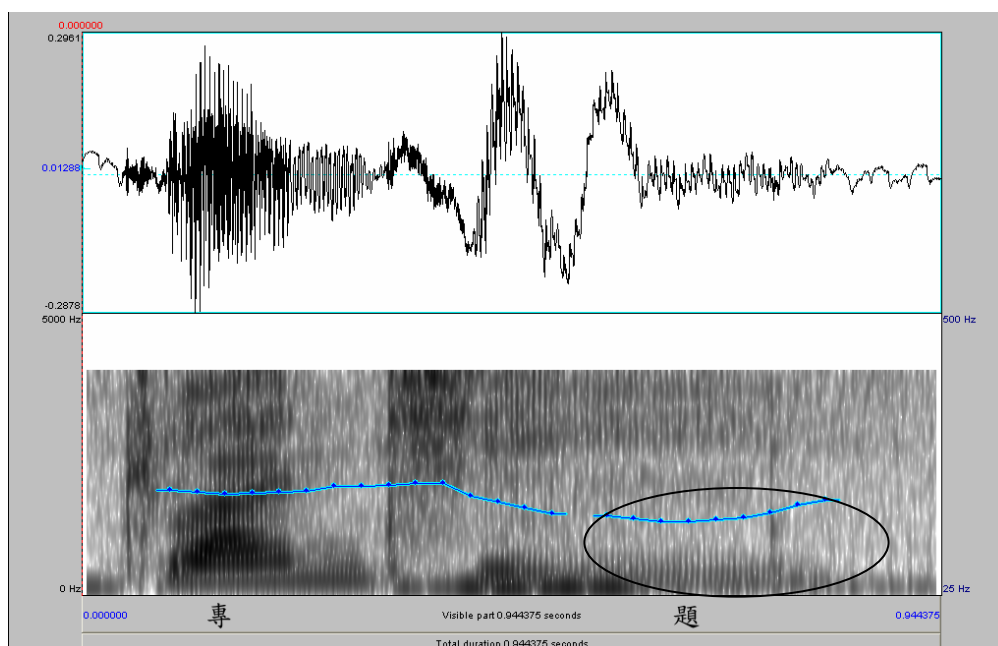


Fig. 4 The normal rising T2 :the “ti2” (題) in “zhuān1 ti2” (專題) (“project”)

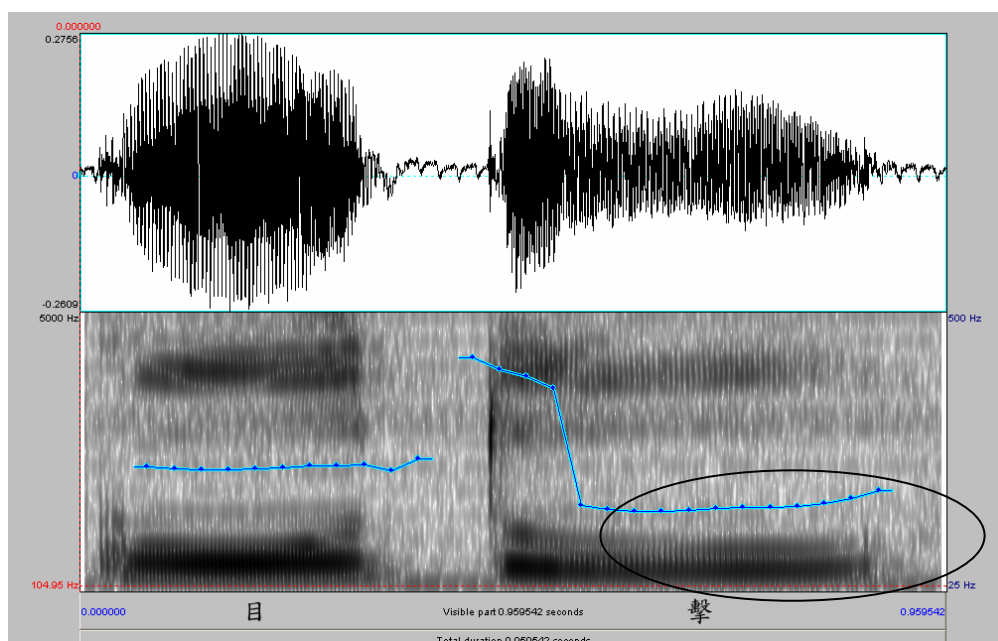


Fig. 5 The normal rising T2 : the “qi” 擊 in “mù4 qī” (目擊) (“witness”)

Level: The level variant has no clear contour but sounds flat. It is lower than the Mandarin high level tone (55) with a relative tone value such as 33, 22 or 11. But the registers are always relative, so all flat tones are considered level.²¹ Figure 6 shows a

²¹ Two subcategories: mid-level and low-level were found. Mid-level starts and ends within the middle register and low-level goes flat at the low register.

sample of mid-level tone. The shapes of low-level tone are similar with those of mid-level but appear in a lower register, as shown in Fig. 7.

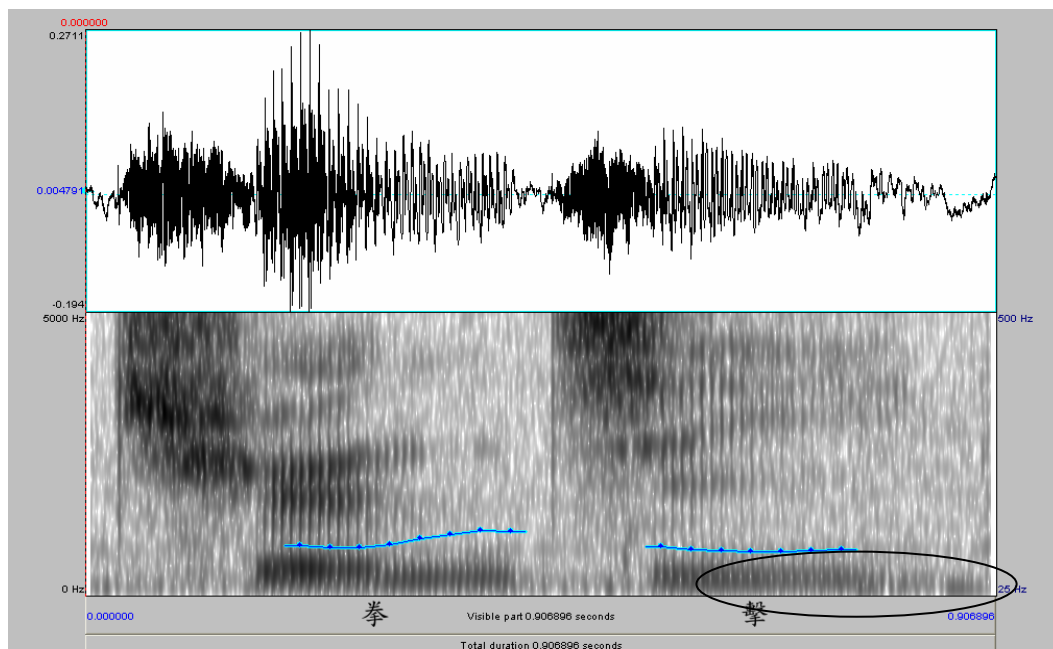


Fig. 6 The mid-level T2: the “qi2” (擊) in “quan2 qi2” (拳擊) “boxing”

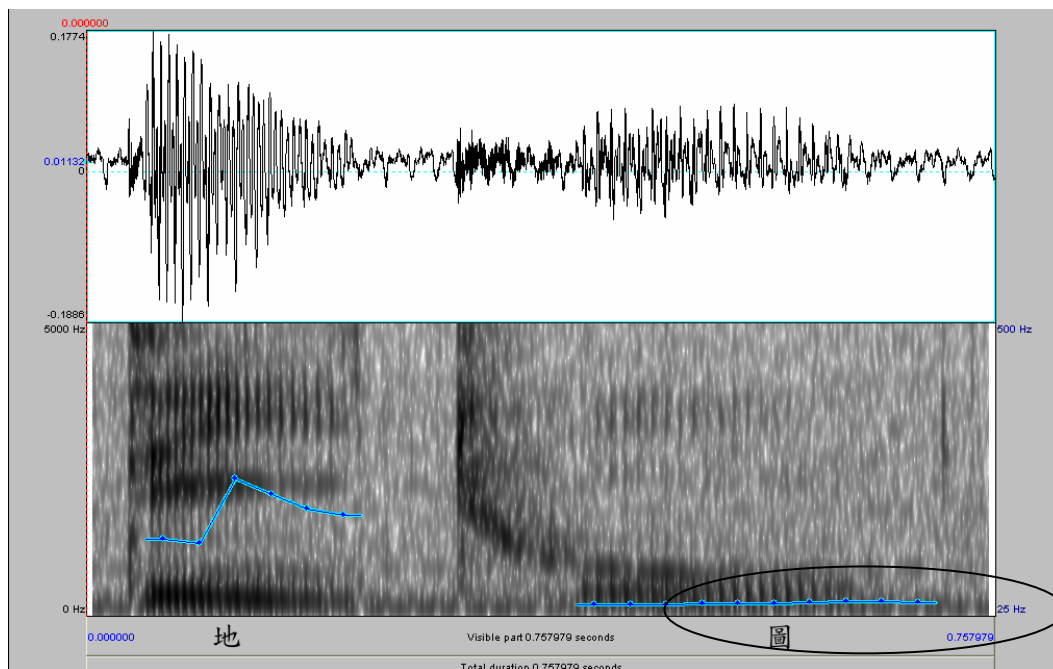


Fig. 7 The low-level T2: the “tu2” (圖) in “di4 tu2” (地圖) “map”

Low-Rising (LR): The low-rising variant rises from the low point of the tonal

range, and slightly goes up. Although the low-rising has a rising contour,” it lacks of a clear and sharp rising portion. The starting point may be at the middle or at the low point, but both rise up a little bit from the bottom of the tonal range. The tonal values contained in this category are 212, 12, 13, 312 and 213. Examples are shown in Fig. 8 and Fig. 9.

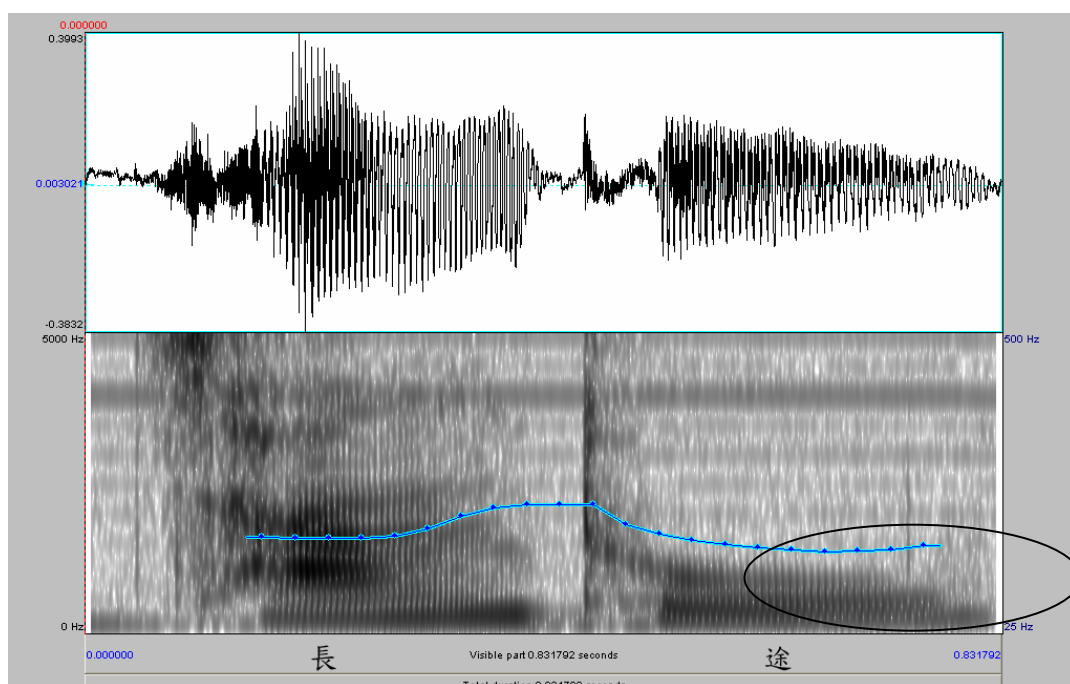


Fig. 8 The low-rising T2: the “tu2” (途) in “chang2 tui2” (長途) “long-distance”

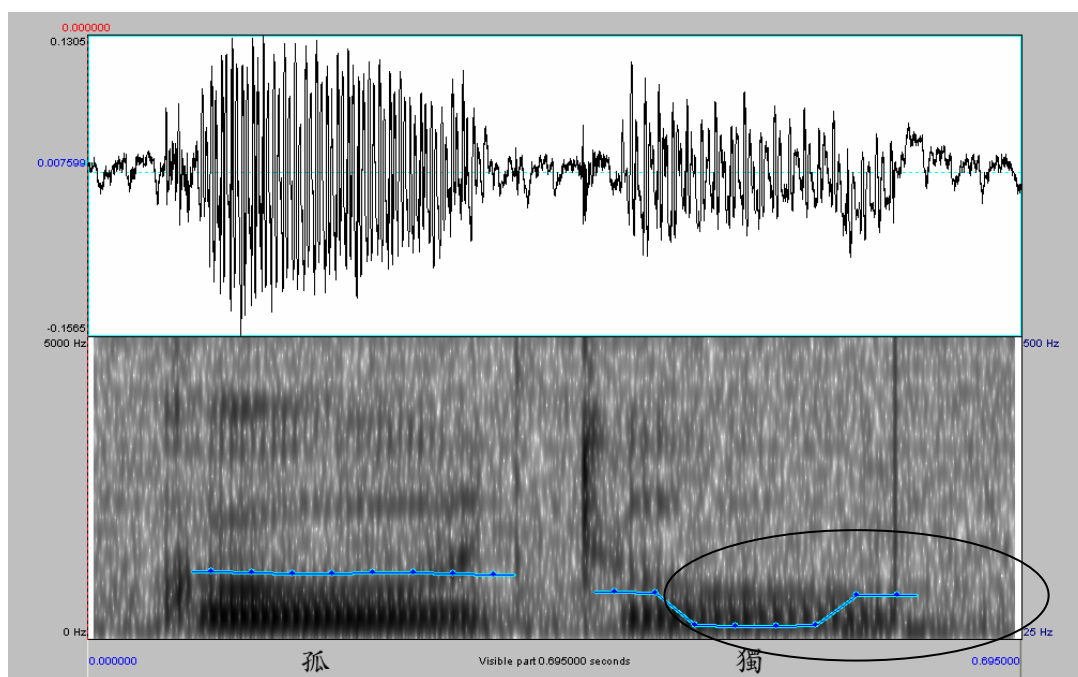


Fig. 9 The low-rising T2: the “du2” (獨) in “gu1 du2” (孤獨) “solitary”

2.7.2 Machine Reading and Reliability Test

The machine reading serves as the further confirmation of some undecided tones. Usually it was only needed when the researcher and the second listener had difficulties categorizing a tone. The machine reading can depict the computerized outcome of the pitch contour and help us to decide the categorization. We use machine reading to help categorize but do not rely entirely on it. The main analysis still depends on the auditory judgment.

After all the speech data judgments were completed, a reliability test was performed through the above procedure to test the inter-rater reliability. In the reliability test, the first grader is the researcher and the second grader is the analyzing software. When the two graders categorized a tone from the sampled data as the same, their judgments were regarded as in agreement. Thirty target tones were randomly

sampled from each categorization. Thus, a total of 90 T2 syllables were randomly examined. The agreement was high. The kappa value reaches $0.9 > 0.8$, which shows a high reliability.

2.7.3 Types of T2 variations

The categorization of T2 variations is not so easy to make. In spite of the definition of each category designated by the researcher, there are always fuzzy boundaries between each two categories. In this case, the researcher followed the auditory judgment in stead of the results of machine reading, in that human perception is more meaningful than the calculation of machine reading here. To sum up, the tonal shapes of each category are not unitary. The same auditory judgment may result in different tonal shapes. Figure 10, 11, and 12 show the possible tone shapes of each category. Each shape in a category is ranked left to right from the most typical one to least typical.



Fig. 10 The tone shapes of “normal rising” (NR) T2

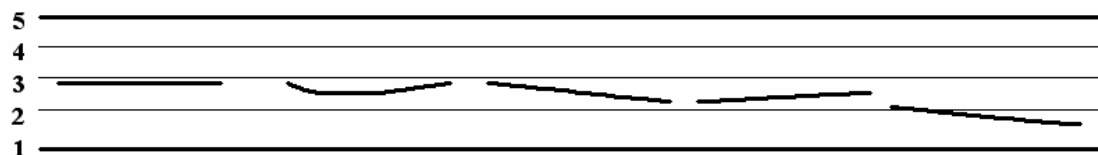


Fig. 11 The tone shapes of “mid-level” (one of the Level) T2

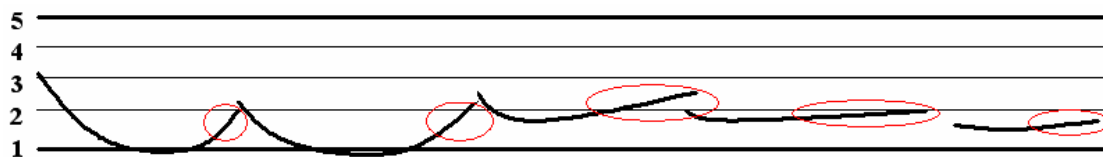


Fig. 12 The tone shapes of “low-rising” (LR) T2