Eötvös Loránd Tudományegyetem
Bölcsészettudományi Kar

DOKTORI DISSZERTÁCIÓ

YE QIUYUE

Pronunciation Problems of Hungarian University Students Learning Chinese: Analysis, Causes and Possible Solutions

NYELVTUDOMÁNYI DOKTORI ISKOLA
DR. BÁRDOSI VILMOS (DOKTORI ISKOLA VEZETŐJE)

SINOLÓGIA DOKTORI PROGRAM
DR. HAMAR IMRE CSC (PROGRAM VEZETŐJE)

A BIZOTTSÁG TAGJAI ÉS TUDOMÁNYOS FOKOZATTUK:

PROF. DR. HAMAR IMRE (C.Sc)
DR. PAP MELINDA (Ph.D)
DR. KÓSA GÁBOR (Ph.D)
DR. CSIBRA ZSUZSANNA (Ph.D)
DR. SZABÓ P. SÁNDOR (Ph.D)
DR. SALÁT GERGELY (Ph.D)
DR. CSOMA MOZÉS (Ph.D)

TÉMAVEZETŐ: DR. BARTOS HUBA (Ph.D)

BUDAPEST, 2013

Table of Contents

Preface

The importance of pronunciation in learning a language is undoubted. My own experience of teaching and learning a foreign language has convinced me that the acquisition of accurate and standard pronunciation is quite possible, but that it is by no means easy. I am also of the conviction that pronunciation is one area in which natural, unconscious acquisition, though not impossible, cannot be relied upon. Particularly with older learners, pronunciation needs to be consciously taught and learned, and one of the ways in which this can be done is through the analysis and deliberate treatment of errors.

However, the teaching of pronunciation in Chinese is a vast field, and ideally one would want to investigate all of the very wide variety of different errors made by learners. Apart from errors of monosyllabic tone, there are issues involving disyllabic tones and sentence intonation patterns. And going beyond tones, there are a great many other kinds of potential pronunciation errors concerning stress and rhythm, as well as vowel errors and consonant errors that deserve investigation. In other words, research intended to provide really comprehensive assistance to the Chinese as a foreign language teaching profession would need to cover all the possible pronunciation errors at both segmental and suprasegmental levels.

Such an endeavour is clearly beyond the scope of a thesis such as this one: it was essential not only to narrow down the field to manageable proportions, but also to prioritise. Therefore, after a brief general introduction of Hungarian and Chinese, the thesis offers an in-depth comparison between the segmental and suprasegmental phonetic systems of the two languages, leading to a discussion of which of the identified differences are likely to cause problems both in terms of acquisition by learners and in terms of communicative efficiency. Applying these criteria for the purposes of selection, the comparative analysis concluded with the choice of monosyllabic tone as a focus for the rest of the thesis. Undoubtedly, the monosyllabic tone is important enough to be the main focus, since on the one hand, Chinese is a tonal language where the same syllable with different tone usually carries different meanings; thus inaccurate use of tones carries the risk of miscomprehension and communicative breakdown; on the other hand, the monosyllabic tone is the basic of disyllabic tone and suprasegmental intonation, and the latter two features cannot really be attempted until the former has been mastered.

As a result, the thesis continued with the detailed analysis and discussion of recorded data concerning learners' and native speakers' pronunciation of monosyllabic tones. This made it possible to identify, within the general field of monosyllabic tones, the particular areas that need attention.

At this point, returning to a somewhat broader perspective, the thesis surveys some of the main ideas circulating in the Chinese language-teaching profession about possible ways of teaching tones, and then offers specific advice in this respect. The thesis concludes with some possible areas for further research and with the prospects of technological solutions to the problem of how to help Hungarian learners to master Chinese tone, both inside and outside the classroom.

This thesis would not have been possible without the support and assistance from my supervisor Dr. Huba Bartos, and Dr. Imre Hamar, director of the Department of Chinese studies, Mr. Christopher Ryan, Dr. Katalin Mady, Ms. Andrea Deme, Dr. Zhu Xiaonong, Dr. Bei Xianming, my friends and my family.

I am indebted to Dr. Huba Bartos for his invaluable comments of earlier drafts of the thesis, and for giving me access to his unpublished notes on the comparison between the two phonetic systems and the categorisation of the similarities and differences according to degree of difficulty.

It is a pleasure to thank Dr. Imre Hamar for his continuous encouragement and support in my research and teaching career.

I owe my deepest gratitude to Mr. Christopher Ryan: he placed his theoretical and practical experience of language teaching at my disposal, and his help contributed a great deal to the accurate and precise expression in the thesis. His patience warmed me and encouraged me to overcome all the difficulties that occurred in the writing task.

I am grateful to Dr. Katalin Mady and Ms. Andrea Deme of the Theoretical Linguistics Institute of the Hungarian Academy of Sciences. Dr. Mady supplied the Hungarian vowel chart and Ms. Deme Andrea gave me great assistance during the recording phase.

It is an honor for me to thank those who offered help although we have never met in person. In particular, I would like to thank Dr. Zhu Xiaonong, who is a remarkable linguist working in the field of experimental phonetics, and Dr. Bei Xianming, for giving invaluable advice about the analysis of experimental data via emails.

I would like to express my gratitude to my friends and family. Their understanding and encouragement inspired me in the whole writing procedure.

1.1. Introduction

Chinese[1] has been growing more and more popular around the world. With the development of the Teaching of Chinese as a Foreign Language (TCFL) in recent years, Chinese teaching has made major advances and achieved a great deal; however the very proliferation of non-native speakers of Chinese has drawn attention to the strength of their distinct foreign accent and the importance of pronunciation is now being re-examined; indeed, in recent years, pronunciation teaching has been greatly revalued. According to the results of questionnaires distributed by the author in the course of small-scale in-house investigations conducted in ELTE and in the Catholic University of Budapest in autumn 2012, 100% of the beginner students in the survey believe that pronunciation teaching plays an important role in the foreign language learning; 15% of them even consider pronunciation to be the most important thing to learn.

As we know, pronunciation can be considered as the "material shell" of a language and is also the most external feature of a language system. What is more, pronunciation reaches beyond mere segmental and suprasegmental features. The language we use to convey messages and the pronunciation which characterizes our speech reflect our identity for others and even for ourselves: as Porter and Darvin say, "... a person's pronunciation is one expression of that person's self-image"[2] (quoted in Dalton and Seidlhofer, 1994, 7). What is more, Dalton and Seidlhofer (ibid) indicate that language is not only used to communicate, but also to express ourselves as members of a group we intend to belong to or we would like to be associated with.

If pronunciation is a reflection of our social as well as individual identity, then consequently it is an essential part of language. Zhou (2006) claims that the teaching of Chinese pronunciation is a basis of TCFL; what is more, he avers, it is the most important prerequisite of training learner's skills of listening, speaking, reading, writing and social communication. The purpose of teaching pronunciation is that learners should master the basic knowledge of

[1] In this thesis "Chinese" is used to refer to *Putonghua* or Mandarin, the variety which is generally regarded as the "standard" and is taught in most schools and universities abroad, including all those in Hungary.

[2] Porter, D. and Garvin, S. 1989.

Chinese phonetics and the correct, fluent pronunciation of Mandarin and, meanwhile, lay a foundation for oral communication. With regard to sequencing, Zhou also indicates that learning a foreign language includes three main parts: pronunciation, grammar and vocabulary, the learning order should be according to these three steps. The pronunciation part, he states, is the most difficult, and is also the most important, because pronunciation is the nature of a language. If pronunciation is incorrect, he claims, then grammar is incorrect, and vocabulary as well[3]. Sheng (2006) agrees with Zhou on this point: in his *Yuyan Jiaoxue Yuanli*, he points out that

> "…in the process of second language teaching, pronunciation is the most difficult part. This is because the adult learners have already mastered one language. If they are not given strict pronunciation training, they will easily use the pronunciation of the first language instead of the second one. The plasticity of their articulators is not as good as children's: learning a new pronunciation is more difficult for them."[4]

Actually, one of the greatest difficulties for a student in learning a foreign language lies in the acquisition of correct pronunciation. Pronunciation is probably one of the most essential components of intelligibility, and as Joanne Kenworthy (1987) writes: "foreign speakers need to be intelligible so that they can communicate"[5].

However, owing to the particularities of the pronunciation of Mandarin, the effects of pronunciation teaching cannot be perceived as immediately as those of vocabulary teaching and grammar teaching. As Bowen and Marks (1992) say, "... pronunciation has often been abandoned while other aspects of language have been paid much attention to. Many teachers are, consciously or unconsciously, reluctant to implement pronunciation teaching in their coursework, because it is regarded as difficult."[6] And subjectively, because learners may agree and make slower progress in this area, it is perhaps not as rewarding as teaching other language skills.

---

[3] Zhou 2006.

[4] Sheng 2006.

[5] Kenworthy, J. 1987: 15.

[6] Bowen, T. & Marks, J. 1992.

As a result, students, whether consciously or unconsciously, tend to shift their learning focus towards the “safer ground” of vocabulary and grammar. In this they are often encouraged, although perhaps not deliberately, by course books and indeed by teachers who focus on pronunciation only at the very beginning of courses and then concentrate on other areas.

A great number of older students have many difficulties with pronunciation, either when trying to make themselves understood or when listening to others. Nevertheless, the necessity and the method of pronunciation teaching is still a controversial issue. The amount of time devoted to pronunciation teaching seems to depend on the teacher. Although for a lot of students it becomes very important, or even a priority, pronunciation is an area that some teachers avoid or are reluctant to teach.

Teachers sometimes justify their avoidance of extensive pronunciation teaching with the argument that for CFL learners perfect pronunciation is not only extremely difficult but actually unnecessary: of course, they claim, one should try to avoid errors that will lead to ambiguity or misunderstanding, but one should be satisfied with comprehensibility rather than aiming at an impossible ideal. Indeed, for learners who only need to use Chinese for example for recreational purposes such as tourism, or for occasional business meetings, this argument would seem to be valid. World-wide, such learners probably represent the majority. On the other hand, students of university departments of Chinese studies belong to a special category: they will eventually receive a degree which qualifies them to act as “professional users of Chinese”. They may well go on to work as translators, interpreters or teachers, and because of this they must aim at a much higher degree of accuracy, in all linguistic areas, including pronunciation, than “ordinary” learners. Some of them may even need to specialise in one or more of the many Chinese dialects, but all of them should be proficient in what is internationally regarded as the acrolect: Mandarin or *Putonghua.* As we mentioned above, students at ELTE and at the Catholic University of Budapest consider learning correct pronunciation to be an important aim, and whether their motivation is purely “aesthetic” or “professional”, it seems to be well-justified. This thesis, in fact, is intended to help university departments of Chinese studies and their teaching staff to fulfil a real responsibility. At the same time, we must acknowledge that perfect, native-like pronunciation, however desirable it may be, is probably an unattainable ambition for most non-native learners of Mandarin: it would be unreasonable to demand it, especially at the elementary level where nearly all ELTE learners begin. It therefore seems necessary to prioritise: to identify aspects of pronunciation that deserve, and will repay, special attention during CFL courses. Without ignoring subjective factors (the learners' own priorities and real/perceived difficulties), we prefer to take a somewhat utilitarian approach to this issue, based on the conviction that if our principle aim is to enable CFL learners to communicate effectively, then our principle duty is to identify and eradicate potential obstacles to effective communication. In other words, we need to identify aspects of pronunciation which are not only difficult for learners to acquire and produce, but which are likely to cause confusion if they are *not* correctly produced.

Since the 1980s linguists have studied pronunciation and pronunciation teaching in TCFL in many aspects and achieved a great deal. Some of them, such as Meng Zimin, have analyzed in depth the scheme of Chinese Pinyin[7] or investigated the underlying strategies and principles of pronunciation teaching, like Zhang Hesheng (2006) and Lu Bisong[8]; others have focussed on the "rules" of pronunciation learning and teaching; these include Zhao Jinming, Yi Hongchuan, and Cheng Tang,[9]. All of this major body of work, and the achievements that it has led to, might have been expected to promote the development of pronunciation teaching. However, the effects on learners' pronunciation are still unsatisfactory. As Cheng Tang says, "We feel that the effect of pronunciation of TCFL is not very ideal. In the past several decades, among the foreign students, there are not so many students who can speak Mandarin in a standard or relatively standard way, actually they are very few."[10] Liu Xun also points out that as a result, perhaps, of the very success of TCFL, many teachers and linguists believe that while standards of verbal fluency have risen, the standard of pronunciation has actually been falling."[11]

---

[7] Meng 1997.
[8] Zhang 2006; Lu 2007.
[9] Zhao 1998; Yi 2001 ; Cheng 1996.
[10] Cheng (ibid).
[11] Liu 2000.

Moving on to the current situation of research into TCFL, the error analysis of pronunciation is even weaker. Recently, there have been some monographs and theses in related fields, but for the most part, they involve error analyses of grammar or vocabulary, rather than of pronunciation. Although learners from different countries have some pronunciation errors in common, specifically, learners with different mother tongues have different interlanguage systems, and therefore have different pronunciation errors. So far, linguists have studied Chinese pronunciation of learners whose mother tongues are English, French, Japanese and Korean and so on. As to the specific case of Hungarian students who are learning Chinese, there has been relatively little systematic error analysis of their Chinese pronunciation.

1.2. The present situation of Chinese pronunciation teaching in Hungarian universities

So far, Eötvös Loránd University (ELTE) and Péter Pázmány Catholic University, are the only two universities in Hungary that have a Chinese department. In addition, The Confucius Institute at ELTE has set up Chinese teaching in nine schools. Since the department of Chinese of ELTE is the oldest and the most professional Chinese teaching institution with the most Chinese learners in Hungary, this department of Chinese is taken as the research object for this thesis. Through the analysis of the current situation of Chinese pronunciation teaching in the department of Chinese of ELTE, the current situation of pronunciation teaching in Hungarian universities is reflected.

Founded in 1924, the Department of Chinese of ELTE was originally aimed at cultivating sinologists, but has developed a double focus including both Sinology research and Chinese language teaching. Graduate students and doctoral students usually take Sinology as their research direction while the Chinese language teaching is mainly conducted for undergraduates. At present, there are 116 undergraduates, including 58 freshmen, 35 sophomores and 23 junior students. Ever since 1984, there have been Chinese teachers sponsored by the Chinese government working in the department of Chinese. In theory, native-speaker teachers from China are expected to ensure that students in the department of Chinese have the possibility to acquire standard Chinese pronunciation. However, due to the limitations of the set curriculum, teaching methods and teaching materials, Chinese pronunciation teaching lacks priority and orderliness and the teaching effect is inevitably unsatisfactory. The main problems existing in the teaching of Chinese pronunciation are as follows.

1.2.1. Concerning Teaching Arrangements

(1) Pronunciation teaching has not been duly valued in the Chinese teaching. In the teaching process prevailing in ELTE, and most probably in other universities, class hours for training pronunciation are relatively few and are unlikely to be enough to enable students to form deep and lasting impressions of Chinese pronunciation in their minds, or for these impressions to become firm habits. At the primary stage of students' learning Chinese, although specialized pronunciation teaching is provided, the training is not comprehensive, since the training time is relatively short, only lasting for two or three weeks (see 1.2.2. "About teaching materials" below).

Mackey (1965) points out that before students' practice and consolidate their oral expression ability, it is quite important to enable them to form correct language pronunciation habits from the very beginning, since every word learned by students will deepen their pronunciation habits and it will be hard and time-consuming to correct the wrong pronunciation habits once they are formed[12]. Weak Chinese pronunciation foundations are bound to affect students' Chinese learning and are likely to reduce the effectiveness of their spoken Chinese.

(2) Requirements for pronunciation teaching are high; probably too high for the relatively short dedicated teaching time lasting only two or three weeks at the very beginning of the course. As Zhao (1985) pointed out, "From the arrangements of pronunciation knowledge in most textbooks, it can be seen that the pronunciation teaching has high and detailed requirements, since in addition to consonants and vowels, the neutral tone, the retroflex final and tone sandhi (the 3rd tone sandhi and semi-3rd tone) are required to be mastered. Moreover, some textbooks require students to master the pronunciation changes of '一 Yi' and '不 Bu '".

---

[12] Mackey W. 1965.

“All in all, pronunciation teaching is designed to be as comprehensive and detailed as possible in terms of teaching contents.”[13]

(3) The pronunciation input and output in the teaching of Chinese are unbalanced. The teaching of pronunciation involves both receptive (input) and productive (output) elements. These two features of pronunciation play an important role in the processing and acquisition of correct pronunciation, and pronunciation training of the second language teaching should combine pronunciation input with the pronunciation output. In spite of the long-term popularity of unconscious “natural acquisition” (as opposed to conscious learning) in for example the teaching of English as a foreign language which is due to the work of such as Krashen (1981)[14], many teachers of Chinese as a foreign language would agree with Ellis (2006) [15] in arguing that Chinese pronunciation needs to be consciously learned and deliberately practiced, especially in view of the notable difference between Chinese and for example Hungarian phonetic and phonological systems (see Chapter 3 below) and that input and output should therefore be regarded as two independent processes, rather than an organically connected continuum. Late or insufficient pronunciation input will directly affect the Chinese listening capacities of learners and further affect the output of Chinese pronunciation. Therefore, if there is little Chinese pronunciation input training in the pronunciation teaching or deviations in the pronunciation input, mistaken pronunciation habits are likely to set in and to be reinforced by students. In some cases, the fossilization of pronunciation deviations and errors may appear. In the meanwhile, the insufficient pronunciation output would result in students’ lacking feedback about their own pronunciation abilities, which is not conducive to the self-improvement and correction of students in terms of their pronunciation.

(4) In ELTE Chinese teaching, little or no planning is devoted to achieving a principled balance between segmental and suprasegmental (sentence-level) phonetic training. The phoneme teaching is usually arranged (according to the teaching material in use – the *Xin*

---

[13] Zhao 1985. (my translation)

[14] Krashen, S. 1981.

[15] Ellis, N. 2006: 164-194.

Shiyong Hanyu Keben edited by Liu 2002[16] see below) to be conducted at the very beginning of the whole teaching process, lasting for two weeks in most cases, and involves the direct teaching of consonants, vowels, and tones, and pronunciation of elements within words, but not of rhythm and intonation. Suprasegmental, sentence-level pronunciation (or “language fluency”) teaching is not considered as a separate topic, although it might be expected to continue, integrated with other areas, throughout the entire teaching process. However, the textbook does not even mention it after the first few chapters, so everything depends on the individual teacher’s (and perhaps the students’) priorities and decisions; as a result this area can easily be ignored and left out altogether. While the initial phoneme teaching does divide tasks in the pronunciation teaching into several independent parts, and is therefore conducive to helping learners to establish favourable basic pronunciation habits within a relatively short time, the lack of planned suprasegmental work reveals the failure to appreciate the importance of this component. Ideally, pronunciation teaching should be in line with real communicative needs and dynamic speech flow teaching can better satisfy the demands of establishing a stable and effective pronunciation system. Compared with phoneme teaching with its clear and relatively limited connotations, speech flow teaching covers rich and extensive contents in terms of the coarticulation, phonetic changes and rhythms. Speech flow teaching is likely to need to continue for a long time as any course progresses, in contrast with phoneme teaching, which is conducted at the very beginning of the teaching process and is unlikely to continue unless remedial work is required. The writer’s own experience within the Confucius Institute and the Chinese Department of ELTE suggests that in the actual teaching process, Chinese teachers tend to attach excessively great importance to the initial phoneme teaching and ignore the significance of the speech flow teaching. Informal discussions with colleagues in the above-mentioned institutions support this impression, as does the nature of the teaching materials in use.

(5) The fact that pronunciation teaching is not normally revised and reinforced several times after the earliest stages of the whole language learning process, is not conducive to students’ learning, consolidating and improving Chinese pronunciation. Informal discussions with

[16] Liu 2006.

TCSL teachers in ELTE indicate that the “costs” of pronunciation teaching are seen as too high. Pronunciation teaching needs a great deal of time and energy on the teacher’s part. Moreover, the pronunciation teaching may be too boring to arouse students’ interest and enthusiasm. Therefore, teachers appear to have formed a consensus that pronunciation teaching is usually best conducted only, or at least mainly, by teachers working with entry-level students. After the initial stage of pronunciation teaching, teachers would basically not correct the pronunciation deviations and errors any more, unless the pronunciation affects the language expression to the extent of causing obvious communication breakdown. The instruction of pronunciation is often regarded as a separate module. In fact, as Zhao (2011) [17] points out, “learning pronunciation is part of language learning. It must be practiced all the time”.

1.2.2. Concerning Teachers

Pronunciation classes are given by native speakers, with the obvious intention of providing more contact with the “real language” and standard pronunciation. Putting native speakers in charge of pronunciation is no doubt a wise decision, because to some extent, they can create real Chinese-speaking circumstances, and in terms of the learners’ pronunciation development this may be effective from a long-term point of view, but it also has some problems:

First, as Zhao[18] indicates, learners’ strong accent mostly results from the interference of their mother tongue (in the case of learners at ELTE, mainly Hungarian), so the teacher who is in charge of pronunciation should ideally have some linguistic knowledge not only about Chinese but also about Hungarian, so that s/he can predict possible difficulties, understand them when they arise, and devise appropriate repair strategies. However, most of the Chinese teachers at ELTE have virtually no idea about Hungarian, so they cannot give a clear and penetrating explanation of the problems troubling learners using, for example, the methodology of contrastive analysis. In other words, they know there is a difference but do not know where the difference comes from. Second, there is the common phenomenon of

[17] Zhao 2011: 367. (my translation)

[18] Zhao (ibid).

“teacher-talk”: in order to make learners understand, teachers often try to slow down their speech and pronounce every word clearly, which means they are speaking in an artificial, “unreal” way. Therefore, learners are likely to get used to teachers’ delivery speed and to feel nervous when they encounter a real conversation. Third, Mao (2002)[19] says, “as a teacher of TCSL, s/he has to possess phonological and phonetic knowledge, the Chinese phonology and the rules governing its application; indeed, s/he also has to understand the rules of phonetic cognition and acquisition. But so far, the people [sic] who specialize in the field of phonetics, especially in applied phonetics are really few.” (my translation)

1.2.3. Concerning Teaching Materials

The Chinese textbooks used by the Department of Chinese of ELTE are *New Practical Chinese Textbooks* edited by Liu[20]. *New Practical Chinese Textbooks* is a series of well-established Chinese teaching materials, which are widely regarded as systematic and reliable. In this series of textbooks, the pronunciation teaching is mainly conducted in the first six lessons. From the seventh lesson, knowledge about pronunciation does not feature any more. In terms of the application of this series, the issue of how to ensure appropriate recycling, revision and continuity of pronunciation teaching is a real problem needing to be solved. With regard to the imperfect arrangement of pronunciation teaching contents in the textbooks, the renowned Chinese linguist Lin Tao once stated (quoted in Jiao, 2001)[21] that he had talked to relevant people in the Office of Chinese Language Council (HANBAN) about this problem, but the textbook compilations had not been changed: pronunciation teaching was still conducted within a short time, and as a result it could not ensure the learning effect. This problem, said Lin, is not as “boring” as it is imagined to be. Students are unwilling to learn pronunciation. Therefore, students are allowed not to learn more about pronunciation for this reason. If one wants to make pronunciation teaching less boring, he concludes, the key lies in whether teachers have the ability to compile suitable teaching materials or to be creative enough to make the textbook attractive.

---

[19] Mao 2002.
[20] Liu 2002.
[21] Jiao 2001.

1.2.4. Concerning Teaching Methods

At present, a popular and commonly used method for teaching Chinese pronunciation to adults is what could be described as the "theory first, then repeated practice" approach. In this approach to pronunciation teaching, as Zhang (1990) says, "teachers often firstly explain the articulation places, manners of articulation and show the picture showing the tongue positions for different pronunciations and the pictures showing the shape of lips and then give demonstration to lead students to practice single tones or distinguish different pronunciations [my translation]."[22] Indeed, it is quite necessary to properly explain the pronunciation principles for adults when teaching them, and the learning, consolidation and if necessary correction of pronunciation on the basis of a proper understanding of pronunciation principles can usually achieve remarkable results. On the contrary, oversimplified or missing explanations of the relevant principles, coupled with a relatively short training time, will inevitably make it hard for students to internalize and consolidate proper pronunciations. The teachers in charge of pronunciation teaching in the ELTE Department of Chinese are from China, and very few of them speak Hungarian at more than basic level. Since more or less serious obstacles to verbal communications exist between teachers from China and Hungarian native students, the theoretical explanations are largely simplified or even simply left out in the pronunciation teaching. Some teachers try to use English for this purpose, but not always very successfully, and in any case by no means all the learners can understand that language.

The simple mouth-to-ear or "listen and repeat" teaching method, unsupported by metalinguistic explanation, that was common in the days of behaviourist, audiolingual approaches most famously supported by Skinner (1957)[23] and typified, at its best, by such English as a Foreign Language coursebooks as *Streamline Departures* (Hartley and Viney, 1978)[24] and the subsequent Streamline series, seems now to be somewhat obsolete. Even if we accept the primacy of natural acquisition, the value of conscious (or artificially encouraged) "noticing", as described by Richard Schmidt (1990) cannot be denied[25]. Nor can we ignore the research achievements gained in experimental phonetics and the extensive

---

[22] Zhang 1990: 262

[23] Skinner, B. 1957.

[24] Hartley, B. & Viney, P. 1978.

[25] Schmidt, R. 1990: 129-158.

applications of advanced technological means such as multimedia computers that have become available and which obviously lend themselves to explicit explanation of the principles of Chinese pronunciation.

Towards the pronunciation teaching approaches, scholars in the field of TCSL hold three kinds of viewpoints: phoneme teaching, speech flow teaching, the combination of phoneme and speech flow teaching (Gu, Wu, 2005)[26]. Cheng (2000) reveals what he describes as the difference between phoneme teaching and speech flow teaching. He claims that the former starts from phonemes, then moves on to syllables, words, phrases, finally reaching conversation practice, while the latter begins from sentence-level utterances and whole conversations, instead of phonemes[27]. In many cases, the teaching approach combining phoneme and speech flow teaching is chosen. The Department of Chinese of ELTE is no exception. However, since pronunciation teaching is not organized as an independent course, but is part of the spoken language learning programme, while the increase and deepening in the contents of spoken language training proceeds throughout the programme, the pronunciation teaching is all too often ignored after intensive pronunciation practice at the very beginning of the course. Experience shows that even if students can correctly pronounce every word in isolation, their foreign accent is likely to emerge, once they start speaking in full sentences, both in segmental tones and in suprasegmental intonation.

1.2.5. Concerning Learning Methods

Students naturally tend to use similar sounds in their native language to replace Chinese pronunciations which are hard to master for foreign students, which is actually caused by the powerful mother-tongue habits of students in their learning of Chinese pronunciation.

As Zhang (2001) argues that when learners can make the oral expression via the foreign language, the nonstandard pronunciation would not exert too much influence upon the success of verbal communication. Affected by the physiological inertia [habits] of the native language pronunciation, he claims, students tend to replace sounds which are hard to master by similar

[26] Gu and Wu 2005.

[27] Cheng 2000.

sounds in their native languages, thereby giving birth to different kinds of pronunciation deviations and errors. According to Zhang this would explain why it is difficult for foreign students to get rid of their foreign accents when speaking other languages[28].

1.2.6. Concerning Teachers' and Learners' Attitudes

The awareness of the importance of pronunciation among teachers of Chinese as a foreign language has fluctuated considerably, with some professionals virtually disregarding it or strongly downgrading its importance. Liu Xun (2000) believes that there is no need to require foreign learners to speak Standard Chinese, since the way most Chinese people speak cannot be regarded as standard Mandarin[29]. And Cheng Tang (2000) agrees with Liu Xun on this point. He also states that the purpose of most students' learning Chinese is to study other majors in Chinese universities (for them, the language is hardly more than a basic tool), so they themselves do not have high requirements of their own Chinese proficiency, especially with regard to Chinese pronunciation[30]. Some teachers of TCFL also think that what learners need is, above all the capacity for understanding and basic communication. As to standard pronunciation, they believe that it does not deserve a great deal of effort: Liu's opinion (2000, mentioned above) referred specifically to the non-standard nature of the pronunciation patterns typically encountered in the wide variety of dialects used by most speakers of "Chinese". "All in all, at the beginning, no matter whether we speak of teachers or learners, they do not pay enough attention to pronunciation teaching". (my translation). While this attitude is understandable, and may be valid for the majority of learners of CFL, we have already argued (see 1.1. above) that it does not apply to the particular group of learners – university students – with whom this thesis is concerned. With the development of TCFL in recent years, Chinese teaching has made great advances and achieved much; however, as we mentioned earlier, the very proliferation of non-native speakers of Chinese has made the strength of their foreign accents especially distinct and the importance of pronunciation is now being re-examined and, indeed, in recent years, pronunciation teaching has been greatly revalued. However, owing to different reasons, it would seem that many Chinese teachers unconsciously or consciously still pay relatively little attention to the teaching of pronunciation. In part, this attitude contrasts with that of their students: according to the

[28] Zhang 2001.
[29] Liu 2000.
[30] Cheng 2000.

results of questionnaires distributed by the author in the course of small-scale in-house investigations conducted in ELTE and in the Catholic University of Budapest in autumn 2012, 100% of beginner students believe that the pronunciation teaching plays an important role in the foreign language learning, 15% of them even consider pronunciation as the most important. However, owing to the particularities of the pronunciation of Mandarin, the effects of pronunciation teaching cannot be perceived as immediately as those of vocabulary teaching and grammar teaching. As a result, students also consciously or unconsciously tend to shift their learning focus towards the "safer ground" of vocabulary and grammar.

1.3. Overview of thesis

Generally speaking, this thesis is intended to explore the relevant differences and similarities between Chinese and Hungarian phonetics and to investigate the following main questions, in the light of existing linguistic descriptions and theories of second language acquisition and through the application of experimental phonetics to data collected from learners.

1. What are the principle differences between the Hungarian and Chinese phonetic systems, and in what areas are these differences likely to cause the most problems?

2. How far has one particular sample university-level learners of CFL been able to overcome these problems?

Apart from the review of the relevant literature, the principle research method used consists of analyzing and comparing recorded speech by Hungarian learners and Chinese native speakers, using one of the latest versions of the PRAAT software originally developed by Boersma & Weenink in the mid-1990s (2013)[31].

Both of the main questions listed above aim at describing and understanding the issues in focus; the thesis does not systematically evaluate attempts to solve the problem of these learners' pronunciation. However, certain implications, relevant to the problems discussed earlier in this chapter, do emerge, including possible pedagogical responses, and these are discussed in the latter part of the thesis.

---

[31] Boersma, P. & Weenink, D. 2013.

Following the present Introduction, Chapter 2 consists for the most part of a discussion of the theories of contrastive analysis, error analysis and interlanguage. After a brief general introduction of Hungarian and Chinese, Chapter 3 offers an in-depth comparison between the segmental and suprasegmental phonetic systems of the two languages, leading to a discussion of which differences are likely to cause problems both in terms of acquisition by learners and in terms of communicative efficiency, and concluding with the choice of syllable tone as a focus for the rest of the thesis. Chapter 4 describes in detail the research procedures used to verify whether and to what extent syllable tones do cause difficulties for a sample of ELTE CFL learners. These procedures consist largely of collecting audio-recordings of learners and comparing them with recordings of native Mandarin speakers through computer-based phonetic analysis, using the PRAAT software which is one of the main components of the research design. Chapter 5 presents the findings of the research in the form of comparative phonetic data on the differences and similarities between the Hungarian and the Chinese subjects, together with notes and comments on these data.

In chapter 6 the focus shifts to the practical implications of the findings for the teaching and learning of Chinese pronunciation in Hungary, mainly in terms of techniques and strategies available to teachers, but with some reference to possibilities of autonomous self-development by learners. The thesis concludes with a summary of findings and implications.

2. Theoretical Foundation of Research

2.1. The theory of contrastive analysis

2.1.1. Behaviourism

Before the introduction of Contrastive Analysis (CA), we must consider the application of behaviourist approaches to language teaching, not only because behaviourism is the core of Contrastive Analysis Hypothesis (CAH), but also because its association with behaviourism gave CA academic respectability, providing the foundation for a certain degree of popular appeal.

In the 1950s and 1960s, the prevalent language teaching techniques were based on the behaviourists' view of language. From the behaviourists' perspective, people learn by responding to external stimuli and receiving appropriate reinforcement. A proper habit is formed by reinforcement, hence learning takes place. Therefore, errors were considered to be a wrong response to the stimulus, which should be corrected immediately after they were made. Unless corrected properly, the error would become a habit and a wrong behavioural pattern would stick or "fossilise" in the learner's mind.

This view of learning greatly influenced the language classroom, where teachers concentrated on the imitation and memorization of the target forms and tried to instil the correct pattern of the form into learners' mind. If learners made any mistakes, the teachers corrected them immediately. Errors were regarded as something to be avoided at all costs and making an error was considered to be fatal to proper language learning process. The ideas of behaviourism can be best found in Skinner's *Verbal Behaviour* (1957)[32], in which Skinner transfers the research methodology used in animal behaviour studies to human verbal behaviour.

However, this belief about how learning takes place learning was eventually largely discredited by the now well-known and completely different perspective proposed by N.

[32] Skinner, B. 1957.

Chomsky (1957)[33]. He wrote in his critical review of Skinner's *Verbal Behaviour* that human learning, especially language acquisition, cannot be explained by simply starting off with a "tabula rasa" state of mind. He pointed out that children learning an L1 do not simply reproduce what they have heard; very often they use language creatively, understanding and producing utterances they have never heard before. They show evidence of internalized rules by producing forms. Therefore, he claimed that human beings must have a certain kind of innate capacity which can guide them through a vast number of sentence-generation possibilities, pointing out that with almost no exceptions all children acquire effective command of the grammar of their native language by the time they reach the age of five or six. Chomsky proposed that all human brains are "hard-wired" with certain basic linguistic principles and constraints, and that all normal humans are able to "re-create" the language of the society into which they are born, on the basis of these principles and constraints, using a natural "language acquisition device". He called this capacity "Universal Grammar" and claimed that it is this very specifically human faculty that linguistics aims to pursue.

Although Chomsky's theory gives us a new perspective concerning language *acquisition*, especially first-language acquisition by young children, there are major questions about whether the natural language acquisition device is still available or effective after early childhood. Moreover, Chomsky's theories are mainly concerned with the acquisition of syntax, and many teachers would agree that some form of behaviourism is also relevant and significant in guiding the second language *teaching and learning*, especially the teaching of pronunciation.

For one thing, Chomsky's theory may help to explain how children or babies grasp the intricate grammatical rules to some extent, but our Chinese learners are second language learners, who are learning Chinese without an effective Chinese environment and they are college students, far beyond the so-called "critical language acquisition period" posited by the Critical Period Hypothesis that became widely known in the 1960s thanks to the work of Lenneberg (1967), and which is still widely believed to limit the effective use of the natural

---

[33] Chomsky, N. 1957: 142-143.

language acquisition device[34]. Besides, the process of learning pronunciation, which requires conscious effort, is quite different from the largely unconscious process of grammar acquisition. As Chomsky himself said, grammar acquisition needs an entirely cognitive process of internalization, whether you are learning L1 or L2. However, the learning of pronunciation involves the physical movement of muscles, the shape of the mouth and the position of the tongue, a set of physical habits that seem to be largely based on stimulus or correct input, response, and positive reinforcement. In this way, the learners will form a specific set of habits. Otherwise, if the errors in pronunciation are not corrected in time, they can be easily fossilized as time passes. Although the majority of young children learn correct pronunciation without much obvious "correction", it is worth noting that they are far more likely to need speech therapy, involving the correction of wrong physical movements, in order to correct poor pronunciation, than they are to need help with poor syntax. People whose childhood pronunciation is non-standard and who are never corrected are likely to retain the "problem" as adults. All of this suggests that habit, including physical habit, can play a significant part in the teaching and learning of pronunciation. Finally, we should recall that Chomsky himself commented that a linguistic theory of the kind he pursued had little to offer for actual language learning or teaching. In fact, the idea of the natural acquisition (as opposed to conscious learning or deliberate habit formation) of syntax has won many devotees in the language teaching world, because it seems to support the naturalistic principles of Communicative Language Teaching (CLT), especially in forms such as the "Natural Approach" proposed by Terrell and Krashen's (1983), later elaborated by Richards and Rodgers (2001)[35]. In its "purest" form, CLT more or less excludes the deliberate teaching and learning of structures. However, as Michael Swan points out (2010), language teaching "fashions" tend to swing backwards and forwards[36], and even in the case of grammatical structures, more recent theories accept the importance of learners having enough conscious knowledge to take note of various features of the target language as they encounter them and

[34] Lenneberg, J. 1967.
[35] Krashen, S. & Terrel, T. 1983.
[36] Swan, M. 2010: 4.

to become aware of their own their problems, with a view to correcting them – the best known proponent of these theorie being Jack Richards, who described his “Noticing Hypothesis” in 2005[37].

In fact, very few teachers would rely on the natural acquisition of accurate pronunciation. As Celce Murcia (1996) pointed out, CLT tends to downgrade the importance of pronunciation on the grounds that perfect pronunciation is unnecessary, but this does not mean pronunciation can be ignored[38]. One communicative approach that recognizes the value of awareness and pattern-practice for pronunciation, and has won considerable support, is Charles Curran’s Community (or Counselling) Language Learning (CLL), admirably described by Richards and Rogers (1986)[39] Interestingly CLL, which strongly favours learner autonomy and places (or relegates) the teacher in a supportive (“counselling”) role, uses awareness and pattern practice for pronunciation development in the opposite of the traditional way: here, it is the leaner who uses the teacher as a provider of the correct form, and who corrects herself by “listening and repeating” until she (the learner) is satisfied. Evidence of the success of this approach comes mainly from English as a Foreign Language teaching (e.g. in Larsen Freeman, 2000[40]): in the case of speakers of a non-tonal language who are trying to learn Chinese, conscious awareness and “noticing” would be all the more important.

As far as the most effective way of actually teaching and learning pronunciation is concerned, the relevant handbook published by the Australian Department of Education Training and Youth Affairs adopts a strongly “communicative stance”, but points out that

> “Learning pronunciation requires an enormous amount of practice, especially at early stages. It is not unreasonable for learners to repeat a particular phrase or sentence twenty or fifty times before being really comfortable with it. Unfortunately, ‘drilling’ has been out of favour in language classes for some

[37] Richards, J. 2005: 85-92.

[38] Celce Murcia, M. 1996.

[39] Richards, J. & Rodgers, T. 1986.

[40] Larsen Freeman, D. 2000.

time, due to association with several bad aspects of the behaviorist method of teaching. Indeed some forms of drilling are at best a waste of time, and can even be a hindrance to learning. However, drilling of real, useful phrases which can actually be used outside the classroom is highly advantageous to learners." [41] (Fraser 2001)

Most of the sources quoted so far refer to the teaching of English as a foreign or second language, but there is no particular reason for not applying the same principles to TCFL. As far as CA is concerned, the renowned linguist Li Wang said, "Teaching Chinese as a foreign language, I think the most effective way is to teach Chinese based on the comparison of two languages." The general consensus among pedagogical theorists seems to favour a largely behaviourist, drill-based methodology, but this does not imply mere mechanical reproduction of syllables: there is also a necessity for general training in articulation, in the gradual tightening and release of the voval cords, for example, as a prerequisite for successful use of pitch in syllabic tone and elsewhere[42].

Therefore, it seems reasonable to adopt a relatively "behaviourist approach" to this area of language learning, and we will continue to develop the theoretical framework of this study by examining the notion of contrastive analysis in more depth, with particular reference to pronunciation.

2.1.2. Contrastive Analysis

Contrastive analysis systematically comparing the differences and similarities between L1 and L2 was conducted mainly from the 1940s to the 1960s. Fries (1945), one of the leading applied linguists of the day, claimed that, the most efficient materials are those that are based on a scientific description of the language to be learned, carefully compared with a parallel description of the native language of the learner[43]. Whorf (1941) first employed the term "contrastive analysis" to denote a comparative study of L1 and L2 which emphasizes

---

[41] Fraser, H. 2001:18.

[42] Guan 2000, Cheng 1997 and Zhao 1997.

[43] Fries, C. 1945.

linguistic differences[44]. Fisiak (1981) defines contrastive analysis as a sub-discipline of linguistics concerned with the comparison of two or more other languages or subsystems of languages in order to determine both the differences and similarities between them[45].

The publication of Lado's *Linguistics across Cultures* marks the real beginning of modern applied "contrastive analysis". In the preface, Lado states that "we can predict and describe the patterns that will cause difficulty in learning, and those that will not cause difficulty, by comparing systematically the language and culture to be learned with the native language and culture of the student" (1957)[46].

CA favours the apparently reasonable assumption that learning a second language is facilitated whenever there are similarities between that language (the L2) and the mother tongue (L1), while learning may be interfered with when there are marked contrasts between the L1 and L2. CA analysis assumes L1 has influence on L2 learning at the phonological, morphological, and syntactic levels and is especially useful when it comes to adequately describing the sounds and grammatical structures of two languages.

However, CA has been contested. Its claim to predictive value (i.e. identifying learning difficulties) has been especially challenged by researchers such as Wardhaugh (1970), who asserted that "contrastive analysis can be of little or no help at all in the learning task" and argued that:

> All natural languages have a great deal in common so that anyone who has learned one language already knows a great deal about any other language he must learn. Not only does he know a great deal about the other language even before he begins to learn it, but also the deep structures of both languages are very much alike, so that the actual differences between the two languages are really quite superficial [47].

Later on, "error analysis" (EA) arose as a counter-theory to CAH (see below). The debate

---

[44] Whorf, B. 1941.

[45] Fisiak, J. 1981: 1

[46] Lado, R. 1957: 7.

[47] Wardhaugh, R. 1970: 123-130.

raged on between CA and EA through the 1970s.

Despite these criticisms, as Larsen-Freeman and Long (1991) point out, Contrastive Analysis continued to be conducted, particularly in Europe, where professionals continued to regard it as a valid methodological option, regardless of whether it was scientifically proven or not: it is hard not to assume that there was no experiential evidence to support this persistence[48]. As Lanlin Zhang (2005, p1) points out: "As an L2 learner, teacher, and researcher for many years, experience and intuition tell me that contrastive analysis of the phonological, grammatical and lexical difference between Chinese and Hungarian can be very useful for L2 learner to pinpoint potential learning difficulties and to better improve the communicative skills so that they can be successful in their academic study and life"[49].

2.1.3. Reflections on the Contrastive Analysis Hypothesis

The Contrastive Analysis Hypothesis (CAH) is the main content of the theory of contrastive analysis.

In Lado's *Linguistics across Culture* (1957), the author states,

> "We can predict and describe the pattern that will cause difficulty in learning, and those that will not cause difficulty, by comparing systematically the language and culture to be learned with the native language and culture of the student."
>
> "Individuals tend to transfer the forms and meanings and the distribution of forms and meanings of their native language and culture to the foreign language and culture---both productively and when attempting to speak the language and to act in the culture and receptively when attempting to grasp and understand the language and culture as practiced by natives."
>
> "… those elements that are similar to this native language will be simple for him, and those elements that are different will be difficult.[…] Where two languages were similar, positive transfer would occur; where they were different, negative

48 Larsen-Freeman, D. & Long, M. 1991.
49 Zhang 2013: 45.

transfer, or interference, would result."[50]

Later, Wardhaugh (1970) was to describe all these statements as the "Contrastive Analysis Hypothesis" (CAH)[51].

The basic pedagogical assumptions of CAH were summed up by Gass and Selinker (1994) as follows:

- Language belongs to behaviour. Learning a second language means establishing a new set of language habits.
- The major source of error in the production and/or reception of a second language is the native language.
- The greater the differences between L1 and L2, the more errors will occur.
- The structural systems of L1 and L2 should be compared and contrasted to determine the areas where they differ most.
- Errors can be predicted to occur mainly with the areas of L2 that differ the most from the L1.
- Language teaching should be focused on the areas where the L1 and L2 systems differ the most.
- One can account for errors by considering the differences between the L1 and the L2.
- What one has to do in learning an L2 is to learn the differences. Similarities can safely be ignored, as no new learning must occur.
- Difficulty and ease in learning are determined respectively by differences and similarities between the two languages.
- CAH covers the area of phonology, grammar, writing system and culture.[52]

Although, as I have already made clear, I find it reasonable to apply many of the assumptions and presumptions of the CAH to the teaching and learning of pronunciation, this does not mean that I agree with the common assumption that similarities between the L1 and L2 will necessarily facilitate learning while marked contrasts will interfere (Stockwell et al., 1965)[53].

---

[50] Lado, R. 1957: 2.

[51] Wardhaugh, R. 1970: 123-130.

[52] Gass, S. & Selinker, L. 1994.

[53] Stockwell, R., Bowen, J. & Martin, J. 1965.

It is true that a quite different system is likely to create difficulty. For example, most Hungarian students will find it difficult to pronounce the consonants written as "zh/ch/sh" in *pinyin*. The difficulty results from the absence of such (or similar) sounds in Hungarian. Yet, are they really the most difficult part for Hungarian students?

In fact, it is often the similarities rather than the differences that give rise to incorrect associations between the L1 and L2, resulting in errors. According to Oller and Ziahosseiny (1970), the real difficulty results from the subtle differences[54]. The more subtle the differences, the more difficult it is to learn the second language. On the other hand, two completely different items will not result in interference. Besides, Oller and Ziahosseiny point out that while learning, people often ignore the small differences but attend to the obvious contrasts. Therefore, the items which are quite similar in form and meaning will most easily bring difficulties. For example, although Hungarian students find [ə] or [ɤ], which are the allophones of /e/, difficult to pronounce, they can hardly tell the difference between [ə],[ɤ] and [e], [ɛ], which are quite similar but not identical nature and which largely result in Hungarian accent of Chinese.

The famous psycholinguist Osgood (1953) developed this point when he commented on such a phenomenon, noting that when two sets of materials to be learned are quite different or are easily discriminated by the learner, there is relatively little interaction; that is, learning one has little effect upon learning the other. If they are similar in such a way that the learning of one serves as partial learning of the other, there may be facilitation, or positive transfer. If, however, the similarities either of stimuli or reponses are such that responses interfere with one another, then there will be greater interference as similarity increases.[55]

2.1.4. Reflections on the Hierarchy of Difficulty

Another opposing voice against the assumption that "the greater the difference, the more difficult it is to learn" comes from the theory of Hierarchy of Difficulty, which posits that

[54] Oller, J. & Ziahosseiny, S. 1970: 183-189.

[55] Osgood, C. 1953.

different types of similarities and differences represent different levels of learning problems. It was first put forward by Briere (1968) in the field of phonological systems[56], later it was applied in other fields of language, such as syntax and grammar by Stockwell, Bowen, and Martin (1965)[57] and finally it was simplified by Prator (1967) and classified into six categories or levels for practical classroom application, with the most difficult category at the top and the least difficult at the bottom[58].

Table 1: *Categories of difference and levels of difficulty, from Prator, 1967*

Level	Category	Explanation
5	split	One item in L1 is split into two or more in L2.
4	new	The item exists in L2 but is absent in L1.
3	reinterpretation	The item is present in L1 but appears in a new form in L2.
2	absent	The item exists in L1 but is absent in L2.
1	coalescence	Two or more items in L1 are combined as one in the L2.
0	correspondence	The items are the same in both L1 and L2.

Although the validity of the hierarchy of difficulty is open to discussion, it does offer insights into the way in which Hungarian students experience problems in their Chinese studies, and in the following sections of my thesis, I will attach great importance to this theory when explaining learners' difficulty in pronunciation.

2.2. The theory of error analysis

The publication of Corder's seminal article on the significance of learner's errors (1967) marks the beginning of second language acquisition research[59]. In this article, he claims that errors can be significant in three ways: (1) they tell the teacher how far the learner has come

[56] Brière, E. 1968.
[57] Stockwell, R., Bowen, J. & Martin, J. 1965.
[58] Prator, C. 1967.
[59] Corder, S. P. 1967: 161-170.

and what he still must learn; (2) they give the researcher evidence of how language is learned; (3) they are a device the learner uses to test out his hypotheses concerning the language he is learning. He argues that first and second language learning share basically the same processes and that whatever differences exist are explainable in terms of motivation. He makes a distinction between a mistake and an error. A mistake is a random performance slip caused by fatigue, excitement, etc., and therefore can be readily self-corrected, an error is a systematic deviation made by learners who have not yet mastered the rule of the L2. A learner cannot self-correct an error because it is systematic in nature and reflects a learner's transitional competence of L2. Errors, according to Pit Corder, are signs that learners are actively engaged in hypothesis testing which would ultimately result in the acquisition of target language rules.

Compared to contrastive analysis, error analysis emphasizes the errors made by learners, instead of systematically comparing L1 with L2. Corder suggests the following steps in EA research:

(1) Collection of errors made by learners.

(2) Identification of errors.

(3) Description of errors.

(4) Explanation of errors.

(5) Evaluation of errors.

Early error analysis mainly studies the reasons that caused the errors. The types of error are as follows: (1) errors caused by L1 interference, termed "interlingual errors" by Richards (1971), who distinguished them from errors commited by SL learners of different origins regardless of their L1 system. He called the latter type (2) "intralingual errors"and explained them as a kind of overgeneralization, caused by learners failing to master the limitations of the rules. [60]George (1972) identified simplification and redundancy reduction errors which resulted when leaners use communicative strategies[61]. These errors were classified by Selinker (1972)

[60] Richards, J. 1971: 12-22.

[61] George, H. 1972.

as (3) "communication-based errors"[62]; Stenson (1974) added the category of (4) "induced errors", which are errors resulting from a teacher's sequencing or presenting two linguistic items in a confusing way.[63]

Error analysis gave a new perspective of errors and plays a very important role in explaining and analyzing the errors made by learners in second language acquisition. However, it also has some defects. First, error analysis cannot observe learners' avoidance strategies, so it fails to account for all the areas of the second language in which learners have difficulty. Second, due to focusing only on errors, researchers were denied access to the whole picture. They studied what learners were doing wrong, but not what made them successful. As Harley (1980) said:

> "the study of errors that L2 learners make can certainly provide vital clues as to their competence in the TL, but they are only part of the picture… it is equally important to determine whether the learner's use of 'correct' forms approximates that of the native speaker. Does the learner's speech evidence the same contrasts between the observed unit and other units that are related in the target syetem? Are there some units that he uses less frequently than the native speaker, some that he does not use at all?"[64]

The limited perspective did not lead to the demise of error analysis: it remains an important theory of second language acquisition, both in teaching and in learning.

Error analysis with regard to Chinese as a second language starts from the 1980s. According to Zhao (2011), the publication of Lu's *Zhongjieyu Lilun yu Waiguoren Xuexi Hanyu de Yuyin Pianwu Fenxi* (1984) marks the beginning of the study of the acquisition of Chinese as a second language acquisition[65]. In his groundbreaking paper, Lu first introduced the concepts of "error analysis" and "interlanguage" into China. At the same time, he analyzed the pronunciation errors made by foreigners in learning Chinese, working within the broad

---

[62] Selinker, L. 1972: 209-231.

[63] Stenson, N. 1974.

[64] Harley, B. 1980: 3-30.

[65] Zhao 2011; Lu 1984.

framework of interlanguage theory. Firstly, he examines negative transfer from the mother tongue in learners' pronunciation of Chinese sounds; secondly, he discusses the transfer that may occur when learning tones among those learners whose mother tongue is a non-tone language; finally, he mentions errors which are caused by insufficient or inappropriate training. All in all, the pronunciation error analysis of foreigners learning Chinese carried out by Lu falls within the framework of interlanguage theory, which we will now consider in detail.

2.3. Interlanguage theory

The language system that the learner constructs out of the linguistic input to which he has been exposed has been variously referred to by Pit Corder (1967) as an *idiosyncratic dialect*[66], by Nemser as an *approximative system*[67] and by Selinker (1972) as an *interlanguage*[68]. These three terms differ somehow in their focus, and *interlanguage* is now most commonly used term. Selinker regards an interlanguage as an independent language system, stating that:

> "We focus our analytical attention upon the only observable data to which we can relate theoretical predictions. The utterances which are produced when the learner attempts to say sentences of a TL. This set of utterances for most learners of a second language is not identical to the hypothesized corresponding set of utterances which would have been produced by a native speaker of the TL had he attempted to express the same meaning as the learner. Since we can observe that these two sets of utterances are not identical, then in the making of constructs relevant to a theory of second-language learning, one would be completely justified in hypothesizing, perhaps even compelled to hypothesize, the existence of a separate linguistic system based on the observable output which results from a learner's attempted production of a TL norm. This

[66] Corder, S. P. 1967: 161-170.
[67] Nemser, W. 1971: 115-124.
[68] Selinker, L. 1972: 209-231.

linguistic system we will call 'interlanguage' (TL)."[69]

Ellis (1997) accepted this terminology, claiming that the learner constructs a system of abstract linguistic rules which underlies comprehension and production of the L2. This system of rules is viewed as a 'mental grammar' and is referred to in his widely influential work as an 'interlanguage'[70].

Obviously, Selinker's definition of interlanguage emphasizes that the learner's linguistic system is an "observable", "separate" linguistic system, that is "based on output". It is also, of course, unstable and liable to develop; both Corder and Nemser believe there is a continuum between L1 and L2 which all learners traverse. Interlanguage is, therefore, a dynamic transitional system and will (or so the learner hopes) be continuously approaching the TL. As Nemser (1971) put it when discussing his "approximative system" concept,

> "Our assumption is three-fold: (1) Learners' speech at a given time is the patterned product of a linguistic system (La), distinct from LS and LT (the source and the target language) and internally structured. (2) La's at successive stages of learning form an evolving series La_1….La_n, the earlier occurring when a learner first attempts to use the LT, the most advanced at the closest approach to LT…. (3) In a given contact situation, the La's of learners at the same stage of proficiency roughly coincide with major variations ascribable to differences in learning experiences."[71]

There are some other views of interlanguage. Adjemian (1976) applied a Chomskyian paradigm to interlanguage competence. He regards interlanguage as linguistic knowledge, and states there is no difference between the task of the linguist modeling the linguistic knowledge of the native speaker and that of the second language learner[72]. The same methodology and data can be used in both cases.

---

[69] Selinker (ibid).

[70] Ellis, R. 1997.

[71] Nemser (op cit): 116.

[72] Adjemian, C. 1976: 297-320.

Tarone (1983) noted that interlanguage not only develops diachronically, but also includes variations at one point in time, and proposed the concept of "variable competence" to describe the learner's IL capability.

> "One phenomenon which must be accounted for by any theory of second-language acquisition is the phenomenon of systematic variability in the utterances produced by second-language learners as they attempt to communicate in the target language [...] Further, so far is seems that the pattern is for the TL feature to be supplied more frequently in the careful style elicited by some tasks, and less frequently in the casual speech."[73]

Tarone argues that the learner's IL capability is not homogeneous, but heterogeneous, made up of a continuum of styles, or "degrees of accuracy" in terms of the target language, depending on the circumstance under which the learner is attempting to use the TL.

In the field of TCFL, there exist several views in terms of the precise definition of interlanguage. First, Lu (1984) agrees with the definition of interlanguage as a linguistic system which is affected by the learner's incorrect generalizations and inferences about TL rules. It is neither the same as the learner's mother tongue, nor the same as the TL[74]. Lu considers interlanguage as an observable, separate linguistic system, in accordance with Selinker's view. Second, Lü (1993) regards interlanguage as a special TL system of second language learners, which differs both from the learner's mother tongue system and from the TL system in phonetics, vocabulary, grammar, culture, communication and so on. It is a dynamic linguistic system and will be continuously approaching the TL[75]. His view corresponds to Corder and Nemser's "transitional system". Somewhat different are Sheng's "Mixed View"[76] and Wen's "Non-connection View"[77]. Sheng states that interlanguage is a mixture of two languages, while Wen believes that there is no connection between

[73] Tarone, E. 1983: 144.
[74] Lu 1984.
[75] Lu 1993.
[76] Sheng 1990.
[77] Wen 1992.

interlanguage and learner's mother tongue or TL. She further claims that interlanguage is not affected by any present languages.

Apparently, however, although they represent different points of view, Chinese scholars have not added new categories to the debate. As Zhao (2011) points out, "the discussion of interlanguage in the TCFL field hasn't gone beyond the range of what the foreign scholars have done.[78]

2.4. The relationships between interlanguage, contrastive analysis and error analysis.

Wardhaugh (1970) proposed a difference between the strong version and the weak version of the contrastive analysis hypothesis. The strong version holds that the ultimate aim of contrastive analysis is to predict errors in second language learning based on a systematically contrastive analysis of the L1 and L2. This is called the "predictive view", but many empirical investigations have shown that predictions based on CA are not always borne out. The weak version starts with learner errors and explains the reasons of errors by comparing the similarities and differences between the two languages; this is called the "explanatory view".[79] According to Zhao (2011) the weak version became a part of error analysis[80].

In the 1960s, Corder gave a new perspective and content to error analysis. In Zhao's (2007) opinion, what Corder advocates actually marks the beginning of the study of interlanguage[81]. He argues that from the perspective of its theoretical foundations, error analysis denies the validity of contrastive analysis, although error analysis does not actually reject the comparison method. However, error analysis compares learners' errors with the target language rather than L1. This shift in focus is not only about research methods, but about the theoretical directions. He further indicates that the theoretical value of error analysis is that its study first focuses on learners and learners' language systems.

---

[78] Zhao 2011: 233. (my translation)

[79] Wardhaugh, R. 1970: 123-130.

[80] Zhao 2011: 233. (my translation)

[81] Zhao (ibid).

As to the relationship among interlanguage, contrastive analysis and error analysis, Selinker (1992) points out that all studies about second language acquisition are based on Corder's error analysis and his study of interlanguage, and Corder's achievements reflect his full understanding of contrastive analysis, error analysis and bilingual research. To some extent, Selinker's view reveals the inheritance relations among the three theories[82]. Interlanguage theory was put forward based on the full understanding of CA and EA. Nemser's approximative system came from a series of phonetic comparisons and experimental research projects; Selinker's interlanguage theory was proposed based on experimental research into phonetic comparison and syntactic comparison; Corder's notion of the idiosyncratic dialect started from error analysis. But as Zhao (2011) notes, none of this means that their researches took over either the theory of CA or that of EA on a wholesale basis[83]. Selinker believes that CA and EA can function as preliminary stgeas of interlanguage study. In other words, interlanguage studies can make use of the methods of CA and EA, but the theoretical nature of the three theories must not be confused. In addition, Zhao (2011) points out that the issues surrounding language transfer somehow reveal the inheritance relations among the three theories, since none of the three scholars mentioned denied the reality of language transfer, although they hold different opinions in terms of the ways in which language transfer is caused and operates.[84]

---

[82] Selinker, L. 1972: 209-231.

[83] Zhao 2011: 236.

[84] Zhao (ibid).

3. The Comparison of Hungarian and Chinese Phonetic Systems

After a brief summary of the overall differences between the Hungarian and the Chinese languages, this chapter focuses on a comparative phonological analysis, not only describing the salient differences but attempting to distinguish which of them are most likely to cause difficulties to Hungarian leaners of CFL. The very nature of pronunciation determines the important position of the pronunciation teaching in the teaching of Chinese as a foreign language. Zhao Yuanren[85] uses his *Guoyu Rumen* (*Mandarin Primer)* as the teaching material when teaching Chinese abroad. "Most teaching is about pronunciation; then move on to the teaching of other aspects". During that period, pronunciation teaching was highlighted. Along with the boom of the teaching Chinese as a foreign language, internally, this discipline has been unevenly developed. Considerable developments have been made in many aspects. However, pronunciation teaching has been on the contrary. As Mr. Lin Dao said, "pronunciation teaching has not made any progress. On the contrary, it has greatly regressed." Possibly as a result, research not only into pronunciation teaching but into pronunciation itself has attracted less attention than that in other fields. Even so, during the past several decades, a considerable body of work, including many excellent theses and papers on pronunciation and pronunciation teaching, has emerged. These theses and papers are of great significance for the theoretical research and teaching practices in pronunciation.

3.1 Introduction

Siptar & Törkenczy (2000) describe Hungarian as "a Uralic language spoken in Central Europe"[86]. In terms of number of speakers, it is the twelfth largest language of Europe. The majority of speakers reside in Hungary itself, but Hungarian-speaking minorities are to be found in the neighbouring states: Slovakia, Austria, Romania, Serbia, Croatia, Slovenia and Ukraine; there are also groups of Hungarian speakers in more distant countries such as Canada and the United States, resulting from waves of emigration during the nineteenth and

[85] Chao 1948.

[86] Siptár, P. & Törkenczy, M. 2000: 13.

especially the twentieth centuries.

Hungarian is not only very different from the majority of (mainly Indo-European) European languages, but also unusual among the members of the Uralic family. Siptár & Törkenczy (2000) point out that it has no close relatives: "The Ob-Ugric languages (Vogul and Ostyak), traditionally bundled together with Hungarian into the Ugric branch Of Finno-Ugric languages, are radically different from Hungarian in their phonology, syntax, and vocabulary."[87]

Hungarian is defined by Kornai (1994) as a language of agglutinating morphology[88], with non-configurational syntax (Kiefer and É. Kiss 1994)[89], and syllable-timed prosody (Roach 1982, Crystal 1995)[90]. Its vocabulary includes large numbers of loanwords.

The variety of Hungarian discussed in this thesis is what Nádasdy (1985) defines as Educated Colloquial Hungarian (ECH) which is typically used by the university-based population in our focus and differs from Standard Literary Hungarian and various types of non-standard speech.

Chinese is a Sino-Tibetan language used mostly in mainland China, Taiwan, Hong Kong and Macau. In terms of number of speakers, it is undoubtedly the "largest" language on the planet, being spoken by about one fifth of the population of the world. Not surprisingly there are many, more or less distinct, varieties of Chinese (see e.g. Li (1972) and Norman (1988) for the division of these varieties into main groups)[91]. Whether these dialects, many of which are not mutually intelligible, should be regarded as separate language is perhaps more of a political than a linguistic question and as such will not be discussed here. Two major varieties are most often taught as second or foreign languages outside China: Standard Chinese and Cantonese; fhe former is far more common. Standard Chinese is known as "普通话 *pǔtōnghuà*", and is officially regarded as the common language of China, which means that it is taught in the schools and employed in all governmental and offical transactions. It is based

[87] Siptár, P. & Törkenczy, M. (ibid): 13.

[88] Kornai, A. 1994.

[89] Kiefer, F. & É. Kiss, K. (eds.) 1994.

[90] Roach, P. 1982.

[91] Li Fang-Kuei 1972; Norman, J. 1988.

on the northern varieties, with the Beijing phonological system as its norm of pronunciation and modern vernacular literary as its syntax specification. It has other names such as "国语 *guóyǔ*", "汉语 *hànyǔ*", "中文 *zhōngwén*" and so on.

While Hungarian is agglutinating, Chinese is an analytic language. It also differs radically from European languages in its character-based writing system and in the existence of phoemic tones: the first, the second, the third and the fourth. For a more detailed description of tones, see section 4.2.3 below.

Since the topic of this thesis is the phonetic system of Chinese, the writing system is disregarded. However, certain pronunciations problems that can be caused by the use of *pinyin*, the artificial writing system with Latin letters much used by learners of Chinese as a Second Language (CSL), are mentioned in 4.2.2.1 and 4.3.3 below. In his *Cong 'Danyin Dongci + Danyin Mingci' Jiegou de Pinxie Kan Zhengcifa de Keguan Yiju*, on the basis of an investigation of the spellings of "Monosyllabic Verbs + Monosyllabic Nouns"structures, Wang Li[92] explores the orthographical rules affecting the spellings of Chinese Pinyin, which is of some guiding significances for the pronunciation teaching practice.

It should also be mentioned that there is some disagreement in this area: in *Xiandai Hanyu Yinjie de Shuliang yu Goucheng Fenbu*, based on three authoritative reference books, *Xinhua Zidian*[93], *Xiandai Hanyu Cidian*[94] and *Xiandai Hanyu Guifan Zidian*[95], the author Lu Wo discussed about modern Chinese syllables, including the amount of syllables on which people disagree, the reasons for these disagreements and the respective rate of openness of basic syllables, so as to provide reliable data reference to those who have chosen to learn standard modern Chinese[96].

Focussing on one particular area, Li Ming in the *ABshi Shuangyinjie Xingrongci Chongdieshi*

[92] Wang 1998.
[93] Xinhua Zidian 1998.
[94] Xiandai Hanyu Cidian 2002.
[95] Xiandai Hanyu Guifan Zidian 2011.
[96] Lu 2001.

de Duyin Kaocha [97], on the basis of investigations and analyses of overlapping pronunciations of 211 adjectives, points out that the overlapping pronunciations of AB-dissyllable are closely related to the tone of the second syllable. If the second syllable does not contain a neutral tone (except for some special situations), he claims, the third and fourth syllables still carry the original tones after being overlapped and appear more regularly; if on the other hand the second syllable does contain a neutral tone, the situations are comparatively more complicated after overlapping. [98] Li Ming goes on to provide detailed analyses of such complicated situations. Owing to research achievements of his paper, research on overlapping pronunciation of AB-dissyllable adjectives reaches new heights.

Research on the prosodic features of Chinese has drawn increasing attention in the field of TCSL, since it is quite difficult for teachers by simply focusing on consonants and vowels as phonemes of Chinese to fundamentally correct the deviant accents of foreign students, though at the same time the very existence and vital important of syllabic tone can make it difficult to identify, let alone to analyse, suprasegmental intonation. Research results in this respect include the following:

In the thesis *Hanyu Shengdiao yu Yudiao de guanxi* which was written by Cao Jianfen (2002), the author clarified the deep essence and structure of Chinese intonation by objectively approaching the production process of tone and intonation and their interaction based on an acoustic-phonetic investigation[99]. The preliminary results indicate that intonation pattern is mainly related to the pitch register movement of the whole utterance, while tone mainly refers to the shape of pitch change of local words and word combinations. In real speech, each tone must be modified by global intonation through adjusting its relative register on one hand, and keeping its basic tone shape on the other hand. At the same time, the global intonation must be manifested through the pitch movement of each local tone. The relation between tone and intonation is an “algebraic sum” of pitch register instead of the tone shape.

---

[97] Li 1996.

[98] Li 1996.

[99] Cao 2002.

In his *Hanyu Jiugong Diaomo de Biaoyi Gongneng*, the author Wu Jiemin[100] exemplifies the ideographic functions of the Chinese "*Jiugong*" intonation modes. Besides, it is believed that tones of Yang (阳) and Yin(阴) are the two basic intonations in Chinese, which vary in one and the same discourse due to the changes of emotion. In the *Qingsheng Yingao Suoyi*, Wang Yunjia discusses changes in the pitch range of neutral tone, the relationship between the tone value of neutral tone and non-neutral tone and changes in the pitch of the neutral tone when it occurs in the middle of the trisyllabic word[101]. This paper is of some significance for making an intensive study on the pitch of neutral tone and guiding the tone teaching.

In her *Qingsheng de Benzhi Tezheng*, Liu Juan offers an analysis of light tone in terms of phonology and attempts to make distinctions between light tone and tones of Chinese characters and modified tone, especially between light tone and unstressed sound[102], while in his *Hanyu Yingao Xitong de Youshengxing he Qubiexing*, Shen Jiong points out that the tone and intonation are in different linguistic levels and have their own patterns of pitch changes which belong to two relatively independent pitch systems[103]. In the *Lun Hanyu de 'Ziran Yinbu"*, Feng Shengli concludes that there exists a purely rhythmic "natural foot" which is not influenced and constrained by syntax, semantics, pragmatics and context and deduces "directional" rules concerning the application of the natural foot, which are of enlightening significance[104].

In his *Hanyu Hanyu Jielü de Ziran Tezheng*, Wu Weishan discusses the interaction between Chinese tone and rhythm, seen as two sets of rhythmic systems and the structural pattern of Chinese rhythm[105], including the following two aspects: 1) building a basic framework of Chinese rhythmic structure; 2) uncovering typical characteristics of Chinese rhythmic structure, which are of some guiding significance for the pronunciation teaching, though

[100] Wu 2000.
[101] Wang 1996.
[102] Liu 1997.
[103] Shen 1995.
[104] Feng 1998.
[105] Wu 2003.

probably at more advanced level than that reached so far by most of the learners in the sample used in the present investigation.

3.2. The contrastive analysis of consonants in Hungarian and Chinese

The following section provides a comparative phonological analysis of Hungarian and Chinese, noting both similarities and differences between the two languages, with particular reference to points where differences are likely to lead to real problems for Hungarian learners. The analysis begins with broad, systemic differences and continues with the detailed examination of specific consonants, vowels and suprasegmental features. It concludes with a summary of the problem areas, ranked according to the categorisation proposed by Prator (1967)[106].

At a general level, five broad systemic differences may be observed. The first and most obvious of these is the fact that Chinese is a tonal language in which the rising, falling, falling-rising or steady pitch of individual syllables regularly carries lexical, and sometimes grammatical, meaning - a feature which is completely absent from Hungarian. The second systemic difference is that Hungarian makes use of the voiced-voiceless contrast, where Chinese does not, while Chinese uses the aspirated-unaspirated contrast where Hungarian does not. The third is that vowel length is contrastive in Hungarian but not in Chinese. The fourth is that Hungarian includes phonemic oppositions between the mid-high vowels [E, o, ↵], and lacks [↔,Φ], while in Chinese all of these are allophones of a single phoneme. The fifth systemic difference is that unlike Chinese, Hungarian in general lacks genuine semi-vowels, and thus opening/falling diphthongs[107].

We will now consider individual speech sounds, starting with the consonants, which are first described language-by-language and then compared with reference first to their place of articulation and then to their manner of articulation. We will then move on to vowels,

---

[106] Prator, C. 1967.

[107] There are exceptions: a limited set of recent loan-words such as "auto", which sometimes preserve the diphthongs used in their original languages.

categorized as high, mid- and low, then to syllable structure, and finally to the suprasegmental issues of stress and intonation.

Except where otherwise specified, the descriptions of Hungarian phonetic features in the rest of this chapter are based on Siptár & Törkenczy (2000)[108], while the descriptions of Chinese phonetic features are based on Huang and Liao (1991). For much of the comparison between the two phonetic systems and the categorisation of the similarities and differences according to degree of difficulty, I am indebted to Dr Bartos Huba for giving me access to his unpublished notes on the subject.

The Hungarian consonant system includes twenty four items: p, b, t, d, ty, gy, k, g, f, v, sz, z, s, zs, h, c, cs, dzs, m, n, ny, l, r, j(ly). Among them, there are nine pairs of consonants in which each pair shares exactly the same articulatory place, the only difference between each pair being whether it is articulated with or without the vibration of the vocal folds. The one which is articulated with the vibration of the vocal folds is named "voiced" consonant, and those articulated without the vibration of the vocal folds are voiceless consonants. The nine pairs are shown in the following table. For more specific classifications of all the consonants, we will discuss later in this section.

Table 2: *Hungarian voiced and voiceless consonants.*

voiced consonants	b, d, g, v, z, zs, (dz), dzs, gy
voiceless consonants	p, t, k, f, sz, s, c, cs, ty

The Chinese consonant inventory consists of twenty-two items; they are as follows: b, p, m, f, d, t, n, l, g, k, h, j, q, x, zh, ch, sh, r, z, c, s, ng. They can be incorporated into different categories in terms of the manner of articulation and the the place of articulation. In contrast to the Hungarian consonant system, voiced/voiceless pairs do not exist in Chinese, but there

[108] Siptár, P. & Törkenczy, M. 2000. *Note that Siptár & Törkenczy do not treat 'dz' as an independent comsonant, but merely as the d + z cluster.*

are unaspirated/aspirated pairs. Members of each pair share exactly the same place of articulation; the main difference between them is whether they are articulated with or without airstream. The distinction between tense "fortis" and lax "lenis" articulation[109] is also an important feature. In each pair the one which is articulated with airstream is named the "aspirated" consonant; the one which is articulated without airstream is the "unaspirated" consonant. The six pairs of unaspirated/aspirated consonants are shown in the following table. More details of categories will be discussed later in this section.

Table 3: *Chinese unaspirated and aspirated consonants*

Unaspirated consonants	b, d, g, z, zh, j
Aspirated consonants	p, t, k, c, ch, q

3.2.1. Contrastive analysis of consonants in Hungarian and Chinese by place of articulation

3.2.1.1 Labials

The Hungarian labials are [b, p, m, f, v], the nearest Chinese equivalents being [β8, p^h, m, f]. In other words, both languages have bilabials and labiodentals, with no notable articulatory difference. However, the second systemic difference mentioned above applies: Hungarian distinguishes between voiced and voiceless unaspirated semi-fortis consonants while Chinese distinguishes between voiceless aspirated lenis and unaspirated fortis, and while Hungarian contrasts the voiced and voiceless fricatives [f, v], Chinese has only the voiceless [f].

3.2.1.2. Dental / alveolar stops

The Hungarian alveolar stops are [d, t, n, l], the nearest Chinese equivalents are [δ8, t^h, n, l]. In other words, both languages have dentialveolars, with no significant articulatory difference. However, the same systemic difference applies as with labials: Hungarian distinguishes between voiced and voiceless unaspirated semi-fortis consonants while Chinese contrasts voiceless aspirated lenis and unaspirated fortis. One particular feature of Chinese that

[109] lenis = pronounced with lax muscles of the oral cavity, fortis = pronounced with tense muscles.

Hungarian learners find hard to master is the fact that in Chinese [n] is pronounced slightly differently when it is in final position, but this does not normally give rise to confusion or cause any loss of comprehensibility.

3.2.1.3. Alveolar affricates and fricatives.

The same systemic difference applies here too: Hungarian with [(dz), ts, s] contrasts voiced and voiceless unaspirated semi-fortis while Chinese with [ts, ts^h, s] contrasts voiceless aspirated lenis and unaspirated fortis.

3.2.1.4. Palatals

Unlike the speech sounds discussed so far, palatals present relatively serious difficulties for Hungarian learners of Chinese because both languages have two articulatory regions here with completely different places (and manner) of articulation. Hungarian contrasts palato-alveorals (or laminal palatals) [τΣ, δZ, Σ, Z] with (dorso-) palatals [c, ⊇⊇|], while Chinese contrasts alveolar-palatal doubly articulated (*apda*): [τ⌡, τ⌡η, ⌡] with retroflex [τ♣, τ♣η,♣, |]. Problems arise because Hungarian learners tend to associate Chinese retroflexes with Hungarian palato-alveolars, and Chinese *apda*'s with Hungarian dorso-palatals.

3.2.1.5. Velars

In the case of velars the place of articulation is identical in the two languages, though once again the second systemic difference mentioned above applies: Hungarian contrasts the voiced and voiceless unaspirated semi-fortis [g, k], while Chinese has the voiceless aspirated lenis and unaspirated fortis [[γ* , k^h]. Problems may arise with the Chinese [N], because although this sound occurs in the speech of Hungarian learners, they are unaware of this fact and tend to pronounce it as [ng]. – their attention must be drawn to the fact that whenenever they pronounce a [ng] sequence, they actually pronounce it as [ng], so they should simply begin to pay attention to this variant of the phoneme /n/ occurring before [g], and try to pronounce it in isolation, without going on and actually pronouncing a [g] after it.

3.2.1.6. Postvelars

In this case there is a marked difference between the Hungarian glottal-laryngeal [h] and Chinese uvular approximant [Ξ] (though in several Chinese dialects such as 上海话), the 'velar' fricative is actually much like Hungarian glottal [h]). It is certainly worth drawing the Hungarian learner's attention to the difference, but in practice there is relatively little perceptible difference and no confusion arises when s/he replaces [Ξ] with [h].

3.2.2. Contrastive analysis of consonants in Hungarian and Chinese by manner of articulation

3.2.2.1. The stop/affricate 'pairs'

Here the second systemic difference applies again: in Chinese the usual opposition is between an unaspirated lenis (rather like a voiceless variant of the English voiced stops) and an aspirated fortis. Both are clearly voiceless, though they may become voiced in certain phonetic environments, especially in toneless/unstressed syllables, e.g.: 的 [δ8ι] → [δ8↔] → [d↔]). In Hungarian there is no such fortis/lenis difference between members of these pairs: both have a medium degree of tenseness ('semi-fortis'), and no aspiration, but there is a clear voiced/voiceless distinction. Hungarian learners tend to equate Chinese unaspirated lenis consonants with the semi-fortis voiceless ones that occur in Hungarian. In practice, this does not often lead to serious problems, as although aspiration is completely absent from Hungarian, if Hungarian learners are made aware of this feature they can easily perceive and reproduce it. The use of *pinyin* can cause problems here, as the letters used in *pinyin* to denote voiceless unaspirated lenis stops and affricates actually denote voiced consonants in Hungarian. Learners need to be made aware of this; they also need to learn to pronounce the Chinese unaspirated lenis consonants with lax muscles.

3.2.2.2. The semi-vowels, functioning as rime- initial/final glides.

Chinese has three such items: [j, w, ǀ], all of which appear in ryhme initial position, though only the first two also appear rhyme-finally. Only one of these, [j] is found in Hungarian with the status of a proper consonant, having a voiced allophonic variant [X] in voiceless contexts ([–voice] _ [–voice] or [–voice] _ #).

Hungarian learners therefore need to be introduced to the glide variants of the full vowels [u] and [y], ([w] and [ǀ], respectively). The [w] sound usually causes fewer problems, since a similar semi-vowel exists in English (the most frequently learned foreign language in Hungary), though it must be noted that the Chinese [w] and English [w] are not identical. The [ǀ] sound, with no easily available equivalent, is somewhat more difficult. Hungarian learners need to beware of the common temptation to replace these glides with the corresponding full vowels ([i, u, y]), resulting in vowel-sequences instead of diphthongs in these rhymes.

3.2.3. Category tables

In terms of place of articulation and manner of articulation, consonants in Hungarian and Chinese can be classified into the following categories:

Table 4: *Categories of Hungarian Consonants (based on Siptár and Törkenczy, 2000)*

		Obstruents						Sonorants			
		Stops		Affricates		Fricatives		Nasals	Liquids		
		Voiced	Voiceless	Voiced	Voiceless	Voiced	Voiceless	Voiced	Voiced		
									Lateral	Trill	Approximant
Labial	Bilabials	b	p					m			
	labiodentals					v	f				
Dental/alveolars	Lamino-dentals	d	t					n	l	r	
	Lamino-alveolars				c	z	sz				
Palatal	Palato-alveolars			dzs	cs	zs	s				
	Dorso-palatals	gy	ty					ny			j ly

Velars	g	k								
Postvelar						h				

Table 5: *Categories of Chinese Consonants (based on Huang and Liao, 1991)*

		Obstruents					Sonorants	
		Voiceless					Voiced	
		Stops		Affricates		Fricatives	Nasals	Liquids (Lateral)
		unaspirated	Aspirated	unaspirated	Aspirated			
Labial	Bilabials	b	p				m	
	labiodentals					f		
Dental/alveolar	Apico-dentals			z	c	s		
	Apico - alveolars	d	t				n	l
	Apico – post-alveolars (retrofle-x)			zh	ch	sh	r	
Alveolar-palatal doubly articulated(apda)				j	q	x		
Velar		g	k				ng	
Postvelar						h		

3.3. The contrastive analysis of vowels in Hungarian and Chinese

There are fourteen vowels in Hungarian. The vowel inventory of Hungarian consists of a, á, e, é, i, í, o, ó, ö, ő, u, ú, ü, ű. A phonetic classification of the Hungarian vowel system is shown in the following table, which is classified in terms of the duration of vowels.

Table 6: *Hungarian short and long vowels*

short vowels	a, e, i, o, ö, u, ü,
long vowels	á, é, í, ó, ő, ú, ű

According to the relative position (high or low) of the body of the tongue in the mouth during articulation, here we also roughly divide them into two groups, see the following table. More specific classifications will be discussed later in this section.

Table 7: *Hungarian high and low vowels*

high vowels	e, é, i, í, ö, ő, ü, ű
low vowels	a, á, o, ó, u, ú

Huang and Liao (1991) list thirty-nine finals in Chinese language. According to the number of vowels and with or without a nasal consonant, they can be divided into four groups as follows:

monophthongs (10): a, o[110], e, ê[111], i, u, ü, -i(zi/ci/si)、-i(zhi/chi/shi/ri), er;

diphthongs (9): ai, ei, ao, ou, ia, ie, ua, uo, üe;

triphthongs (4): iao, iou, uai, uei;

nasal finals (16): an, en, ian, uan, üan, in, uen, ün, ang, iang, uang, eng, ing, ueng, ong, iong.

In the traditional way, the final is divided into a medial, a main vowel and a final/ending. Not every syllable has a medial or an ending, only the main vowel is obligatory. To take "*普通话 pǔtōnghuà*" and "*想 xiǎng (think)*" for examples, the details are shown on the table:

Table 8: *Chinese medial, main and final vowels*

syllable	medial	main	final
pǔ		u	
tōng		o	ng
huà	u	a	

110 This sound occurs only in interjections

111 This sound occurs only in interjections

xiǎng	i	a	ng

3.3.1. The contrastive analysis of monophthongs in Hungarian and Chinese

There are fourteen vowels in Hungarian and seven vowels in Chinese.

A pair of vowel charts (based on the “quadrilateral” or “trapezium” method of depiction devised by the British linguist Daniel Jones (1917) can be used to present every vowel both in Hungarian and Chinese. Note that high vowels are at the top and low vowels at the bottom, back vowels are on the right and front vowels on the left. A contrast between rounded and unrounded vowels in the same position cannot be easily represented: they appear in the same position on the chart.

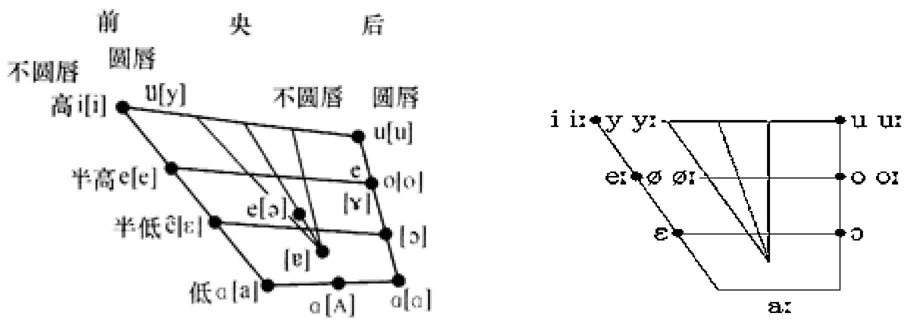


Figure 1. Vowel charts of Mandarin Chinese (left) and Hungarian (right)[112].

The discriptions of the vowels in Hungarian are shown on the following table, based on Siptár & Törkenczy (2000)[113].

Table 9: *Hungarian vowels according to lip-shape and tongue-position.*

Vowel and IPA	Shape of lips	Tongue position	
	Rounded/ Unrounded	Front/Back	High/Low

[112] The author is indebted to Dr. Katalin Mady of the Theoretical Linguistics Institute of the Hungarian Academy of Sciences for supplying the Hungarian vowel chart.

[113] Siptár, P. & Törkenczy, M. 2000.

a [ɔ]	Rounded[114]	Back	Low-mid
á [a:]	Unrounded	Central	Low
e [ɛ]	Unrounded	Front	Low-mid
é [e:]	Unrounded	Front	High-mid
i [i]	Unrounded	Front	High
í [i:]	Unrounded	Front	High
o [o]	Rounded	Back	Mid
ó [o:]	Rounded	Back	High-mid
ö [ö]	Rounded	Front	Mid
ő [ö:]	Rounded	Front	High-mid
u [u]	Rounded	Back	High
ú [u:]	Rounded	Back	High
ü [ü]	Rounded	Front	High
ű [ü:]	Rounded	Front	High

The descriptions of the vowels in Chinese are shown in the following table, which is based on Huang and Liao (1991)[115]:

Table 10: *Chinese vowels according to lip-shape and tongue-position*

Vowel and IPA	Shape of lips	Tongue position	
	Rounded/ Unrounded	Front/Back	High/Low
a [A]	Unrounded	Central	Low
o [o]	Rounded	Back	High-mid
e [ɤ]	Unrounded	Back	High-mid
i [i]	Unrounded	Front	High
u [u]	Rounded	Back	High
ü [y]	Rounded	Front	High
er [ɚ]	Unround	Central and curl	Central

Note that according to Huang and Liao (1991) /a/ has two allophones, [a] and [ɑ]; /e/ has three allophones, [ə], [ɛ] and [e]; /i/ has two allophones, [ɿ] and [ʅ]. Let us now consider the phonetic conditions under which these allophones occur (note that o [o], u [u], ü [y] and er [ɚ]

[114] According to Siptár & Törkenczy 2000, though not all speakers of Hungarian agree.

[115] Siptár, P. & Törkenczy, M. 2000; Huang and Liao 1991.

do not have allophones in Chinese):

Table 11: *Allophones of /a/, /e/ and /i/ in Chinese*

Vowel	Allophone	Phonetic conditions
/a/	[a]	/ai/ [ai], /an/ [an]
	[Λ]	/a/ [Λ], /ia/ [iΛ], /ua/ [uΛ]
	[ɑ]	/ao/ [ɑu], /ang/ [ɑŋ]
/e/	[e]	/ei/ [ei]
	[ε]	/ie/ [iε], / üe/ [yε]
	[ə]	/en/ [ən], /eng/ [əŋ]
	[ɤ]	/e/ [ɤ]
/i/	[i]	/ji/ [tɕ], /qi/ [tɕ‘], /xi/ [ɕ]……
	[ʅ]	/zi/ [ts], /ci/ [ts‘], /si/ [s]
	[ʅ]	/zhi/[tʂ] , /chi/ [tʂ‘], /shi/ [ʂ], /r/ [ʐ]

In order to further analyze the phonetic systems of Hungarian and Chinese, it will be useful to present a comparative table of the vowels of these two languages, based on Huang and Liao, (1991) and Siptár & Törkenczy (2000)[116].

Table 12: *Comparison of vowels in Hungarian ("H") and Chinese ("C")*

		Blade vowel					Retroflex-ed vowel	Apical vowel	
		Front		Central	Back		Central	Front	Back
		Unround-ed	Round-ed	Unround-ed	Unround-ed	Round-ed	Unround-ed		
High	H	i [i], í [i:]	ü [y], ű [y:]			u [u] ú [u:]			
	C	i [i]	ü [y]			u [u]		-i [ʅ]	-i [ʅ]
High-mid	H	é [e:]	ő [ø:]			ó [o:]			
	C	e [e]			e [ɤ]	o [o]			
Mid	H		ö [ø]			o [o]			

[116] Huang and Liao 1991.

	C			e [ə]			er [ɚ]		
Low -mid	H	e [ɛ]				a [ɔ]			
	C	e [ɛ]							
Low	H			á [a:]					
	C	a [a]		a [A]	a [ɑ]				

The above table summarises the similarities and the differences between the vowels in the two languages. We will now describe the salient differences in detail.

3.3.1.1. High vowels.

Hungarian has [i, y, u, i˥, y˥, u˥]; the Chinese equivalents are [i, y, ɿ, ʅ, u]. The issue of vowel length was mentioned earlier as a systemic difference between the two languages, and will be further discussed below. Apart from this feature, [I, y, u] are exactly alike in the two languages. Chinese [ɿ, ʅ] are different, but they are easily perceived as being different from both [i] and [y] by Hungarian learners, and it is important to teach the latter that in order to pronounce them properly the tongue must simply be kept in the same position as when the preceding consonant (alveolar and retroflex, respectively) is pronounced.

3.3.1.2. Mid vowels

Hungarian has [E, ↵, o, e˥, O˥, o˥], while Chinese has [E, e, o, Y, Φ, ↔]. Some of these are very similar; for example Hungarian [E] corresponds to Chinese 月 [yE], and [o] to 多 [δ8Yo], but there are some key differences and problems here. First of all, the vowel of the *-ong* rime falls halfway between Hungarian [o] and an [u] (IPA [Y]). Hungarians tend to perceive it as [u], though once they have been made aware of the difference they usually have little difficulty in pronouncing the Chinese sound. On the other hand, [Φ] can cause serious problems: Hungarians usually perceive it as something close to their own [↵] sound, but in fact the two sounds are very different: the front rounded [↵] is quite unlike the back unrounded [Φ]), and once Hungarian learners fall into the habit of pronouncing [Φ] as [↵]) is formed, it is probably the hardest vowel mispronunciation to correct.

3.3.1.3. Low vowels

Hungarian has [, a⌉][117], while Chinese has [a, ℘, A]. Hungarians tend to equate both Chinese [a] and [℘] with their own [a⌉] in Hungarian, and this causes no problems in understanding as the difference is relatively small, and there is no chance of confusion with any other phoneme: the issue has more to do with the esthetics of pronunciation. Hungarians are also often familiar with [a] because of their knowledge of foreign languages where it occurs and because of its presence in loan-words that have become part of their own language. Problems can also occur with Chinese [A] (as in 王 [wAN]): Hungarians learners tend to equate it either with [a] or with [ ]. In both cases this can result in “ugly” pronunciation, but not in misunderstanding.

3.3.2. Vowel length

One of the main systemic differences between Hungarian and Chinese, referred to above, is that in the former vowel length is contrastive (compare, for example, tör [t↵r], meaning “break” or “burst” with tőr [tOr], meaning “dagger”). In Chinese vowel length carries no meaning and is purely derivative: a vowel is long if and only if it is a syllable-final monophthong in a stressed syllable. Although this may be considered a major difference, and can be expected to cause communication problems for Chinese speakers learning Hungarian, it does not cause such difficulties for Hungarians speaking Chinese: most long Chinese vowels exist in Hungarian as well: [i⌉, y⌉, o⌉, u⌉], the two exceptions being [Φ⌉], and [℘⌉]. The latter is usually replaced by [a⌉] by Hungarian learners which leads to “ugly” pronunciation but not to misunderstanding, but the former needs to be practised, especially since its short form is also difficult for Hungarian learners, as already mentioned.

3.3.3. Syllable-initial vowels

The last “individual sound difference” to be noted is that concerning syllable-initial vowels, where the syllable does not begin with a consonant (声母). In Chinese the pronunciation of such vowels begins with the glottis open (‘smooth’ initiation), while in Hungarian they are initiated with a closed glottis (‘abrupt’ initiation). This explains why Hungarian learners are liable to hear an initial [j] in *yi* 以 or *yu* 鱼, as if they were *[ji⌉],*[jy⌉], rather than just [i⌉]

[117] Note that in present-day Hungarian [ ] tends towards being pronounced as [ ].

and [y˥], and it should be noted that *pinyin* spelling reinforces the confusion, denoting the 'smooth' initiation with an extra letter: *y*-). The situation is rather similar for *wu* 五, which is erroneously perceived by Hungarian learners as *[wu˥] or even *[vu˥], instead of [ʷu˥]. However, although this represents a distinct difference, it seldom leads to actual miscomprehension.

3.3.4. The contrastive analysis of compound vowels in Hungarian and Chinese

As we mentioned before, there are thirteen compound vowels in Chinese, nine dipthongs and four triphthongs. Strictly speaking, standard Hungarian has continuous vowels, but not normally compound vowels, although diphthongs do appear in some loan-words. Compound vowels and continuous vowel are totally different concepts. A compound vowel is a part of a syllable and cannot be divided. For example, in 好 *hǎo* (good), 有 *yǒu* (to have), 家 *jiā* (home), "ao", "ou" and "ia" are compound vowels. A continuous vowel occurs when two vowels appear side by side in a sequence and can be divided into two syllables. For instance, Hungarian *tea* (tea), *piac* (market), *diák* (student), *fiatal* (young) can be divided into syllables, *te-a*, *pi-ac*, *di-ák*, *fi-a-tal* respectively. The pronunciations of a compound vowel and a continuous vowel are also different. Take "ia" for example, as a continuous vowel, "i" and "a" both keep their own characteristics, the durations of the two sounds are the same. As a compound vowel, "i" and "a" can not be divided. In a same syllable, when "ia" is pronounced, the position of tongue slides from "i" to "a", and the durations of the two sounds are not the same. The "i" is short and weak, the "a" is long and strong. Note that there are three types of compound vowels in terms of the position of the main vowel. They are falling diphthongs, rising diphthongs and middle-rising triphthongs with a main vowel in the middle, as shown in the following table:

Table 13: *Compound vowels in Chinese*

compound vowel	Members	Type	Explanation
Diphthongs	ai, ei, ao, ou	falling diphthong	The former is strong, the latter is weak.
	ia, ie, ua, uo, üe	rising diphthong	The former is weak, the latter is strong.

Triphthongs	iao, iou, uai, uei	Middle-rising triphthong	The main vowel is in the middle, so the middle part is strong.

The differences between a continuous vowel and a compound vowel are illustrated in the following table.

Table 14: *Compound and continuous vowels in Chinese*

Name	Tongue-position feature	Type	Loudness/Strength
continuous vowel (H)	fixed	V V	equal
compound vowel (C)	slide	V v	not equal
		v V	
		v V v	

3.4. Tones and syllable structure in Chinese

3.4.1. Tones

Before we move on to the comparison of syllable structures in Hungarian and Chinese, it is worth considering syllabic tone, a well-known characteristic of Chinese (and of many other languages) which does not exist at all in Hungarian.

The tone, which pertains to the entire syllable and distinguishes different meanings, is primarily characterized by voice pitch contour, although length and intensity also play a role in its perception. There are four regular tones in the standard Chinese language: the first tone (or *Yīnpíng)*, the second tone (or *Yángpíng*), the third tone (or S*hǎngshēng*) and the fourth tone (or *Qùshēng*).

The system devised by Chao (1983)[118] for describing the four tones is as as follows (the “neutral tone” or *Qīngshēng,* will be discussed separately):

Pitch is plotted on a vertical scale which covers the normal voice (pitch range) of a speaker.

[118] Chao 1983.

The scale is devided into five points, such that 1 shows the lowest point (the lower limit of a speaker’s normal voice range) and 5 the highest (the upper limit of a speaker’s normal voice range); 3 is mid pitch, 2 half-low and 4 half-high. From 1 to 5, means from the lowest to the highest. A tone can be described by indicating its beginning and ending point（If it is a falling-rising tone, the point which connects falling and rising is also indicated. The four tones can be described as follows:

The first tone, T1 (also called *Yīnpíng*) is at a continuous high level. The value of tone in the Chao system is 55, which means it starts at point 5 and ends at the same level 5. In *pīnyīn* the diacritical mark of the first tone is ˉ.

The second tone T2, also called *Yángpíng*, is at mid-to-high level. The value of tone in the Chao system is 35, which shows it begins at mid pitch 3, and then rises to the highest pitch 5. The diacritical mark of the second tone is ˊ.

The third tone T3, also called *Shǎngshēng*, begins at the half low point 2 and falls to the lowest level 1, then rises to the half high level 4. The value of the tone in the Chao system is therefore 214 and its diacritical mark is ˇ.

The fourth tone T4, also called *Qùshēng*, starts at the highest level and falls to the lowest. The value of tone is therefore 51 and its diacritical mark is ˋ.

The following table shows the details of the four tones:

Table 15: *The four main syllabic tones in Chinese*

type of tone	value of tone	shape of the tone	diacritic	examples
The first tone T1 (*Yīnpíng*)	55	high level	ˉ	摸 *mō* (touch), 星 *xīng* (star)
The second tone T2 (*Yángpíng*)	35	high rising	ˊ	桃 *táo* (peach), 红 *hóng* (red)
The third tone T3 (*Shǎngshēng*)	214	falling and rising	ˇ	马 *mǎ* (horse), 水 *shuǐ* (water)
The fourth tone T4 (*Qùshēng*)	51	high falling	ˋ	坏 *huài* (bad), 二 *èr* (two)

They may be represented graphically, as in the following figure, where they are exemplified by the four different words *mā* (mother 妈), *má* (numb 麻), *mǎ* (horse 马) and *mà* (scold 骂):

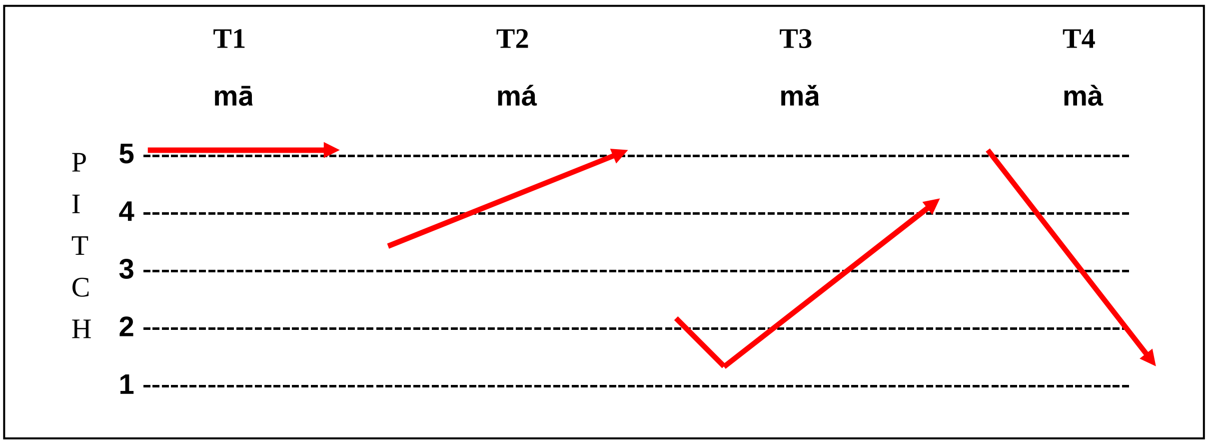


Figure 2. Graphical representation of the four tones.

3.4.2 The "neutral tone" or *Qīngshēng*

The "neutral tone" or *Qīngshēng* is a somewhat controversial feature. In *Xiandai Hanyu* (1991) Huang and Liao argue that *Qīngshēng* should not be regarded as the fifth tone, but rather as the weakened form of tones. In fact, *Qīngshēng* cannot exist in a single syllable; it occurs only in combinations of syllables such as words or phrases, and it does not have a fixed pitch. According to Huang and Liao (1991), "Generally speaking, the *Qīngshēng* which occurs after a third tone (or *Shǎngshēng*) usually has relatively high pitch (at point 4 in the pitch range), the one that follows a first (*Yīnpíng*) or second (or *Yángpíng*) tone has relatively lower pitch (at point 2 or 3, respectively), and the one that appears after a fourth tone (or *Qùshēng*) therefore has the lowest pitch (at point 1)."[119]

Huang and Liao also list the categories of morphemes in which *QīngShēng* occurs, which are as follows:

- Particles "的 *de*、地 *de*、得 *de*、着 *zhe*、了 *le*、过 *guo*"and interjections "吧 *ba*、嘛 *ma*、呢 *ne*、啊 *a*";

---

[119] Huang and Liao 1991.

- Reduplicated words and the second verb which exist in a Overlapped verb, such as “娃娃 *wáwa*、弟弟 *dìdi*、看看 *kànkan*、玩玩 *wánwan*”;
- Word suffixes like “子 *zi*、头 *tou*”and the word “们 *men*” which indicates plurality;
- Words which indicate directions such as “来 *lai*、去 *qu*、起来 *qĭlai*、下来 *xiàlai*”;
- The measure word “个 *ge*”;
- Orientation morphemes or words which occur after nouns or pronouns;
- The second syllable in some very commonly used bisyllabic words is customarily uttered with *Qīngshēng*.

Qīngshēng can carry semantic value. For examples, 东西 *Dōngxī* means “east” and “west”, but 东西 *Dōngxi* (with *Qīngshēng* on the second syllable) refers to “thing”.

The importance of tone in both the receptive and the productive use of Chinese is indicated by the existence of the four words used in the preceding figure (*mā 妈, má 麻, mǎ 马* and *mà 骂*), which have completely different meanings but are distinguished only by their syllabic tones. Nor should we forget that there is yet another “ma 吗” , the *Qīngshēng* version without a tone, which does not occur in isolation but can be added to the end of a declarative sentence to make it into a question. Such extreme cases are relatively rare, but it is not difficult to find examples of minimal pairs distinguished by tone, such as: *mǎihuà* 买画（to buy paintings）and *màihuā* 卖花（to sell flowers); *dōngxi* 东西（things) and *dōngxī* 东西（east and west); or *jiéshù* 结束 (to finish) and *jièshū* 借书 (to borrow books)

3.4.3 Sandhi (*biàndiào*)

As we have seen, standard Chinese has four tones: the first tone (or *Yīnpíng)*, the second tone (or *Yángpíng*), the third tone (or S*hǎngshēng*) and the fourth tone (or *Qùshēng*). Generally speaking, a Chinese syllable corresponds to a Chinese character, so as Huang and Liao (1991) point out, tones are also called “character tones (*zìdiào*)”. However, as Huang and Liao also note, single tones affect each other in words, phrases and sentences. This phenomenon is called “sandhi (*biàndiào*)”. Sandhi (*biàndiào*) takes various forms, including: sandhi of S*hǎngshēng*, sandhi of *Qùshēng*, sandhi of “一 *yī*” and “不 *bù*”, sandhi of reduplication adjectives. One very common example occurs in the greeting phrase 你好 *nǐhǎo*. Both of the words in the phrase carry the third (or S*hǎngshēng*) tone, and in isolation they are pronounced accordingly. But when they are combined in the greeting phrase the sandhi effect causes the word *nǐ* to be pronounced *ní* instead. For sandhi of *Qùshēng*, in the combination of *Qùshēng*, like 介绍 *jièshào (introduce)*, the value of the first word *jiè* should be 53, rather than 55.

3.4.2. Syllable structure in Hungarian and Chinese

The syllable is the basic unit of a phonotactic system. One or more phonemes constitute a syllable.

A Hungarian word can be divided into syllables, each of which contains a vowel. In other words, how many syllables a word contains is decided by how many vowels it has. A syllable which ends with a vowel is defined as an open syllable, such as, *ceru<u>za</u>* (pencil), *ze<u>ne</u>* (music), *aj<u>tó</u>* (door), etc. And a syllable which ends with a consonant is called a checked syllable, for example, *i<u>gen</u>* (yes), *vo<u>nat</u>* (train), *kö<u>vér</u>* (fat), etc. A syllable can consist of a vowel, or a vowel and a consonant, or a vowel and more than one consonant. There are no words in Hungarian which do not contain at least one vowel.

A Hungarian syllable can contain an onset, a nucleus, and a coda. The onset is not compulsory; therefore, both vowel-initial and consonant-initial syllables are possible. Disregarding the possible complexity of the onset, the nucleus and the coda, the basic types of Hungarian syllable are as follows:

Table 16: *Syllable patterns in Hungarian*

	Word-initial	Word-medial	Word-finial
CV	**ce.**ru.za (pencil) **Kí**.na (China)	fe.**ke**.te (black) vi.**lá**.os (bright)	ka.to.**na** (soldier) sző.**lő** (grape)
V	**a**.pa (father) **í**.ró (buttermilk)	fi.**a**.tal (young) i.di.**ó**.ta (idiot)	szi.**a** (hello) rá.di.**ó** (radio)
VC	**asz.**tal (table) **é.**let (life)	a.**or.**ta (aorta) ki.**ál**.tás (shout)	i.**tal** (drink) di.**ák (**student)

CVC	**lec**.ke (lesson) **lám**.pa (bulb)	ke.**men**.ce (oven) ta.**nár.**nő (female teacher)	ti.**los** (forbid) kí.**ván** (to hope)

The above table reveals that in Hungarian:

(1) Any type of syllable can occur in any position in the word, no matter whether in initial, medial or final position.

(2) The distribution of long and short vowel syllables is the same in the word.

(3) Neither open syllables nor closed syllables are restricted to word-final position.

Note that the examples listed above are polysyllables; monosyllables also exist in Hungarian, such as *ír* (to write), *jó* (good), *és* (and), *nap* (day), *fal* (wall) etc. In addition, besides the four basic syllable types, there are some other extra types due to the existence of compound consonants. Two-member and three-member consonant clusters occur in word-initial position, but only in loan-words. A selection of loanwords with two-member and three-member initial consonant clusters is shown in the following tables, based on Siptár & Törkenczy (2000)[120], which show all the possible initial CC[121] and CCC clusters:

Table 17: *Word-initial CC syllable clusters in Hungarian*

Word-initial CC clusters	Examples
pt	ptózis (ptosis)
psz	pszichológus (psychologist)
pn	pneumatikus (pneumatic)
pl	plakát (poster)
pr	prém (fur)
tv	tviszt (twist)
tr	tréfa (joke)

[120] Siptár, P. & Törkenczy, M. 2000.

[121] Here “C” represents a consonant, “V” represents a vowel.

ks	xilofon (xylophone)
kv	kvarc (quartz)
kn	knédli (dumpling)
kl	klór (chlorine)
kr	krém (cream)
bl	blúz (blouse)
br	bronz (bronze)
dz	dzéta (zeta)
dr	drukkol (cheer)
gv	gvárdián (Father Superior)
gn	gnóm (gnome)
gl	gladiator (gladiator)
gr	gróf (count)
cv	cvekedli (type of pasta)
ft	ftálsav (phthalic acid)
fl	flóra (vegetation)
fr	friss (fresh)
fj	fjord (fjord)
szp	szpícs (speech)
szt	sztár (star)
szty	sztyepp (steppe)
szk	szkíta (Scythian)
szc	szcéna (scene)
szf	szféra (sphere)
szv	szvetter (sweater)
szm	szmog (smog)
szn	sznob (snob)
szl	szláv (Slav)
sp	sport (id)
st	stáb (staff)
sk	skorpió (scorpion)
scs	scsí (Russian soup)
sv	svéd (Swedish)

sm	smink (makeup)
sn	snassz (passé)
sl	slussz (finished)
sr	sróf (screw)
vl	vlach (Walachian)
zl	zlotyi (Polish currency)
zr	zrí (trouble)
mn	mnemonika (mnemonics)
ng	nganaszán (Nganasan)
hr	Hradzsin (proper noun)

Table 18: *Word-initial CCC syllable clusters in Hungarian*

Word-initial CCC clusters	Examples
Sztr	sztrájk (strike)
Szkl	szklerózis (sclerosis)
Spr	spriccel (spray)
Str	strázsa (guard)
Skr	skrupulus (scruple)

Thus, there are two more types of word-initial which are CCV/CCCV and CCVC/CCCVC.

Chinese is a monosyllabic language, where almost every syllable corresponds to a morpheme. A syllable consists of an initial, a final and a tone. The initial is the consonant which is at the beginning of the syllable. For example, "普通话 *pǔtōnghuà*" has three syllables, the initials of which are p, t, and h, respectively. Chinese has twenty-two consonants, all of which can occur as initials except for "ng ", which can only be used as the ending of a final, as in "听 *tīng*" (listen), "中 *zhōng*"(middle) and so on. In other words, there are twenty one consonantal initials in Chinese language. They are as follows: b, p, m, f, d, t, n, l, g, k, h, j, q, x, zh, ch, sh, r, z, c, s. There are some syllables which have no consonant as an onset, they are defined as "zero-initial syllable". For instance, "爱 *ài*" (love) or "儿 *ér*" (son). Note that in *pinyin* the letters "y" and "w" only occur at the beginning of a zero-initial syllable, as in "一 *yī* (one) " or

"五 *wǔ*" (five) ", but they do not represent initials.

The final is the remainder of the syllable minus the initial and the tone. A final can be a single vowel or can consist of more than one vowel, as in "老鼠 *lǎoshǔ*" (mouse). There is only a single vowel "u" functioning as the final in the syllable 鼠 shǔ, but there is a combination of the two vowels "a"and "o" serving as the final in the syllable "老 *lǎo*". Some finals also contain a consonant, as in "冬天 *dōngtiān*" (winter), where the finals include the consonants "ng" and "n".

The various types of syllable structure in Chinese can be roughly divided into two categories: with an initial and without an initial. Then, according to the distribution of consonants and vowels, they can be classfied as more detailed types. See the following tables[122]:

Table 19: *Syllable patterns in Chinese (with initial)*

Types of syllable structure	Examples
CV	妈 *mā* (mother); 米 *mǐ* (rice)
CVG	白 *bái* (white); 飞 *fēi* (fly)
CGVG	快 *kuài* (fast); 票 *piào* (ticket)
CVC	看 *kàn* (look); 很 *hěn* (very)
CGVC	脸 *liǎn* (face); 穷 *qióng* (poor)

Since zero-initial syllables exist in Chinese phonetic system, the types of syllable structure can also be as follows:

Table 20: *Syllable patterns in Chinese (without initial)*[123]

Types of syllable structure	Examples
V	啊 *a* (interjection)；哦 *o* (interjection)
VG	爱 *ài* (to love)；傲 *ào* (proud)
GVG	有 *yǒu* (to have)*；外 *wài* (outside)*

[122] As above, "C" represents a consonant, "V" represents a vowel. "G" represents a glide, a shorter vowel that immediately precedes or follows the main vowel of a syllable.

[123] Note that originally you, wai, yuan, wen are iou, uai, uan and uen, respectively.

VC	暗 *àn* (dark)； 昂 *áng* (expensive)
GVC	远 *yuǎn* (far)*； 问 *wèn* (to ask)*

3.4.3. Contrastive analysis of syllable structure in Hungarian and Chinese

From the introduction of syllable structure in Hungarian and Chinese presented above, we can make a contrastive table, comparing the available syllable structures in the two languages.

Table 21: *Comparison of Hungarian and Chinese syllable patterns*

	Distribution of consonants and vowels	tone
Hungarian	CV, V, VC, CVC, CCV, CCCV, CCVC, CCCVC	-
Chinese	CV, CVG, CGVG, CVC, CGVC, V, VG, GVG, VC, GVC	+

The table presented above summarises the similarities and differences between syllable structure in Hungarian and Chinese. The main similarities are as follows:

(1) Both Hungarian and Chinese syllables can begin with a vowel.

(2) The two languages share some same types of distribution of consonants and vowels, such as CV, V, VC, CVC, though it should be noted that the range of consonants that can occur in final position is more limited in Chinese than in Hungarian.

(3) A vowel is normally obligatory in a syllable in both Hungarian and Chinese, though the latter language contains a few exceptions in the form of interjections which consist only of consonants, such as 嗯 *ng*.

As for the differences, they are as follows:

(1) Chinese is a tone language while Hungarian is not. This is one of the most notable differences between these two languages. Every syllable in Chinese has a tone, even if this is the neutral tone (*QīngShēng*) described in 4.4.1 above.

(2) Chinese is a monosyllabic language, where almost every syllable corresponds to a morpheme. Although monosyllables do exist in Hungarian, it is not a monosyllabic language.

(3) Compound consonants (consonant clusters) exist in Hungarian, where they can occur in

initial, mid- or final position, but they are absent in Chinese.

(4) The maximum number of phonemes that a Chinese syllable can contain is four, for example, the CGVG and CGVC types; the minimum number of phonemes in a Chinese syllable is one, such as the V type. In Hungarian the number of phonemes in a syllable ranges from one (e.g. the V-type in *ó* "old") to six or more, especially in loan-words with initial consonant clusters such as CCCVVC *sztrájk* "strike".

3.5. Stress and intonation in Hungarian and Chinese

3.5.1. Stress in Hungarian and Chinese

Stress is a feature of pronunciation: the opposite of weakening. It refers to the degree of force used in producing a syllable. In most languages two kinds of stress can be distinguished: word stress and sentence stress. Every multisyllable word when it is read separately may have its own stress pattern, which is called word stress, while in a sentence, some (relatively important) words are stressed and others are not. This phenomenon is therefore called sentence stress.

Concerning word stress in Hungarian. Siptár and Törkenczy (2000, 21) point out that "In its citation form, a Hungarian word typically has a single primary stress, which falls on its initial syllable, no matter whether the word is simple (e.g. ***is****kola* "school") or derived (e.g. ***for****rósodik* "grows hot") or a compound (e.g. ***szé****nanátha* "hay fever")"[124].

Meanwhile, there are two major types of stresslessness: spontaneous enclisis and stress eradication. The following examples which describe and compare these two types are from Kálmán and Nádasdy (1994)[125]:

(1) a. 'Géza 'táncolni *akar*

to-dance want

"Géza wants to dance"

[124] Siptár, P. & Törkenczy, M. 2000: 21.

[125] Kálmán, L. & Nádasdy, Á. 1994.

b. 'Géza 'táncolni *akar* a 'magas 'fekete 'lánnyal.

the tall black girl-with.

"Géza wants to dance with the tall black(-haired) girl"

c. 'Géza *bácsi*

"Uncle Géza"

d. 'Géza *bácsi* 'táncolni *akar* a 'magas 'fekete 'lánnyal.

"Uncle Géza wants to dance with the tall black(-haired) girl"

(2) a. 'Jenő 'táncolni *imád*

to-dance love

"It is to dance that Jenő loves"

b. 'Jenő 'táncolni *imád* a magas fekete lányyal.

"It is to DANCE with the tall black girl that Jenő loves"

c. 'Jenő 'táncolni *akar*

"It is to dance that Jenő wants"

d. 'Jenő 'táncolni *akar* a magas fekete lányyal.

"It is to DANCE with the tall black girl that Jenő wants"

In (1), the italicized words "*akar*", "*bácsi*" are enclitic, they join the stress domain of the previous word. The stress on "táncolni" eradicates the rest of the lexical stresses in its whole domain in (2). With regard to stress eradication, Siptár and Törkenczy (2000) indicated that "two important facts about eradicating stress are that it need not be stronger than a non-eradicating stress; and that it cannot be followed by another stress within the same sentence unless that other stress is also of the eradicating type. A sentence with no eradicating stress is said to have flat prosody, corresponding to neutral interpretation; a sentence with eradicating prosody has a contrastive or emphatic interpretation." [126]

Chinese is, as we have seen, a tone language. To some extent, tone is such a salient and distinguishable feature of the Chinese language that the latter is often classified as a non-intonation language. According to Jerry Norman's *Chinese* (1988), "Some people apparently think that pitch cannot function at the lexical level (tone) and at the syntactic level (intonation) at the same time". This is of course not true: "In fact, in addition to tone, Standard Chinese possesses both stress and intonation"[127].

The rule of the stress of Chinese words is relatively clear. As Luo and Wang (2002) point out, "About disyllabic words, there is a primary stress on the second syllable, and the first syllable is relatively light, except for those which end with a neutral tone", such as 老师 *lǎoshī* (teacher), 汉语 *hànyǔ* (Chinese language), 中国 *zhōngguó* (China)[128]. Based on Luo and Wang's statement, Lu (2010) further indicates, "As for Chinese words or phrases which consist of three or more than three syllables, the primary stress is on the last syllable"[129], such as 普通话 *pǔtōnghuà* (Mandarin), 语言教师 *yǔyán jiàoshī* (language teacher), 中华人民共和国 *zhōnghuá rénmín gònghéguó* (People's Republic of China). The words or phrases which end with a neutral tone syllable are the exceptions, because the primary stress is on the syllable which is next to the neutral tone syllable, for example, 妈妈 *māma* (mother), 你好吗 *nǐhǎo ma* (how are you?), and 非常喜欢 *fēicháng xǐhuan* (like s.thing very much).

The primary stress mentioned above operates at lexical level and indicates a possible locus for stress. Whether or not the syllable concerned actually contains primary stress in a sentence depends on syntactic structure and on intended meaning. In other words, it depends on the relative importance of a word. The more important a word is, the stronger is its stress. In Hungarian, content words such as nouns, adjectives, main verbs, adverbs, and demonstrative and interrogative pronouns are likely given more stress and energy. Other categories of words like auxiliary verbs, conjunctions, prepositions, etc. are usually unstressed. The same is true

[126] Siptár, P. & Törkenczy, M. 2000: 21.
[127] Norman, J. 1988.
[128] Luo and Wang 2002: 156-157.
[129] Lu 2010.

of Chinese. When we speak Chinese, we tend to emphasize the content words including all the pronouns, by uttering them with more stress, but not the grammatical or function words.

3.5.2. Intonation in Hungarian and Chinese

To use the famous phrase coined by Roger Kingdon (1958), “the phoneme is the body of a language, yet intonation is the soul”[130]. Intonation plays a vital role in the correct and appropriate use of spoken language. So, what exactly is intonation? The concept has been described in various ways, ranging from Trask’s (1996) very broad definition which includes stress and rhythm under the general umbrella of intonation: “When speaking people raise or lower the pitch of their voice, forming pitch patterns, they also give some syllables in their utterance a greater degree of loudness and change their speech rhythm. These phenomena are called intonation”[131], to the very limited version offered by The Oxford Advanced Learner’s Dictionary of Current English with Chinese Translation (Hornby & Zhang, 1984) which defines intonation as “the rise and fall of the pitch of the voice in speaking”[132]. For our purposes, we shall use the latter definition, adding that we are using “intonation” to refer to phrase- or sentence-level rather than word-level patterns.

Intonation is an important way to express meanings and feelings. The same sentences with different intonations can carry very different messages. Mainly, the intonation of Hungarian has two types: falling tone and rising tone. Specifically, the intonation of narrative sentences contains a falling tone. The intonation of wh- interrogative sentences usually contains a falling tone at the end of the sentence, and a primary stress accompanied by a rising tone on the interrogative word as well. The intonation of yes/no questions is a little bit complicated: according to Siptár and Törkenczy (2000) it “involves a rise-fall pattern (LHL) which spreads over the last three syllables provided that the major stress occurs on the antepenultimate (or earlier) syllable of the utterance. Thus, given a question whose focus is well before the third-last syllable, a bisyllable final word will have a pitch on its initial syllable, whereas a

[130] Kingdon, R. 1958.

[131] Trask, R. 1996.

[132] Hornby, A. & Zhang, F. 1984.

trisyllable word will have one on its medial syllable"[133].

As regards intonation in Chinese, briefly, there are two kinds: falling and rising. With regard to the actual pitch of intonation itself, Jin Song (1992) distinguishes between high intonation, relatively low intonation and low intonation[134]. But as far as the available patterns and actual significance of intonation are concerned, there are several different opinions. According to the tone of comments, Chao Yuenren (1929) divides Chinese intonation into as many as forty different types[135]; Hu Mingyang (1987) lists eight kinds of intonation according to the final pitch pattern of a sentence, defining them according to function as follows: *statements, questions, commands, imperatives*, *expressions of amazement*, *sighs,* invocations and *pauses preceding continuation*[136]. Still working on a functional basis, Shen Jiong (1994) divides intonation into functional intonation and tone of comments intonation[137]. Lin Maocan argues that Chinese intonation has two variables: pitch accent and boundary tone; he claims that only boundary tone plays the role of differentiating between questions and statements. Whether the first tone, the second tone, the third tone, or the fourth tone, pitch pattern in boundary tone with question keeps its citation form[138].

Therefore, intonation of Chinese is complicated. It shares the basic features of falling pitch and rising pitch that are common to most languages, but the way in which these operate are open to discussion. What makes the intonation even more complicated in Chinese is the existence of syllabic tones. The relation between tone and intonation is always a heated topic for linguists. As we know, tone and intonation are two different concepts. The former pertains to the syllable, and is primarily characterized by voice pitch, although length and intensity also play a role in its perception; the latter, to use Trask's (1996) explanation, means "the use of pitch, and possibly of additional prosodic phenomena such as loudness, tempo and pauses, over a stretch of utterance generally longer than a single word for the purpose of conveying

[133] Siptár, P. & Törkenczy, M. 2000: 17.
[134] Jin 1992.
[135] Chao 1983.
[136] Hu 1987.
[137] Shen 1994.
[138] Lin 2004.

meaning."[139] To sum up, (1) the pitch is the determinant in both tone and intonation; (2) tone pertains to a syllable while intonation pertains to an utterance.

What, then, is the relationship between them? The most influential theory was put forward by Chao (1983), who used "small wave" and "big wave" as metaphors to describe tone and intonation, respectively. He proposed that the relation between tone and intonation is the algebraic sum of the "small wave" and the "big wave"[140]. Based on Chao's theory, Zongji Wu (1997) points out that the "algebraic sum" refers to the algebraic sum of register. In other words, the algebraic sum of "small wave" and "big wave" can be explained as the algebraic sum of the average pitch of tone and the average pitch of intonation, meanwhile, the shape of tone remains unchanged[141].

Although tone and intonation are independent concepts, in terms of actual pronunciation they are interrelated. On the one hand, intonation cannot be seperated from tone, in fact, it is shown through the pitch movement of tone. On the other hand, within the intonation of an utterance, syllabic tones basically remain unchanged, but the register and tone shape are restricted by the intonation.

The semantic value of suprasegmental intonation becomes clear if we consider, for example, the difference between .今天是你的生日。Jī tiān shì nǐ de shēngrì. (Today is your birthday.) and 今天是你的生日?(Is today your birthday?), in which the interrogative is indicated only by intonation. Further examples of semantic differences carried by intonation are 以前是喜欢一个人，现在是喜欢一个人。Yǐqián shì xǐhuan yígèrén, xiànzài shì xǐhuan yígèrén. ((I) liked a person before, now (I) like to be alone.); and 夏天能穿多少就穿多少，冬天能穿多少就穿多少 。Xiàtiān néng chuān duōshǎo jiù chuān duōshǎo, dōngtiān néng chuān duōshao jiù chuān duōshao. (In summer, wear as little as you can; In winter, wear as much as you can.)

3.5.3. The measurement and plotting of intonation

In general, intonation patterns can be described in terms of the linguistically significant pitch

---

[139] Trask, R. 1996.

[140] Chao1983.

[141] Wu 1997.

contours. In the charts produced by the application of PRAAT software (Boersma & Weenink, 2013) to relevant sound samples, pitch contours are described by the F0 (or Fundamental Frequency) contour[142]. Naturally, in Hungarian – a non-tone language - the rising and falling of pitch register corresponds only to rising and falling intonation, thus intonation determines the rising and falling of F0. In other words, in Hungarian (suprasegmental) intonation can be visualised using PRAAT, without the risk of interference by rising and falling syllabic tone. With a language like Chinese, characterised by both syllabic tone and suprasegmental intonation, the situation is less satisfactory, because tones prevail and determine the rising and falling of F0. Since F0 is "occupied" by tones, it cannot be used to present other changes in pitch and as a result intonation remains "invisible".

3.6. Categorization of differences by difficulty

The following table presents Prator's (1967) categorisation, which ranks differences between languages according to the degree of difficulty that they may be expected to cause to speakers of one language attempting to learn the other[143]. Prator includes six degrees, with 0 as the least problematic category and 5 as the most problematic. Whether these predictions are accurate will be examined in the relevant chapter of this thesis where the recorded data are presented and analysed using PRAAT software.

"Correspondence", referring to cases where there is no difference between the two languages and consequently no difficulty experienced by the learner, is ranked "0"; the "split" category, where a single item in the first language corresponds to two or more items in the target language, causes the most serious problems and comes at the top of the scale. Obviously the ranking of specific differences and similarities between languages depends on which language is "first" and which is "second": a problematic "split" from Hungarian to Chinese would correspond to a less problematic "coalescence" from Chinese to Hungarian and so on. It should also be noted that this is merely an ordinal, not an interval or ratio scale: it indicates

[142] Boersma, P. & Weenink, D. 2013.

[143] Prator, C. 1967.

for example that the “split” category is “more problematic” than the “new” category, but not “how much more problematic”. Finally, we should also remember that the degree of difficulty experienced by the learner in recognising and/or reproducing different sounds is not necessarily proportionate to the resulting problems in understanding or being understood by others. In cases where there is no risk of ambiguity the learner’s inability to perceive differentiated sounds will hardly matter; his inability to produce sounds accurately will lead to “aesthetic” rather than communicative problems. For further discussion of this issue and its relative importance for different groups of learners, see Chapter 1 “Introduction”.

Table 22: *Categories of difference by level of difficulty, based on Prator (1967)*

Level	Category	Explanation
5	split	One item in L1 is split into two or more in L2.
4	new	The item exists in L2 but is absent in L1.
3	reinterpretation	The item is present in L1 but appears in a new form in L2.
2	absent	The item exists in L1 but is absent in L2.
1	coalescence	Two or more items correspond to one in the L2.
0	correspondence	The items are the same in both L1 and L2.

Evidently categories 0, 1 and 2 are unlikely to be the direct causes of phonological problems for learners; the following summary of the problematic cases identified in the previous sections will therefore focus on categories 3, 4 and 5, with reference both to receptive (perceptual) features and productive (active pronunciation).

3.6.1. Split:

3.6.1.1. Vowels:

- The contrast between Hungarian [i] and Chinese [i, ɿ, ʅ]. Hungarian learners have little difficulty in recognising the difference, and are generally aware of what they need to produce, but find it hard to actually do so. This may lead to awkward or “odd” pronunciation, but seldom causes misunderstanding.

- The contrast between Hungarian [a] and Chinese [a, ℘]. Hungarian learners do have real difficulty in distinguishing between these, but the difference is neither phonemic nor salient in Chinese, so the difficulty is not likely to cause communication problems.
- The contrast between Hungarian [u] and Chinese [u, Y]. As with the previous item, this causes few problems because although Hungarians find it hard to differentiate, in Chinese the difference is neither phonemic nor salient.

3.6.1.2. Consonants:

- The contrast between Hungarian voiceless stops/affricates and Chinese aspirated and unaspirated pairs. This is relatively easy to distinguish, both in production and in perception, because of salience of the articulatory differences.

3.6.2. New:

3.6.2.1 Vowels:

- The Chinese [Φ] is difficult for Hungarians to produce, and hard to distinguish from [↔]: it is very difficult to prevent (or stop) Hungarian learners pronouncing it like their own [↵]. However, since the two sounds are allophones in Chinese, no misunderstanding is likely to result.
- The Chinese retroflecized schwa [™] is very difficult to master for Hungarian learners, who often replace it with the standard Hungarian sequence *[↵r], as in *sör* 'beer'. However, this is not too serious as the resulting problems are aesthetic rather than communicational.

3.6.2.2 Consonants:

- Chinese Retroflexes are difficult for Hungarians to produce accurately, and hard to distinguish perceptionally from Hungarian palato-alveolars, but are easy to distinguish from any other place-of-articulation in Chinese.

- Chinese *apda* consonants are very difficult to produce, and to distinguish from Hungarian dorso-palatals, but, like retroflexes, easy to distinguish from any other place-of-articulation in Chinese.
- The Chinese uvular approximant [Ξ] is not difficult for Hungarians to distinguish perceptionally, but is somewhat difficult to produce. However, its usual replacement by Hungarians learners with [h] does not lead to confusion.

3.6.2.3 Semi-vowels, diphthongs:

- Hungarian normally lacks proper semi-vowels or glides, and (with very few exceptions) has no diphthongs. As a result Chinese diphthongs are often replaced with sequences of the corresponding full vowels in Hungarian learners' speech. This in itself does not cause confusion, and CFL teachers at ELTE report that once the difference has been explained, their Hungarian students find it relatively easy to produce the requisite diphthongs.

3.6.2.4. Tones and syllable structure

- The very existence of syllabic tone falls into the "new" category because this feature is completely absent in Hungarian. Once a speaker of a non-tone language has become familiar with the notion of the Chinese tone system, further difficulties may be caused by the existence of neutral tones and sandhi.
- Chinese is a monosyllabic language where each syllable constitutes a single morpheme, while the majority of Hungarian morphemes contain more than one syllable. This might tempt Hungarian learners to split a Chinese syllable that contains a diphthong into a sequence of vowel sounds. CFL teachers at ELTE report that although the results may sound "odd", they do not lead to confusion during communication.

3.6.2.5. Stress and Intonation

- The existence of syllabic tone tends to obscure the importance of suprasegmental intonation, to the extent that some Hungarian learners may find it difficult to perceive both and may concentrate on producing accurate tones in their speech, at the expense of intonation. This can cause problems in communication, given that for example in certain cases the difference beween a statement and a question may be indicated purely by intonation, without the use of a final “question word”.

3.6.3. Reinterpretation:

3.6.3.1 Vowels:

- The vowels [E, e, o] exist as independent phonemes in Hungarian, while in Chinese they only have allophonic status (realizing one and the same morpheme: “–high, –low V”) in a context-dependent fashion. (*Modulo* the other allophonic occurrence of [E] realizing the phoneme “+low V” in the context ‘i_n’, as in the rhyme -*ian*: [jEn]). This is likely to cause difficulties for Chinese speakers learning Hungarian, but not vice versa.

3.6.3.2 Consonants:

- The Hungarian palatal fricatives [Σ, Z] have the Chinese retroflex fricatives [♣, ⌈] as their nearest equivalents. Hungarians tend to replace the Chinese sounds with these “counterparts” from their own language. The result may be “ugly”, but it is unlikely to cause confusion.
- The velar nasal [N] exists in Chinese in its own right as a phonemic element, while in Hungarian it occurs only as an allophone of [n], most frequently in the sound sequence [Ng]. Hungarian learners need to be careful to pronounce it as a “bare” [N] and not as the full sequence *[Ng]. Once again, however, failure to produce the appropriate sound will result in awkward speech, but not in misunderstanding.

3.7. Choice of phonetic features for further consideration

On the basis of the previous section it is possible to identify the phonetic features of Chinese which are theoretically likely to cause difficulties to Hungarian learners. In Chapter 1 we discussed the issue of which potentially problematic features deserve particularly close attention on the part of teachers and learners of CFL. As we saw, it can be argued that the particular category of learner in focus in this study, the university student who is a future professional user of Chinese, should aim at the highest possible level of proficiency in every aspect of the language, including all of its phonetic features. The implication of this is that all the differences listed in 3.6 above require attention inasmuch as they are likely cause problems sooner or later, and perhaps in the long term this is the case. However, time and other resources available for TCFL are limited, and it makes sense to start, at least, by trying to establish which need to be dealt with first, in order to provide all learners with a sound foundation on which they can, given the necessary motivation and opportunities, build further progress.

The choice could be made according to various criteria. One approach, typical of traditional structuralist "grammar translation" language teaching methods, is to start with linguistic elements that can be regarded as "relatively easy", and once these elements have been acquired, to use them as "building bocks" from which to construct the rest of the language. While superficially attractive, this approach suffers from a number of weaknesses. Intuitive judgements (especially by native speakers) about what is "relatively easy" and what is "relatively difficult" are notoriously unreliable. Even if they are accurate - or if we accept the basic notion that the greater the "distance" between corresponding elements of two languages, the greater the likely difficulty, and therefore adopt the scale that emerges from Prator's (1967) categories[144]), there is no guarantee that a progressive structural syllabus based on them will be effective. In this respect we need only consider what Dulay and Burt (1974) demonstrated about the striking differences between the natural order of the acquisition of morphemes in English as a Foreign Language and planned text-book sequences[145].

A different, learner-centred view, would be imply the need to start with the learners' own priorities and expectations, originating partly in generally available knowledge about the target knowledge or skills (what might be described as "hearsay evidence") and partly in personal experience. Considering the vital importance of individual motivation and attitudes

[144] Prator, C. 1967.
[145] Dulay, H. C., & Burt, M. K. 1974: 24, 37-53.

in any learning process, this approach has much to recommend it; indeed, it has become a central pillar of the powerful humanistic ethos in education that reaches back to such great nineteenth-century figures as John Dewey, and beyond, and has more recently been successfully propagated by writers such as Carl Rogers (1969)[146]. To put it briefly in practical terms, a learner of Chinese (or anything else, for that matter), is unlikely to make much progress unless s/he believes that what s/he is expected to learn is worth the effort. However, this is not in itself a sufficient basis for a reliable language-learning program. The learner's choices may be unrealistic or partly mistaken, and it is part of the teacher's and/or material writer's task to form, or at least to guide, the learner in this respect.

In attempting to do this, the teacher will probably find utilitarian considerations to be useful - instrumental motivation can be created and fostered by pointing out the practical value of what needs to be learned, and this leads us to the third criterion for choosing phonetic features that deserve priority. The criterion may be summed up in the phrase "potential for communicative efficiency"; or in this case, since we are considering phonetic differences that maight cause difficulty "potential for avoiding communicative confusion". A choice made on this basis fulfils both the need to cater for the learners' perceived needs and the institution's educational responsibility for ensuring that students graduate from ELTE with (at the very least) an efficient working command of Chinese.

Looking at the categorised list of differences in 3.6 above, we find two that represent both potential learning difficulty and potential communicative inefficiency: syllabic tone (3.6.2.4.) and, closely linked to it, interference by syllabic tone in suprasegmental intonation (3.6.2.5). Both of them belong to the "new" category, and should therefore represent a roughly equal degree of difficulty from the point of view of the Hungarian learner. However, as we saw in 3.5.2 above, suprasegmental intonation in Chinese is a complex and debatable issue: the authors recognise its existence, but do not always agree over how it should be defined and described, and over how it interacts with syllabic tone. In experimental phonetics it is also difficult to plot Chinese intonation separately from syllabic tone, and therefore to analyse it

[146] Rogers, C. 1969.

satisfactorily. Above all, although intonation can certainly carry meaning, it seems to be less pervasively significant than syllabic tone, which affects virtually all Chinese words and very often carries semantic value.

To sum up, there would seem to be a clear case for choosing syllabic tone as the feature of Chinese phonetics that deserves the most attention from both teachers and learners, and it is this feature that will be prioritised in the remainder of this thesis.

4. Research Design and Procedures

4.1 Introduction

In the previous chapter, detailed descriptions were provided of the syllabic tone system of Chinese. In this chapter, we provide an overview of the methods by which Hungarian learners' pronunciation of the different tones of Chinese was recorded, analysied and compared with that of native speakers', in order to identify areas of difficulty.

4.2. Experimental Design and Procedure

4.2.1. Experimental Subjects

It is of great importance for both research on pronunciation and studies made on acoustics to obtain an adequate sample of subjects. People greatly differ from each other in terms of of their voice pitches, duration, range of the fundamental frequency, shape of the fundamental frequency curve, sound intensity and other acoustic properties. A set of people should be chosen as experimental subjects, so as to avoid the interference of personal characteristics and gain information about linguistic and phonetics, that is, both average or mean values and the extent to which those means represent a common, shared reality as opposed to the average of distant extremes: in statistical terms, the standard deviation (Rose 1993)[147]. According to Zhu Xiaonong (2010), at least 6 individuals should be chosen for an experiment which involves describing tones, and if possible, it would be better to choose 12 people for the experiment[148]. For the purposes of the current investigation, 10 students, including four females and six males, were chosen. All of them are students from the Department of Chinese of Eotvos Lorand University, who have studied Chinese in the department for a minimum of nine months and a maximum of three years. The four females were labeled from F1 to F4, while the six males were labeled from M1 to M6. The information on the experimental subjects is as follows:

[147] Rose, Phil 1981.
[148] Zhu 2010: 279.

Table 23: *Hungarian male and female subjects involved in the investigation.*

Articulator	Age	Educational Background	Time spent learning Chinese
F1	19	Hungarian	9 months
F2	21	Hungarian	2 years
F3	20	Hungarian	9 months
F4	19	Hungarian	9 months
M1	23	Hungarian	3 years
M2	23	Hungarian	3 years
M3	19	Hungarian	9 months
M4	20	Hungarian	9 months
M5	21	Hungarian	9 months
M6	22	Hungarian	3 years

In addition, as a reference, pronunciations of two people, one male and one female, from Beijing are recorded. The male was labeled MB and the female FB.

Table 24: *Chinese native-speaker subjects involved in the investigation*

Subject	Age	Educational Background	Time spent in Hungary
FB	26	BA Degree	3 years
FM	27	BA Degree	8 years

4.2.2 Chinese words used for investigating Tones 声调字表

It was also essential to select an appropriate set of syllables for the subjects to pronounce. Zhu Xiaonong, (2010) recommends that in choosing Chinese words for the purposes of an investigation into tones, vowels such as *a, i* and *u* are to be preferred, while *b* and *d* as voiceless and unaspirated consonants are favored[149]. In addition, zero initials can also be chosen for experiments. The duration of tone is not equal to the time taken to pronounce the whole syllable bearing this tone or the time taken to pronounce the final (the whole syllable minus the initial), but to the time taken to pronounce the rhyme. In fact, it is the time taken to

[149] Zhu 2010: 279.

pronounce the final minus time taken to pronounce the medial or the time taken to pronounce the nucleus plus the time to pronounce the ending. Therefore, according to Zhu Xiaonong, it would be better for the syllables to be measured without the medial, and to avoid syllables starting with semi-vowels, but to choose syllables consisting of fully voiceless, plosive *b, d* with *i*, *u* or *a* as vowels[150]. Syllables based on the combinations mentioned above fall into nine categories: *ba*, *bi*, *bu*, *da*, *di*, *du*, *ya*, *yi* and *wu*. Then, these syllables are coupled with different tones. Frequently-used free morphemes are preferred. More than one tone can be chosen for some syllables with various free morphemes. However, for syllables with fewer or no morphemes, we can either choose a non-free morpheme or keep the corresponding position in the table vacant. Morphemes that each tone contains should be balanced in number. Zhu Xiaonong recommends that times for reading each kind of tone should be no less than 12ms.; times of 24 to 30 ms are preferable.

The set of syllables selected for the present investigation consists of a total of 60 monosyllabic words, shown in the following table. The numbers of first-tone “Yinping” (T1) and fourth-tone “Qusheng” (T4) are the same: 15 of each. The numbers of second-tone “Yangping” (T4) and third-tone “Shangsheng” (T3) are 14, 16, respectively.

Table 25: *Matrix of 60 Chinese words representing the 4 tones*

	a				i				u			
	T1	T2	T3	T4	T1	T2	T3	T4	T1	T2	T3	T4
b	巴 八	拔	把	爸 罢	逼	鼻	比 笔	必 币	逋		哺 捕 补	部 布
d	答 搭	达	打 沓	大	滴 低	迪 笛 敌	底 抵	第 弟	督 都	读 独 毒	堵 赌	度 杜
Zero Initials	阿				衣 一	姨 宜 咦	已 以	易 意	乌 屋	吴 无	五 午	误 物

[150] Zhu 2010: 281.

4.2.3 Procedures for obtaining recorded data

Words chosen for the test were printed on two sheets of A4 paper (see Appendix); each subject read out both sheets. Exactly the same words appeared on both sheets, but in different order. Before the recording, subjects were required to practise by reading the words on their paper several times to ensure that they could read these words fluently and naturally. For the recording itself, subjects were required to read each monosyllabic word in the word table only once. They were told to take a two-second pause before they read the next word, in order not only to avoid the possible sound changes caused by an excessively short interval between two syllables, but also to make the subsequent data extraction and analysis more manageable.

Recordings were conducted in the recording room of the Institute of Linguistics of the Hungarian Academy of Social Sciences, which, to a great extent, ensured the effectiveness of the recordings and the quality of the resulting data[151]. The recording devices include a laptop, dedicated microphone and PRAAT software (version 5.3.51, Boersma & Weenink, 2013)[152].

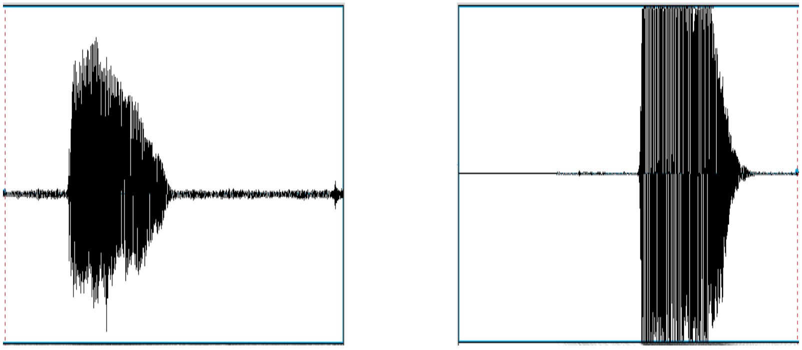
Since the sound volumes produced by the subjects varied, before recording, the volume of the laptop had to be adjusted to suit the volumes produced by different people, in order to avoid excessively loud or quiets sounds in the recording. In particular, excessively high volume-levels might have led to overloading, thus producing clipping. The expression "overloading" refers to what happens when the upper and lower ends of the amplitude go beyond the window. That is to say, the amplitude is greater than 1 or less than -1. As a result, to use Zhu Xiaonong's wording, the upper and lower ends are flattened[153], which can be shown by the graph below.

---

[151] The author is indebted to Andrea Deme, of the Institute of Linguistics of the Hungarian academy of Sciences, for her invaluable help and advice concerning the use of the technical data collection equipment.

[152] Boersma, P. & Weenink, D. 2013.

[153] Zhu 2010: 280.



Audiograph for normal sound volume Audiograph for overloaded sound volume

Figure 3. The effect of overloading on audiographs

In spite of this, due to the subjects' nervousness and other factors in the recording process, it is quite possible that some subjects' sound volumes will still be too loud or low. In this case, it is suggested that to ensure the proper sound volume, it is prefereable to adjust the volume of the laptop, rather than asking subjects to speak louder or lower, so as not to further affect subjects. PRAAT can monitor and indicate whether the sound volume is appropriate or not, as shown in the figure below. The green area refers to samples with suitable sound volumes; the yellow area implies a warning; the red area means "overloaded" samples with excessively sound volumes (as shown in the following figure).

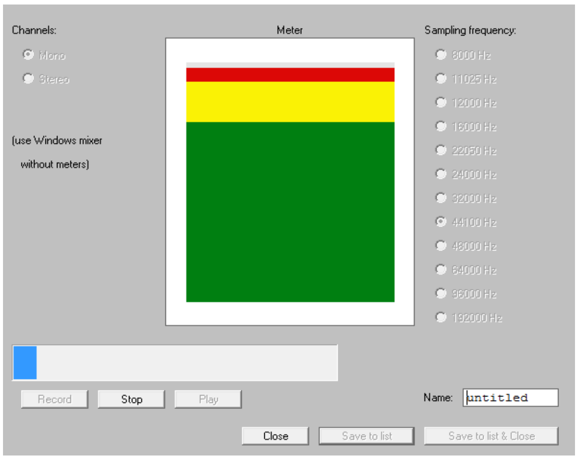


Figure 4. PRAAT volume-level indications

In the recording, the sampling frequency is 44100; audio files are saved into WAV format.

4.3 Data Extraction and Analysis

The heart of experimental phonetics lies in analyzing the voice samples after recording. In research on tones, the key issue is to measure and analyze the fundamental frequency, since the fundamental frequency is the most universal and indispensable distinctive feature for any tone system. The pitch range (both the absolute value and the distance between the highest to the lowest pitch produced) is another essential parameter for describing tones, while the duration is the third integral parameter (Zhu Xiaonong, 2010)[154].

In this investigation a total of 720 voice samples (multiplying 12 people by 60 syllables) are analyzed. Differences and similarities between 10 Hungarian learners of Chinese and 2 native speakers from Beijing are analyzed in terms of the duration, the fundamental frequency and tone range of the monosyllabic tones they produced, thus seeking to identify problems in tone acquisition experienced by this sample of Hungarian university students. In view of the central importance of this phase of the investigation, the technical procedures employed will be described here in some detail, largely on the basis of the published instructions for using PRAAT software (Boersma & Weenink, 2013)[155] but using examples related specifically to the Chinese tone patterns under investigation and some of the issues that arise in connection with them, as discussed by Zhu Xiaonong (2010)[156].

4.3.1 Determining the Duration of Tone

As already mentioned above, the duration of tone is not equal to the total time taken to pronounce the syllable bearing this tone or the time to pronounce the final (to pronounce the whole syllable minus the time to pronounce the initial), but the time to pronounce the rhyme (that is, the time taken to pronounce the final minus the time taken to pronounce the medial or the time taken to pronounce the nuclear vowel plus the time taken to pronounce the ending). Thus, it is essential to determine the start and the end of a tone.

A tone actually starts from the starting point of the nucleus (the vowel proper). In the

---

[154] Zhu 2010: 275

[155] Boersma, P. & Weenink, D. 2013.

[156] Zhu (ibid).

spectrogram, it starts from the second pulse of the vowel[157] .

There exist unified standards for determining the end of a tone. In addition, supporting standards are also formulated for different tone patterns. Unified standards include the following aspects. On the one hand, the amplitude of the spectrogram can be seen to undergo a sharp decline. On the other hand, one may observe whether the second formant in the wideband spectrogram is clear or not. If the second formant has been blurred, it can be considered as the end of this tone. (Zhu Xiaonong, 2010)[158]

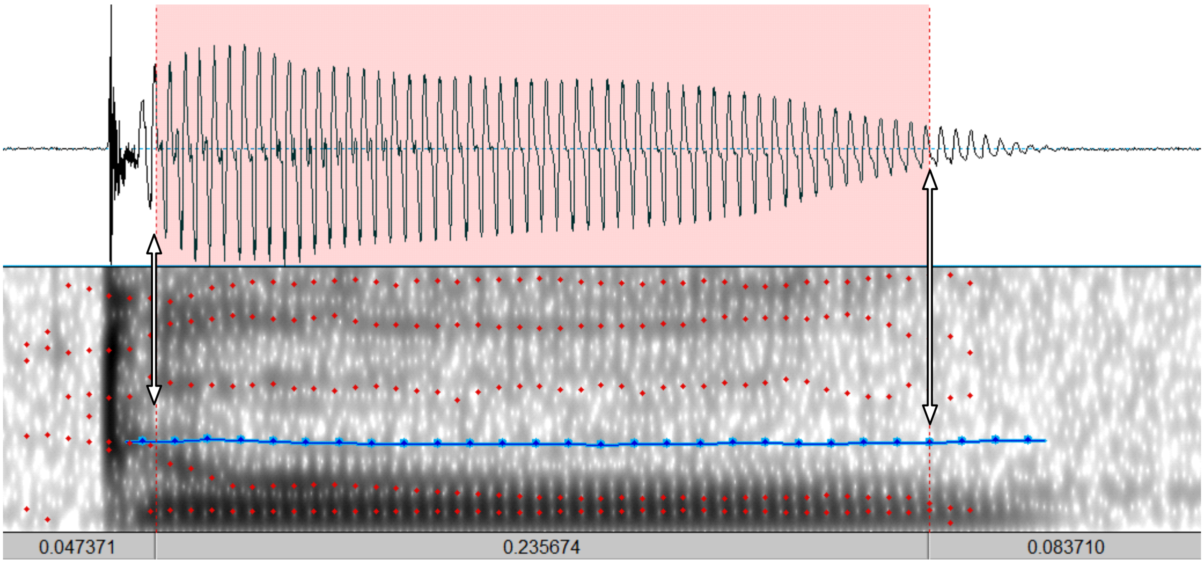


Figure 5. Spectrogram for the syllable "dū"; arrows indicate the start and ending of the tone.

Rising and falling tones have their own respective supporting standards. For rising tones, especially at the end of a high rising tone, there often exists a non–phonemic glottal stop. At this point, an inaudible falling in the end of the fundamental frequency will appear in the spectogram (Rose 1989)[159]. The end of a rising tone is set at the peak of the fundamental frequency in the narrowband spectrogram (Zhu Xiaonong, 2010)[160].

[157] Lisker Abramson 1963: 416; Baken 1987: 376.

[158] Zhu 2010: 281-282.

[159] Rose, Phil 1981.

[160] Zhu (ibid).

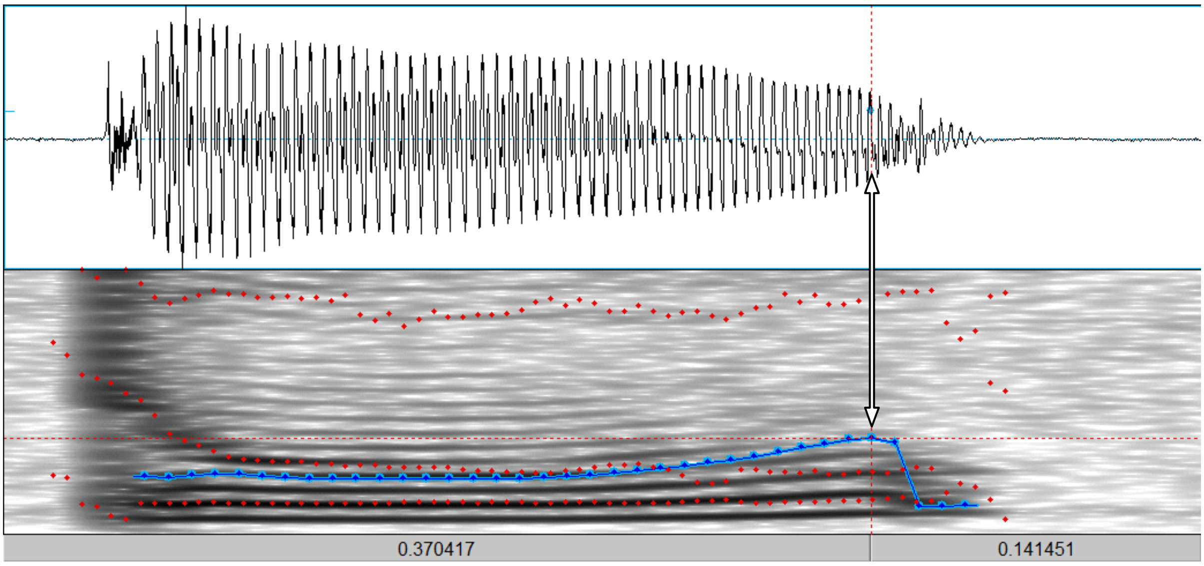


Figure 6. Pectogram for the syllable “dú”; arrow indicates the end of the rising tone.

The end of a falling tone corresponds to the point where fundamental frequency ceases to be marked by regular and proportional intervals in the wideband spectrogram (Zhu Xiaonong, 2010)[161].

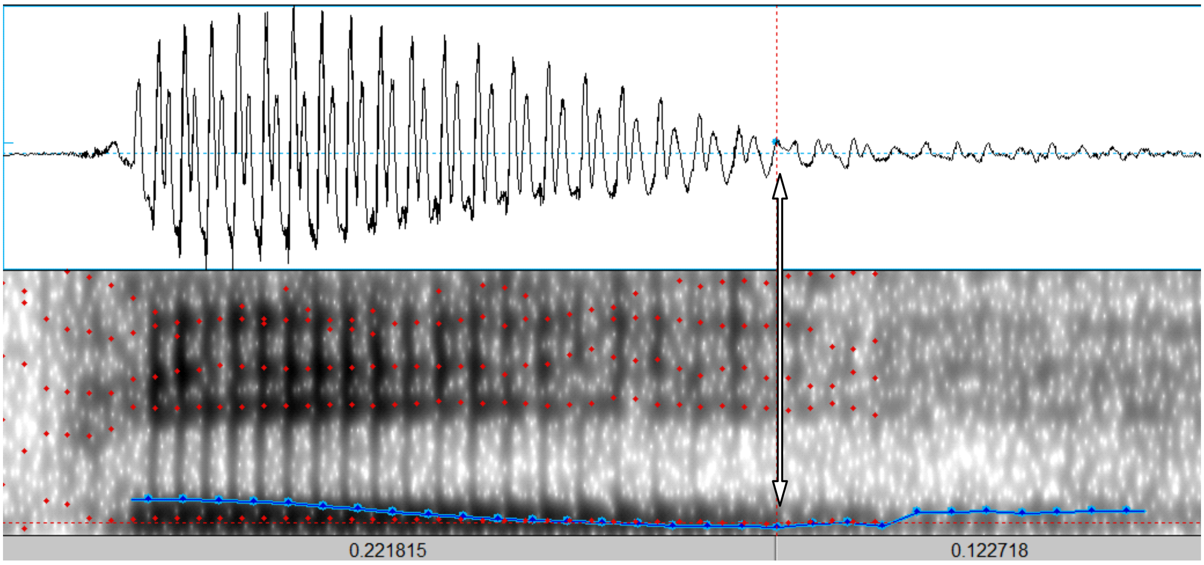


Figure 7. Spectrogram for the syllable “bì”; arrow indicates the end of the falling tone.

[161] Zhu 2010: 281-282.

4.3.2 Fundamental Frequency Measurement and LZ Normalization

4.3.2.1 Fundamental Frequency Measurement

In this investigation PRAAT was used mainly to automatically measure the pitch. After the duration of a tone has been determined, it is divided into 10 equal parts. This gives 11 measuring points at equal intervals, and data about the fundamental frequency is obtained for each of these measuring points. Taking "bá" for example, the specific steps followed for measuring the fundamental frequency (as prescribed by the PRAAT instructions[162]) were as follows.

1. First, the duration of the tone was determined as described above, then the relevant period, from the start and end of this tone, was selected by clicking and dragging the cursor.

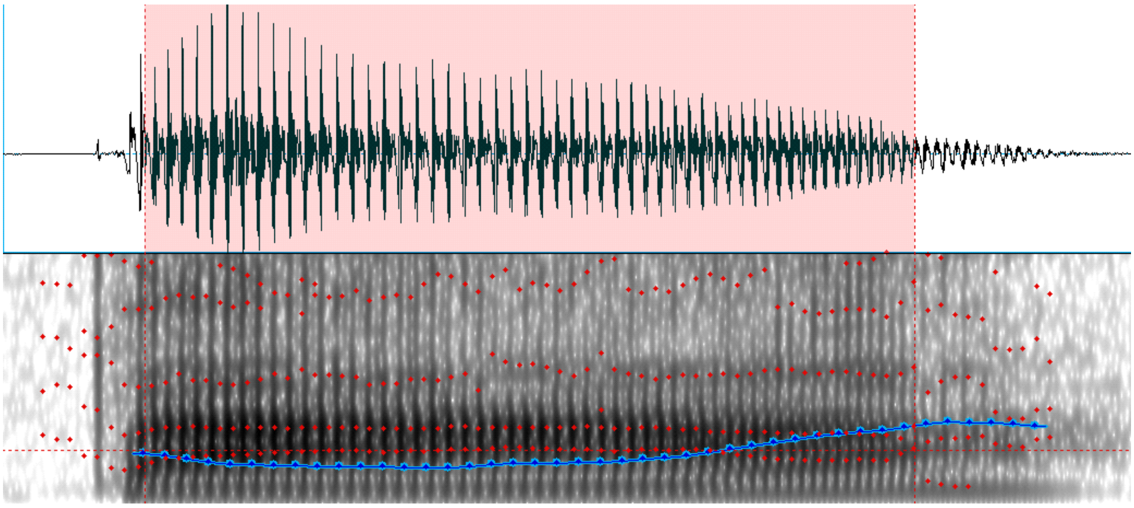


Figure 8. The whole syllable, with the actual tone highlighted

2. Then, by clicking on "sel" in the bottom left corner, the selected part was enlarged. In this case, the selected part lasts for 0.353 seconds; that is, the tone duration is 0.353 seconds.

[162] Boersma, P. & Weenink, D. 2013.

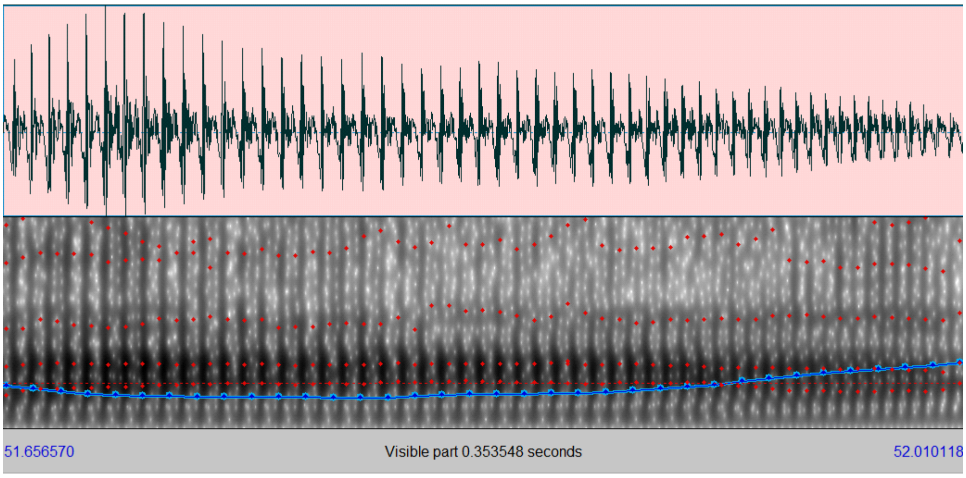


Figure 9. The duration of the tone, selected for analysis.

3. Thirdly, clicking on "View" on the top left corner and selecting the option "Time Step Setting", elicited a pop-up window as shown in the following figure. After selecting "fixed" in the "Time Step Strategy", 0.0345 (approximately one tenth of 0.353 as the duration of this tone) was typed into "Fixed time step(s)". To ensure that 11 testing points are covered, the input value was often slightly less than one tenth of the tone's duration.

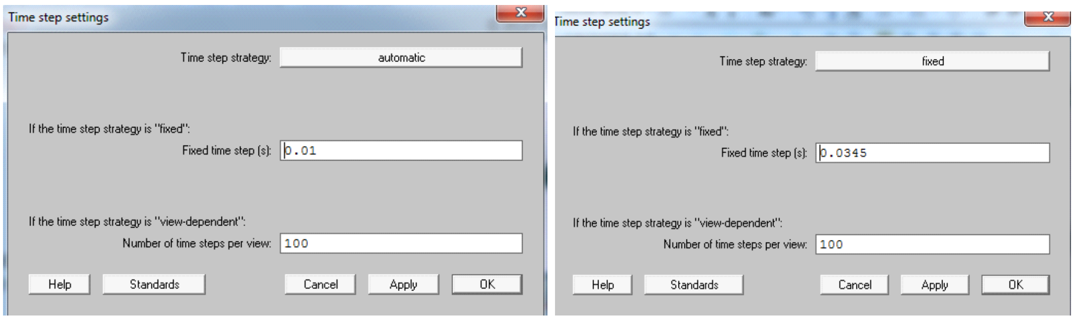


Figure 10. Dialogue box for inputting time-step values

4. Clicking on "Ok' fixed the division of the fundamental frequency curve into 10 equal parts, marked by 11 measuring points.

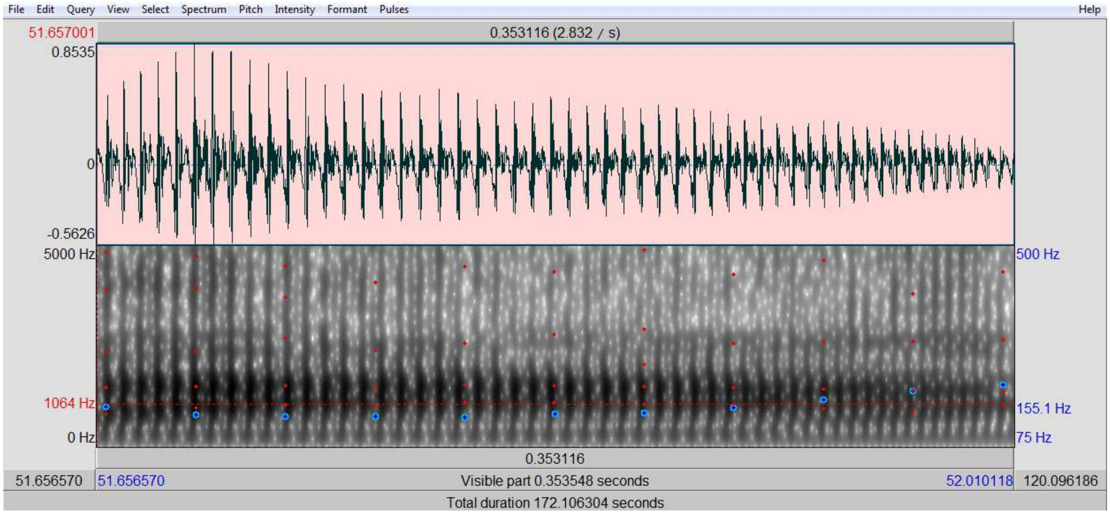
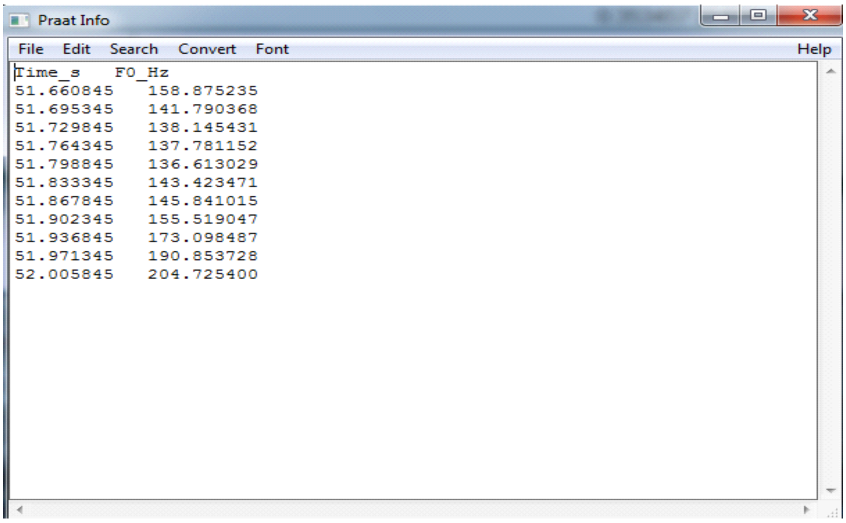


Figure 11. Measuring points, marked by blue dots

5. By clicking on "Pitch" in the top left corner and selecting the option "Pitch Listing", data on the fundamental frequency at the 11 measuring points were obtained.



Praat Info

File Edit Search Convert Font Help

Time_s	F0_Hz
51.660845	158.875235
51.695345	141.790368
51.729845	138.145431
51.764345	137.781152
51.798845	136.613029
51.833345	143.423471
51.867845	145.841015
51.902345	155.519047
51.936845	173.098487
51.971345	190.853728
52.005845	204.725400

Figure 12. Tabulated F0 data for each measuring point.

6. Finally, data obtained through the steps detailed above were exported to Microsoft Word and Excel files for subsequent analysis.

However, in some cases, PRAAT cannot automatically obtain data about the fundamental frequency. To deal with this, Zhu Xiaonong (2010) puts forward two supplementary methods

[163] to manually measure the fundamental frequency. One is the traditional approach, implemented by measuring the harmonic. Taking this recording of “dà” as an example, we note that the fundamental frequency curve is disrupted between the seventh and ninth measuring points, as shown below.

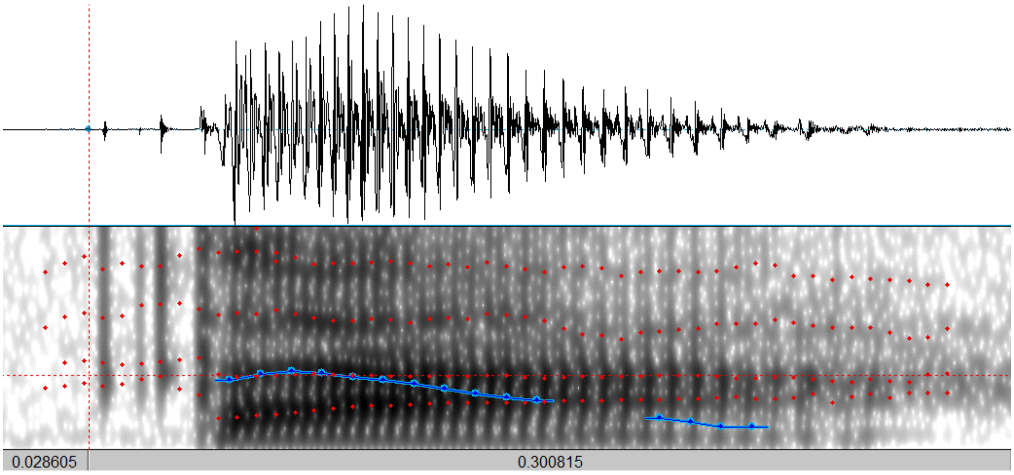
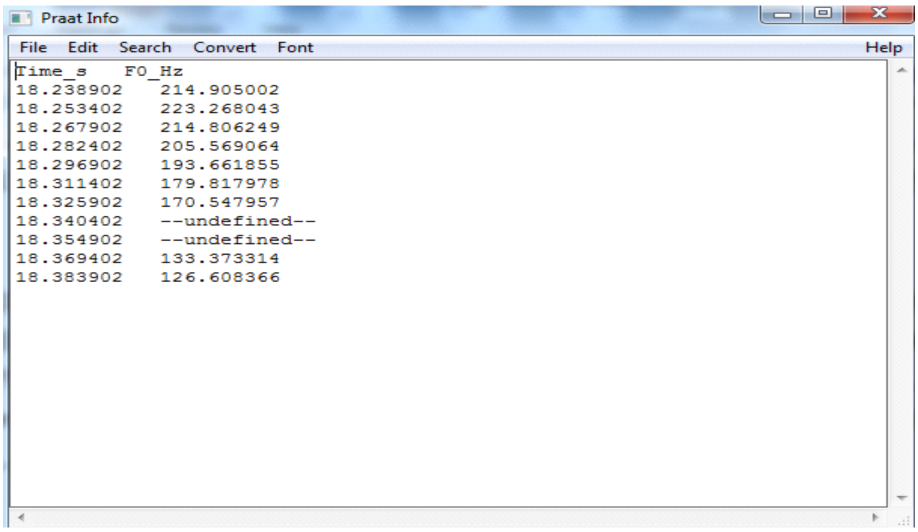


Figure 13. Fundamental Frequency Curve for the Syllable “dà”

In accordance with the above measurement steps, data on the fundamental frequencies at the 11 test points is obtained as shown in the following figure.



Time_s	F0_Hz
18.238902	214.905002
18.253402	223.268043
18.267902	214.806249
18.282402	205.569064
18.296902	193.661855
18.311402	179.817978
18.325902	170.547957
18.340402	--undefined--
18.354902	--undefined--
18.369402	133.373314
18.383902	126.608366

Figure 14. Data on the Frequency Curve for the Syllable “dà”, with disruption at the 8th and 9th measurement points.

Due to the disruption of the fundamental frequency curve, the fundamental frequencies at the

[163] Zhu 2010: 285-286.

8th and 9th measurement points are given as “undefined”. In order to obtain data about the fundamental frequencies at these two points, the following steps may be taken.

(1) First of all, the wideband spectrogram is changed into a narrowband one. The wideband spectrogram is a design sketch for PRAAT. In order to do this, on the basis of the 4th step in the “fundamental frequency measurement”, we select “Spectrum” in the menu and click on “Spectrum Settings” to change the “ View Range” in the pre-set wideband spectrogram from “ 0-5000’ to “ 0-3000” and the “ Window Length” from “ 0.005” to “ 0.04”.

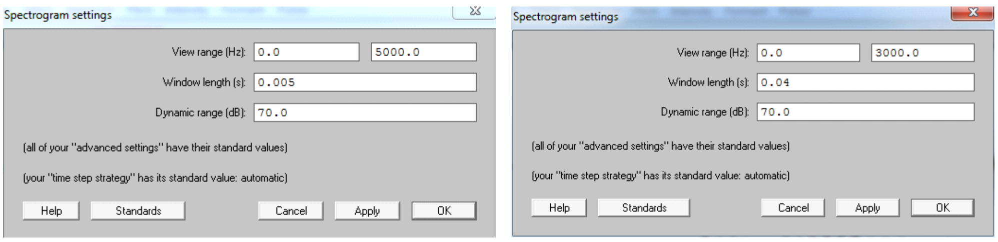


Figure 15. Dialogue boxes for changing from wide- to narrowband spectrogram

(2) Clicking on “OK” effectively changes the wideband spectrogram for the syllable “dà” into the narrowband one as shown below.

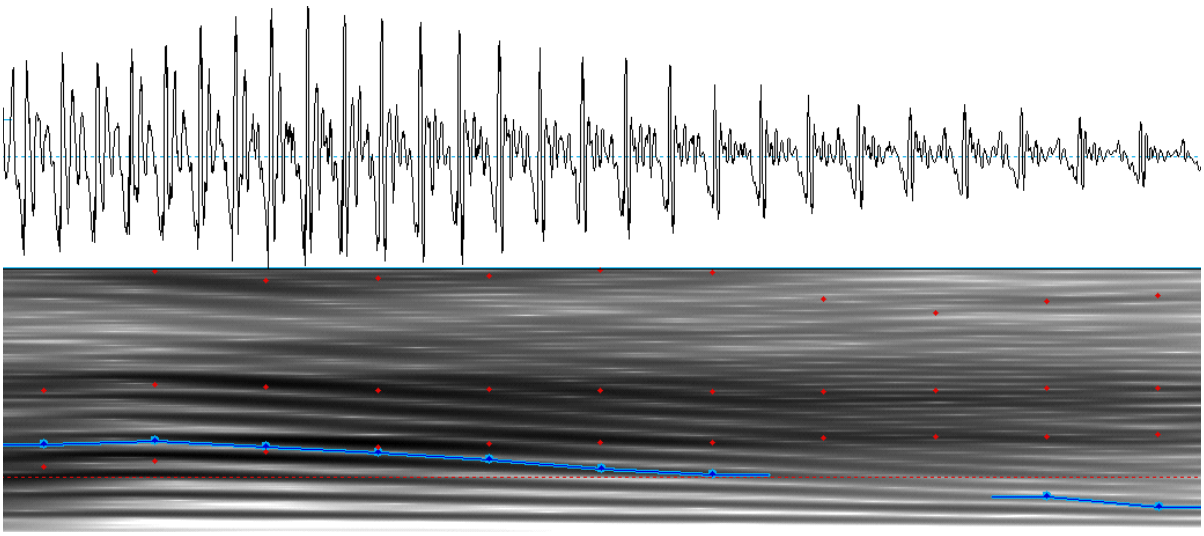


Figure 16. Narrowband version of the Fundamental Frequency Curve for the Syllable “dà”

(3) The cursor is now placed at the 8th measurement point. Two cursor dotted lines in the spectrogram intersect in the center of the 4th harmonic H4. From the longitudinal axis,

it can be seen that the fundamental frequency of H4 is up to 638.2 Hz. The frequency of a harmonic is an integer multiple of the fundamental frequency (Zhu Xiaonong, 2010, 239). The frequency of H4 is divided by 4 to get 159.55, which is the fundamental frequency at the 8th measurement point. Using the same method, the fundamental frequency at the 9th measurement point can also be obtained.

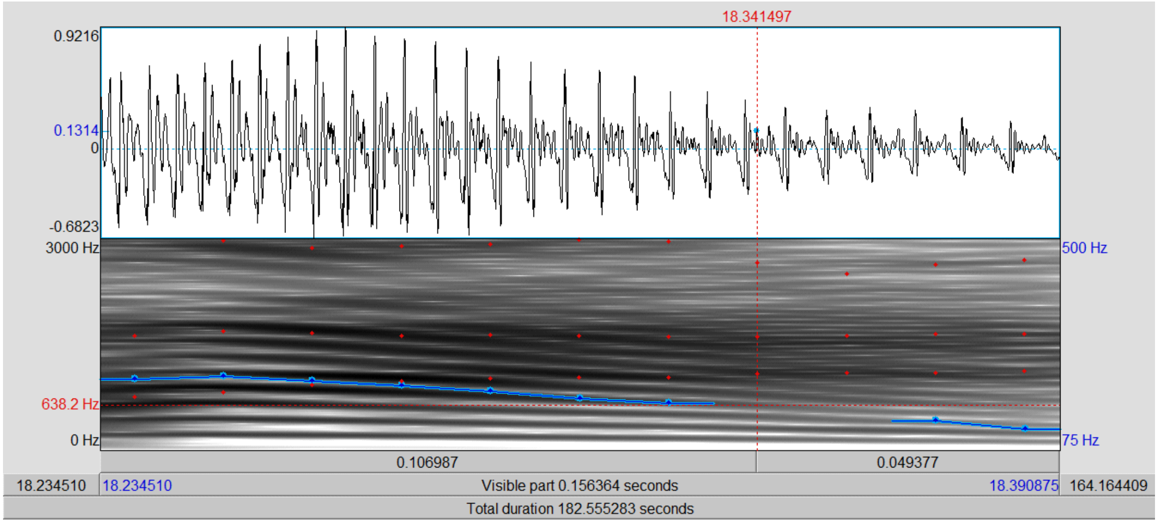


Figure 17. Stage 3 of the procedure for reconstructing disrupted data.

The other method is implemented by measuring the periodic length of the audiogram. Taking "dì" for example, the last part of the fundamental frequency curve for the syllable "dì" is missing. According to the above-mentioned principle for determining the end of a tone, the end of the tone of "dì" should be in the end of the fundamental frequency with regular and proportional intervals as shown below.

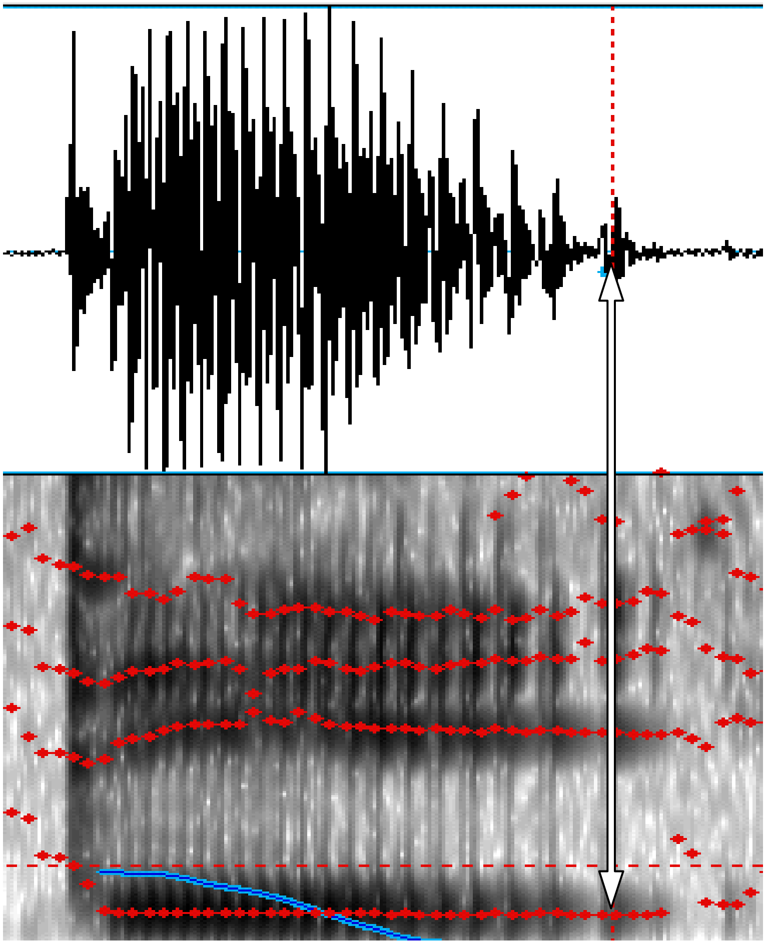


Figure 18. Audiogram of "dì"; arrow indicates the end of the tone

In order to obtain data on the fundamental frequency of the end of tone of the "dì", the last part of the syllable may be selected (see following figure) and enlarged by clicking on "sel". After that, two peaks of the last period in the audiogram are selected. From the horizontal axis, it can be seen that the duration of the selected part is 0.025209s. The fundamental frequency is the reciprocal of the duration. Therefore, the fundamental frequency for the end of this tone can be calculated at approximately 40 Hz.

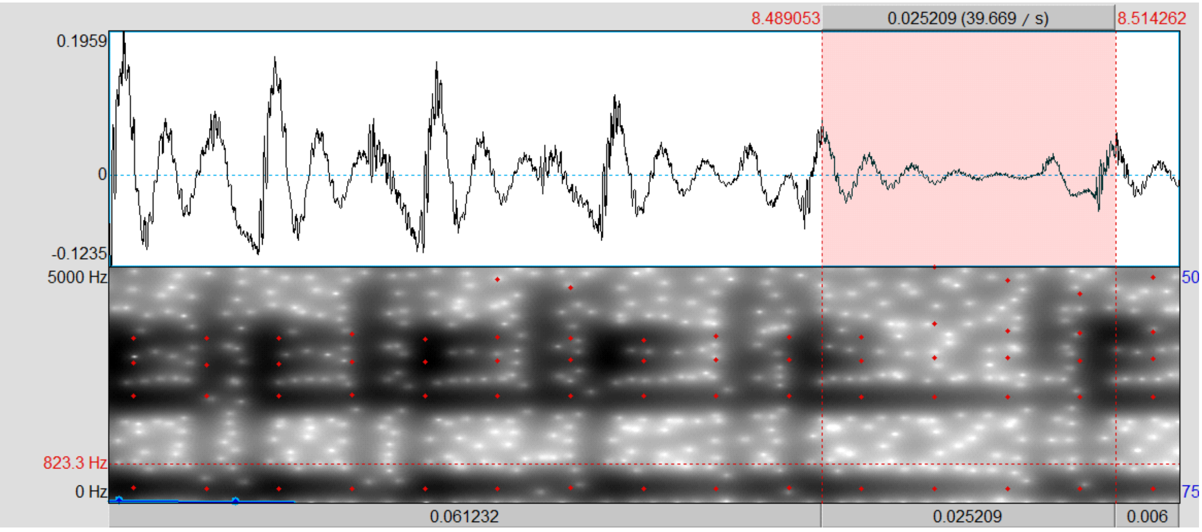


Figure 19. Disrupted section of audiogram selected for repair

Another potentially problematic phenomenon that should be mentioned here is that of "Creaky voice". As Zhu Xiaonong (2010) pints out, due to the thick and stiff nature of an individual's vocal folds, the fundamental frequency of creaky voice is quite low and far below the minimum of the voice sphere of subjects[164]. Since the fundamental frequency of the creaky voice is low and quite irregular, there may be no way to measure it (. Creaky voice usually occurs during the production of vowels, and frequently appears in the "*Shangsheng*" of the Beijing dialect. In the spectrogram, creaky voice is often manifested in form of disruption in the fundamental frequency curve as shown below.

[164] Zhu 2010: 93.

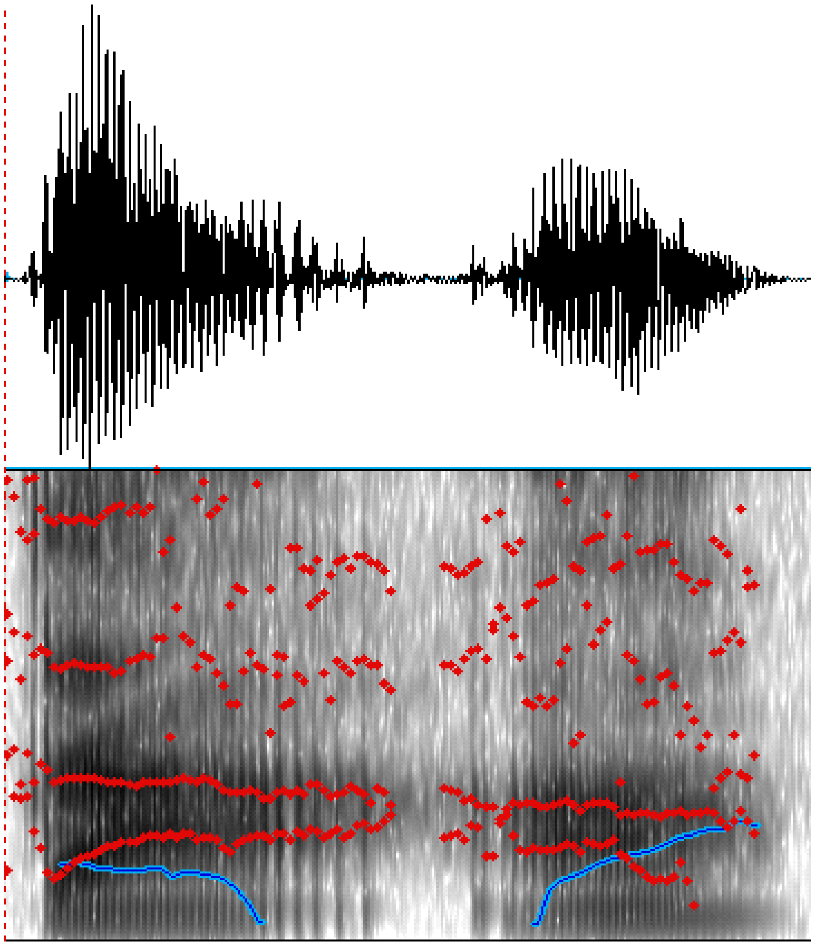


Figure 20. “da” pronounced by FM, disrupted by creaky voice.

The diagram above is a spectrogram for the syllable “dǎ” pronounced by one of the native-speaker subjects in the current investigation: FM. Due to the appearance of creaky voice, the fundamental frequency curve is disrupted. In this case, the author chose to discard this voice sample and use the corresponding voice sample in the other version to replace it (as mentioned above, each subject was asked to make two recordings based on two sheets of paper with the same group of words in a different order). If the voice sample for “dǎ” provided in the other version was also affected by creaky voice, another voice sample with “Shangsheng” would be chosen to replace it. If creaky voice was pervasive in the voice samples of a subject, then all the samples provided by this subject were discarded in order to ensure the reliability and validity of data obtained from experiments.

4.3.2.2. LZ Normalization

Normalization, according to Rose (1993) serves to eliminate the random interpersonal differences so as to extract constant parameters, that is, eliminating personal characteristics and obtaining linguistically meaningful information[165]. Zhu Xiaonong (2010) states that the primary role that normalization plays is to describe the tone based on standardized quantitative descriptions; the other role is to reduce the differences between pronunciation styles such as the formal, the casual and the nervous ones. Normalization makes it possible to find the constant in the interpersonal[166] differences and find the commonalities in the interlanguage variations, thus making research on interpersonal and interlingual comparisons possible. Normalization procedures involve the following two basic stages. In the first stage, parallel movement is made on the coordinate; in the second, the frequency domain is set through compression or expansion. The key to normalization lies in determining the frequency domain. In this invetigation, the logarithmic z-score (LZ) method created by Zhu Xiaonong (2010)[167] was adopted. LZ normalization follows the following steps: (1) the average value of the fundamental frequency is calculated; (2) average value is converted into the logarithm; (3) the average value and standard deviation of the logarithm are calculated; (4) Z-score normalization is performed; (5) the normalized result of each subject is averaged and the standard deviation of the average value is calculated.

Data for the four tones of Mandarin provided by the native-speaker subject FB will be used to exemplify the process employed to perform LZ normalization. The whole process is carried out using Microsoft Excel.

(1) First of all, we calculate the average value of the fundamental frequency (using Hz as its unit) of each voice sample provided by the articulator MB at each measurement point, as shown in the following table.

---

[165] Rose, Phil 1981.

[166] Zhu 2010: 286.

[167] Zhu (ibid).

Table 26: *MB's average F0 values at each measurement point for all tones.*

	0%	10%	20%	30%	40%	50%	60%	70%	80%	90%	100%	ms
T1	131.2667	131.1333	131	130.9333	131.2667	130.6	129.9333	129.9333	129.4	129	128.0667	479.2
T2	105.6429	103	102.7857	104.1429	107.4286	112.2143	119.7857	127.4286	138	147.1429	156	520.8571
T3	100.5625	93.3125	88.375	87.625	87.3125	88	89.0625	91.625	96.5625	102.125	105.875	607.75
T4	153.8	153.4667	151.0667	145.0667	136.0667	126.3333	115.6667	105.8667	98.06667	92.73333	87.13333	281.6667

(note: ms refers to the duration)

(2) All the F0 data in the above table are then converted into logarithms. For example, the fundamental value of T1 at its start ("0%") is 131Hz. Coverting this value into its logarithm, we obtaining the value 2.12 (lg (131) ≈ 2.12). Conversion results are as follows:

Table 27: *MB's average F0 values converted into logarithms*

	0%	10%	20%	30%	40%	50%	60%	70%	80%	90%	100%	ms
T1	2.118154	2.117713	2.117271	2.11705	2.118154	2.115943	2.113721	2.113721	2.111934	2.11059	2.107436	479.2
T2	2.02384	2.012837	2.011933	2.017629	2.03112	2.050048	2.078405	2.105267	2.139879	2.167739	2.193125	520.8571
T3	2.002436	1.96994	1.946329	1.942628	1.941076	1.944483	1.949695	1.962014	1.984809	2.009132	2.024793	607.75
T4	2.186956	2.186014	2.179169	2.161568	2.133752	2.101518	2.063208	2.024759	1.991521	1.967236	1.940184	281.6667

The average value μ and standard deviation σ of the logarithms is now calculated. In calculating these two values, the values of the tone at the 1st measurement point are ignored, as are values for the falling tone (T4) at the 11th measurement point (Zhu Xiaonong, 2010)[168]. Therefore, data shaded in blue in the previous table are discarded. According to the table, it can be seen that μ is up to about 2.06244, while σ is about 0.076867. The formula for calculating the standard deviation is as follows;

$$\sigma = \sqrt{\frac{1}{N}\sum_{i=1}^{N}(x_i - \mu)^2}$$

(3) Z-score normalization is then performed, using the formula (xi-μ) / σ. Results of the normalization of the data in the previous table are shown below:

[168] Zhu 2010: 287.

Table 28: *Z-score normalised version of MB's data*

	0%	10%	20%	30%	40%	50%	60%	70%	80%	90%	100%	ms
T1	0.714019	0.714019	0.714019	0.714019	0.714019	0.714019	0.67071	0.67071	0.627066	0.627066	0.583082	479
T2	-0.4828	-0.64507	-0.64507	-0.59046	-0.42974	-0.17162	0.218321	0.538754	1.008233	1.36531	1.701162	521
T3	-0.75589	-1.22229	-1.53462	-1.53462	-1.59922	-1.53462	-1.47076	-1.28339	-0.98428	-0.70021	-0.4828	608
T4	1.628234	1.628234	1.517046	1.287886	0.925723	0.494075	0.026715	-0.4828	-0.92631	-1.22229	-1.59922	282

4.3.3 Determining Tone Range

Owing to differences of gender, vocal folds, timbre and other objective conditions of different individuals, their sound volumes, that is, the fundamental frequency of their tones, are not the same. Since the pitch of the tone refers to the relative pitch, the research object should be the tone range, that is, the range within which the tone fluctuates. For "*Yinping*", the the issue is how flat or straight and steady the contour is; the degree to which the contour rises is the focus for "*Yangping*"; the degree of falling-and-rising is the focus for "*Shangsheng*", studying both the degree of the decline in the first half part and the degree of the rise in the latter half; the degree of the decline is the focus for "*Qusheng*". As an example, we will consider the contour of the average values of the fundamental frequency of "*Yinping*" syllables produced by the native-speaker subject MB.

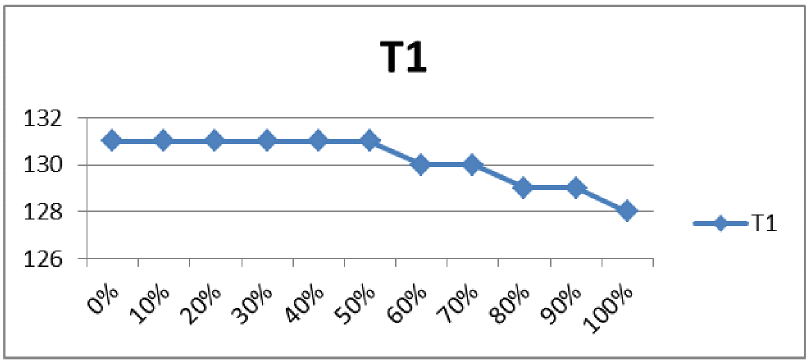


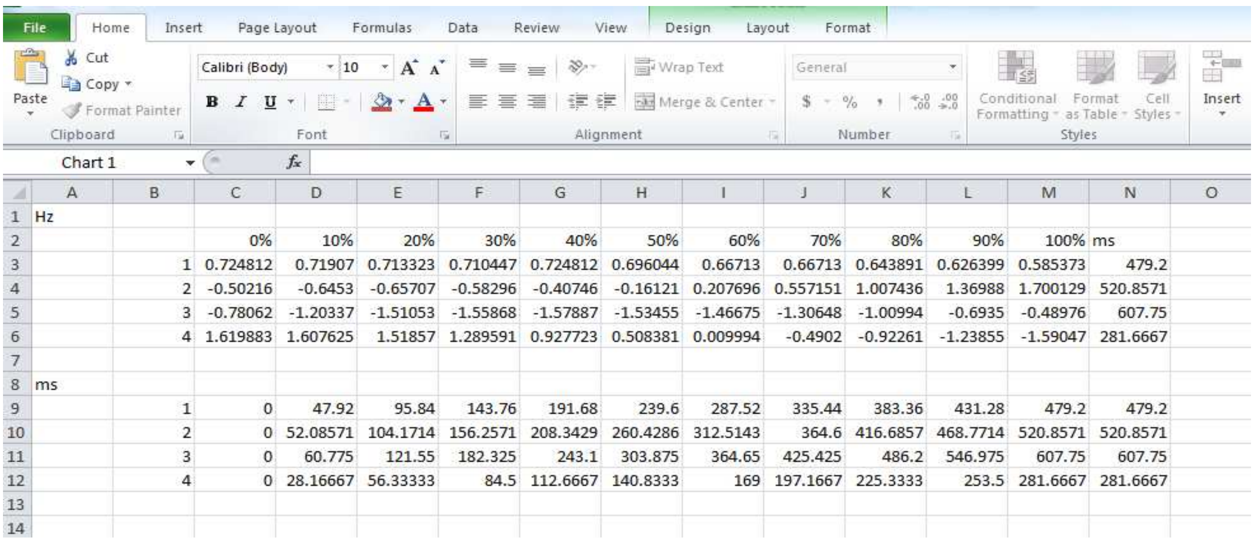
Figure 21. Mean of Yinping (T1) contours produced by MB

The highest value for the fundamental frequency of the high-level tones by the articulator MB is about 131Hz, while the lowest one is about 128 Hz. The tone domain is therefore approximately 3 Hz (131-128 = 3). That is to say, the tone range is 3 Hz. For "*Yinping*", the smaller the tone range is, the higher the degree of straightness of this tone is. Thus it can be seen that the "*Yinping*" contour produced by MB is quite straight.

4.3.4 Tone Mapping

Tone mapping is the process by which each subject's recorded data for all four tones are summarized and plotted together on a single graph. The data conversion and mapping involved in this study were implemented using Microsoft Excel. Tone mapping was performed in the same way. As explained above, LZ normalization makes interpersonal and interlingual comparisons possible. Therefore, in this section, the fundamental frequency curve of the absolute duration of data, obtained by performing normalization, is introduced. The method for drawing the fundamental frequency curve for the absolute duration will be demonstrated by taking data from MB used in the previous section.

(1) Data about the duration of tone were tabulated, on the basis of the Excel table about normalization; that is to say, input the durations of tone in 11 equal division points into corresponding tables. For example, if the duration of T1 is 479.2ms, the duration in the 1st equal division point is 0 ms; 47.92 ms in the 2nd equal division points. The rest can be deduced in the same manner, with the result shown in the following diagram.



	A	B	C	D	E	F	G	H	I	J	K	L	M	N	O
1	Hz														
2			0%	10%	20%	30%	40%	50%	60%	70%	80%	90%	100%	ms	
3		1	0.724812	0.71907	0.713323	0.710447	0.724812	0.696044	0.66713	0.66713	0.643891	0.626399	0.585373	479.2	
4		2	-0.50216	-0.6453	-0.65707	-0.58296	-0.40746	-0.16121	0.207696	0.557151	1.007436	1.36988	1.700129	520.8571	
5		3	-0.78062	-1.20337	-1.51053	-1.55868	-1.57887	-1.53455	-1.46675	-1.30648	-1.00994	-0.6935	-0.48976	607.75	
6		4	1.619883	1.607625	1.51857	1.289591	0.927723	0.508381	0.009994	-0.4902	-0.92261	-1.23855	-1.59047	281.6667	
7															
8	ms														
9		1	0	47.92	95.84	143.76	191.68	239.6	287.52	335.44	383.36	431.28	479.2	479.2	
10		2	0	52.08571	104.1714	156.2571	208.3429	260.4286	312.5143	364.6	416.6857	468.7714	520.8571	520.8571	
11		3	0	60.775	121.55	182.325	243.1	303.875	364.65	425.425	486.2	546.975	607.75	607.75	
12		4	0	28.16667	56.33333	84.5	112.6667	140.8333	169	197.1667	225.3333	253.5	281.6667	281.6667	
13															
14															

Figure 22. Excel chart of BM's LZ-normalised valuses for all four tones

(2) The fundamental frequency curve for the absolute duration can be obtained by taking Hz as the vertical axis and ms as the horizontal axis, which is shown below.

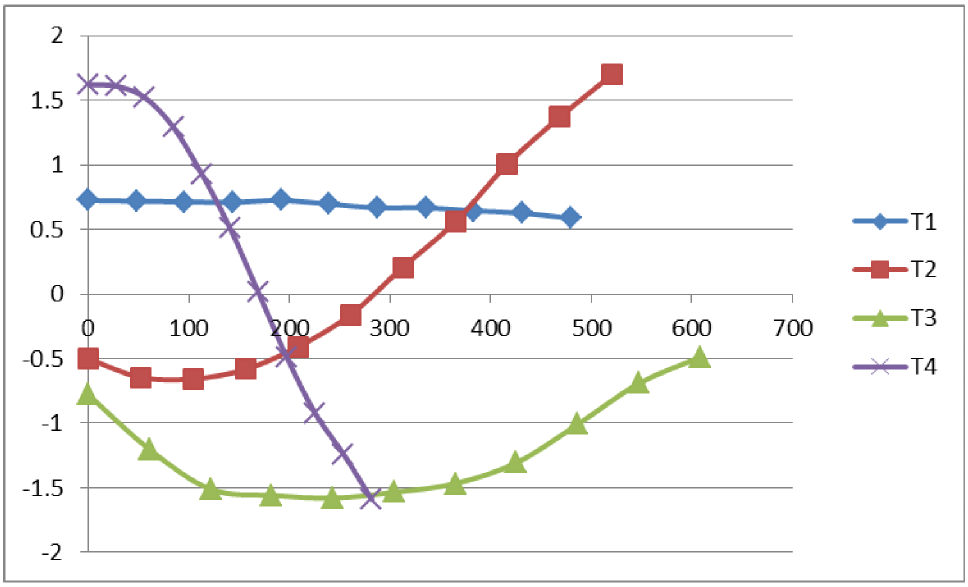


Figure 23. Tone map for BM

4.4. Judgements

Recording, processing and analyzing the data from the Hungarian and Chinese subjects was relatively easy, though it was time-consuming and the use of PRAAT and Excel required some expertise. Greater problems arose when it came to judging the Hungarian subjects' pronunciation in comparison with the Chinese subjects'. The small size of the sample and the absence of generally accepted quantitative measures of correctness in this area meant that purely statistical judgements could not be reliably applied. If the experiment were to be repeated at a later date with a much larger sample, one of the first tasks would be to establish quantitative criteria, probably by using the approach commonly applied to create criteria for criterion-referenced productive language tests. In this case a suitable panel of native-speaker listeners would judge the recorded data using key terms such as "incomprehensible" or "acceptable", and the sum of their subjective judgements, duly processed, would be matched with PRAAT data to constitute a set of standard quantitative criteria. Whether or not these could be considered truly objective is open to debate, but they could certainly be used as if they were. In the present research there were cases in which Hungraian subjects produced tone contours, for example, which deviated so widly from any possible native speaker norm (see, for example, M4's Yangping data) that the problematic nature of the subject's pronunciation could not be ignored. In many cases, however, judgements were based on the author's subjective opinion about whether a particular subject's pronunciation was within acceptable limits, noticeably deviant but not confusing, or actually likely to cause communicative problems.

5. Analysis of Data

Before we start to analyze data, there are a few related concepts that need to be clarified. The concepts discussed here are based on Zhu Xiaonong's seminal publication *Yun Yinxue* (2010)[169]. The first of these is the tone composition. A tone is made up of three components: start, body and end. The start is the pitch section which constitutes about 10%-20% of the whole length of the tone in question. The end is the pitch section which makes up the final 10%-20% of the tone in question. Apart from the start and the end, the rest is the body.

Secondly, we must mention the linguistic goal of tone. The part of the tone which is important in terms of the recognition of a tone by the listener is what Zhu Xiaonong calls the *shengdiao de yuyanxue mubiao* 声调的语言学目标 or "linguistic goal of tone" [my translation], which is conveniently abbreviated by the same author as *shengdiao mubiao* 声调目标 or "tone goal" [my translation]. The part which is unimportant to recognizing a tone is the *shengdiao de yuxian bufen* 声调的余羡部分 or "redundant part of the tone" [my translation]. Generally speaking, the tone goal is the *fengdian* 峰点 or "peak point" of a tone, no matter whether it is a level tone, an oblique tone, a high tone or a low tone.

5.1. *Yinping*

The first tone, also called *Yīnpíng*, is at a continuous high level. The value of tone in the Chao system is 55, which means it starts at point 5 and ends at the same level 5. *Xiandai Hanyu* defines *Yinping* as a high even tone with the highest start and a high and even voice. (Feng Zhichun, 2008)[170]

Zhu Xiaonong notes that theoretically speaking, the whole level tone itself constitutes a linguistic goal on the basis of the tone shape. The reason is that it will become a rising tone if the start is pronounced at a low pitch. It will become a falling tone if the end is pronounced at a low pitch. It will become an *aodiao* 凹调 or "dipping" tone (fall-rise tone) if the body is

169 Zhu 2010.

170 Feng 2008: 99.

pronounced at a low pitch. On the contrary, it will become a *jiangdiao* 降调 or falling tone, a *tudiao* 凸调 or raised tone (rise-fall tone) or a *shengdiao* 升调 or rising tone if the start, body or end is pronounced at a high pitch. Therefore, as Zhu Xiaonong (ibid, 277) points out, the level tone should keep a steady pitch continuation. In this section, the linguistic goals of *Yinping* which are the start, the body and the end will be investigated.

The main aim of this section is to present the voice samples of the first tone, *Yinping*, which were recorded by 12 subjects, including the measurement, extraction, analysis and comparison. The recordings are examined in terms of the duration, the fundamental frequency (F0), the tone range and the tone value of the monosyllabic tone produced by the recorded subjects, thus creating a detailed profile of the *Yinping* pronunciation of this sample of Hungarian university students (six male, labeled M1, M2 etc., four female, labeled F1, F2 etc.), juxtaposed with the pronunciation of one male and one female native speaker of Chinese (labeled MB and FB respectively). The actual methodology used in the measurement of duration, F0 and tone range was detailed in 4.3 above; note that the initial or “onset” of each syllable was discarded; the 11 measuring points all fall within the nucleus or rime. Thus the word “start” used in the comments about the subjects’ performance refers to the beginning of the nucleus, and not to the beginning of the whole syllable.

The 15 *Yinping* syllables used in the recordings, as explained in 4.3 above, fit into the following matrix:

Table 29: *Matrix of Yinping syllables*

	ā	ī	ū
b	巴、八	逼	逋
d	答、搭	滴、低	督、都
Zero-initials	阿	衣、一	乌、屋

5.1.1. Duration

The method used for determining the duration of the syllable has been presented before. The following table shows the durations of each *Yinping* syllable pronounced by 12 subjects, as well as average durations. The unit of duration is the millisecond (ms).

Table 30: *Duration of 15 different Yinping syllables pronounced by 12 different subjects.*

Syllables	F1	F2	F3	F4	M1	M2	M3	M4	M5	M6	FB	MB
都 dū	402	450	314	511	300	432	265	315	531	304	344	433
乌 wū	431	506	354	512	294	409	312	351	559	302	379	489
督 dū	335	560	322	638	258	320	248	360	720	305	370	467
阿 ā	424	361	290	534	229	365	270	306	649	293	297	457
答 dā	353	461	362	497	301	341	255	353	642	317	443	439
滴 dī	379	482	346	520	267	331	249	337	570	314	411	487
衣 yī	453	631	351	541	331	385	295	577	540	401	526	618
低 dī	407	591	463	562	302	351	269	371	581	379	496	464
屋 wū	444	472	344	457	259	340	285	353	600	373	535	490
逼 bī	360	438	442	482	263	291	284	280	540	333	374	432
一 yī	480	502	425	511	282	351	319	468	579	411	426	588
巴 bā	491	415	505	430	283	339	288	293	532	376	410	524
搭 dā	398	528	429	419	321	355	270	350	606	364	439	397
八 bā	432	470	391	550	280	299	302	271	714	370	403	471
逋 bū	437	505	385	557	276	307	301	300	591	330	466	432
Average	415	492	382	515	283	348	281	352	597	345	421	479

On the basis of the averages displayed in the bottom row of the above table, we made a chart for further comparison.

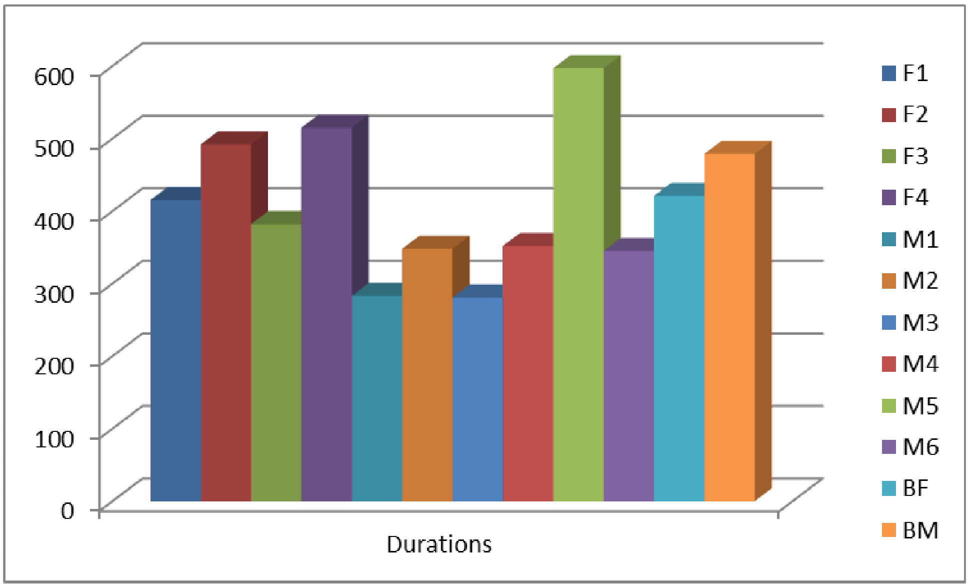


Figure 24. Average duration of *Yinping* syllables pronounced by 12 different subjects

The chart clearly shows that the average duration of M5 is the longest, while the average duration of M3 is the shortest, at less than half of M5's. The average durations of FB and MB are 421ms, 479ms, respectively. The average durations of F1, F2, F3, and F4 are close to FB's and MB's, the average durations of M1, M2, M3, M4, and M6 are considerably shorter than FB's and MB's, but the average duration of M5 is much longer than FB's and MB's. In other words, 60% of Hungarian university learners – all the male subjects in the sample - differ considerably from the "model" native speakers in terms of the duration of *Yinping* syllables, most often in producing much shorter durations.

Undoubtedly, producing a tone in too lengthy or too short a manner may be the cause of problems in duration, but it is probably not the only reason. The inappropriate pronounciation of the vowel which is bearing a tone also can lead to duration problems. Generally speaking, according to Zhu Xiaonong, the duration of a vowel is associated with its height, and each vowel has an intrinsic vowel duration (referred to as IVD).[171]. Catford (1977) believes that the duration of a high vowel is longer than that of a low one. His statement is described as IVD "i>a"[172].

The following table shows the average durations (m) of *Yinping* produced by 12 subjects, in terms of different vowels /i/, /u/ and /a/ which are bearing the *Yinping* tone, as well as the average durations of all female subjects (F) and male subjects (M). The unit of duration is the microsecond (ms), and "n" refers to the number of sample syllables. The last column in the table shows the subject's "duration order" of /i/, /u/ and /a/, with the longest vowel first and the shortest vowel last. For example, F1's average durations for the *Yinping* tone which is contained by vowel /i/, /u/ and /a/ are 416ms, 410ms and 420ms, respectively. The "duration order" of /i/, /u/ and /a/ is /a/>/i/>/u/.

[171] Zhu 2010: 45.
[172] Catford 1977: 45.

Table 31: *Average durations (m) of vowels /i/, /u/ and /a/ with Yinping tone produced by 12 subjects*

	/i/		/u/		/a/		the order
	m	n	m	n	m	n	
F1	416	5	410	5	420	5	/a/>/i/>/u/
F2	529	5	499	5	447	5	/i/>/u/>/a/
F3	405	5	344	5	395	5	/i/>/a/>/u/
F4	523	5	535	5	486	5	/u/>/i/>/a/
M1	289	5	277	5	283	5	/i/>/a/>/u/
M2	342	5	362	5	340	5	/u/>/i/>/a/
M3	283	5	282	5	277	5	/i/>/u/>/a/
M4	407	5	336	5	315	5	/i/>/u/>/a/
M5	562	5	600	5	619	5	/a/>/u/>/i/
M6	368	5	323	5	344	5	/i/>/a/>/u/
F	468	5	447	5	437	5	/i/>/u/>/a/
M	375	5	363	5	365	5	/i/>/a/>/u/
FB	447	5	419	5	398	5	/i/>/u/>/a/
MB	518	5	462	5	458	5	/i/>/u/>/a/

From the above table, we can see: (1) FB's average durations of /i/, /u/ and /a/ are 447ms, 419ms and 398ms, and MB's durations of these three vowels are 518ms, 462ms and 458ms, respectively. In other words, the orders of both FB's and MB's durations of these three vowels are /i/>/u/>/a/, which corresponds to the "i>a" principle. (2) Only F2's, M3's and M4's order of durations of /i/, /u/ and /a/ are the same as "model" native speakers', which is /i/>/u/>/a/. (3) The order of durations of /i/, /u/ and /a/ produced by F4 and M2 are in the same order, which is /u/>/i/>/a/. Since both /i/ and /u/ are high vowels, the order in /u/>/i/>/a/ makes sense. (4) 50% of Hungarian subjects produce deviant orders of durations of /i/, /u/ and /a/. F3's, M1's and M6's are in the order of /i/>/a/>/u/, F2's in /a/>/i/>/u/ and M5's in /a/>/u/>/i/. (4) The average durations of /i/, /u/ and /a/ produced by all Hungarian female subjects are sorted by /i/>/u/>/a/, which is identical to that of the "model" native speakers, while the average durations of /i/, /u/ and /a/ produced by all Hungarian male subjects are not. All in all, the deviant orders of durations of /i/, /u/ and /a/ indicate that the inappropriately producing vowels are likely to cause problems in duration.

It is worth mentioning that the average durations of /i/, /u/ and /a/ produced by all Hungarian female subjects (F) and male subjects (M) are 451ms [(467＋447＋437) ÷ 3], and 368ms [(375＋363＋365] ÷ 3], respectively. In other words, while the sample is too small to allow generalized statements, the data support the possibility that a connection may exist between duration and gender. This conclusion differs from the views expressed by Zhu Xiaonong in 2005[173], when he expressed the belief that there is no connection between duration and gender. However, Zhu Xiaonong was mainly concerned with Shanghainese, which differs in various ways (some regarding tone) from standard Mandarin, so the contradiction may not be important.

However, it should be emphasized that duration is a relative concept. It is hard to define a standard or universally "desirable" value of duration, because it depends so strongly on the individual's general rate of speech, not to mention the nature of each specific utterance and the conditions under which it is made. Discussing the duration of specific syllables within the parameters of an individual's own tone system probably makes better sense. This issue will be discussed later in this chapter.

5.1.2. Fundamental frequency

The fundamental frequency ("F0"or "F_0" in PRAAT-processed data), namely tone pitch, is a universal, indispensable and very distinctive feature of pronunciation in Chinese. In this section, firstly, each subject's F0 values for every *Yinping* monosyllable, together with the means for all the syllables, are presented in a combined table. Secondly, each subject's F0 contours for all the *Yinping* monosyllables are plotted in one chart for further analysis and comparison. In each chart the vertical (Y) axis shows the value of fundamental frequency, with Hz as the basic unit of measurement. The horizontal (X) axis presents the 11 measuring points. For better comparison purposes, the vertical axis covers a fixed value range based on

[173] Zhu 2005: 77.

the maximum and minimum frequencies recorded by female and male subjects respectively. For female subjects, the range is from 200Hz to 320Hz; for male subjects, from 100Hz to 200Hz. Thirdly, the average contour of all the subject's *Yinping* syllables is plotted in another chart. (For details of the methods used for measuring fundamental frequency (F0), see 3.2.1 above)

5.1.2.1. Individual subjects' F0 data

The F0 data table and charts of every subject will now be provided.

(1) F1

The following table shows the F0 data for subject F1, including 15 *Yinping* monosyllables. Each monosyllable has 11 F0 values which correspond to the 11 measuring points. The means of the measurements made at each point each point are presented in the final row of the table.

Table 32: *F1's F0 values for 15 Yinping syllables at 11 measuring points*

Syllables	1	2	3	4	5	6	7	8	9	10	11
都 dū	251	252	250	249	251	251	250	248	247	249	253
乌 wū	237	239	241	242	242	243	245	243	243	242	242
督 dū	252	250	248	248	247	247	247	247	246	246	245
阿 ā	235	231	232	230	228	227	227	227	227	227	227
答 dā	235	227	227	227	225	225	227	227	229	229	233
滴 dī	242	239	237	236	236	236	234	233	234	234	234
衣 yī	231	231	230	228	226	227	228	229	228	229	228
低 dī	241	237	232	230	229	231	229	228	228	229	230
屋 wū	236	239	243	244	244	242	241	242	240	241	240
逼 bī	255	252	248	246	246	245	245	245	245	244	238
一 yī	235	233	236	238	238	238	238	238	239	240	236
巴 bā	243	234	234	234	235	235	235	233	232	233	239
搭 dā	244	238	238	238	238	238	237	236	236	237	238
八 bā	238	235	238	237	234	233	232	233	236	240	245
逋 bū	248	246	244	242	242	242	242	242	242	244	239
Average	241.5	238.9	238.5	237.9	237.4	237.3	237.1	236.7	236.8	237.6	237.8

On the basis of the above table, we made the following two F0 contour charts of F1's *Yinping* pronunciation. The first chart shows the contour for every monosyllable, the second shows the mean of all the contours. The two charts are followed by brief comments on the data.

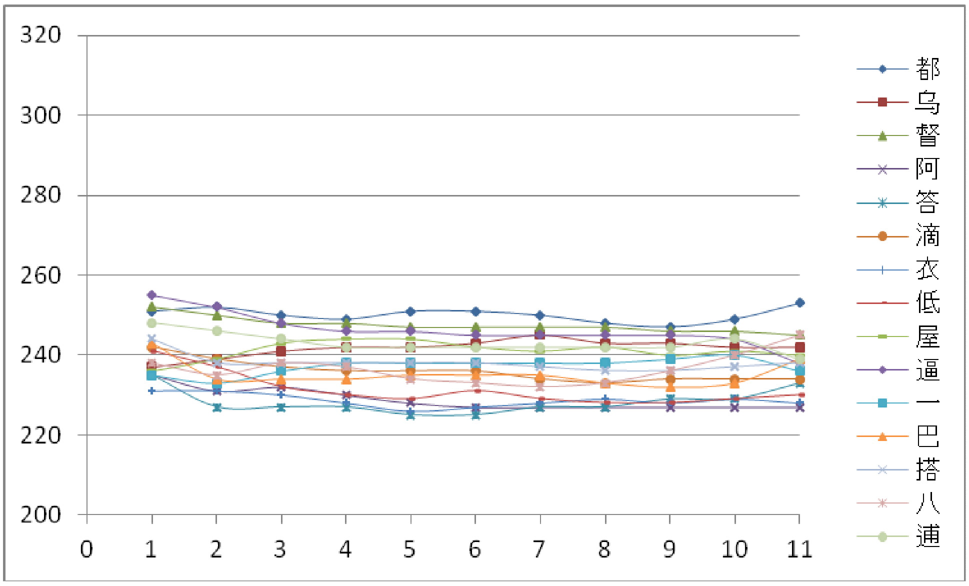


Figure 24. The F0 contour of 15 *Yinping* syllables pronounced by F1

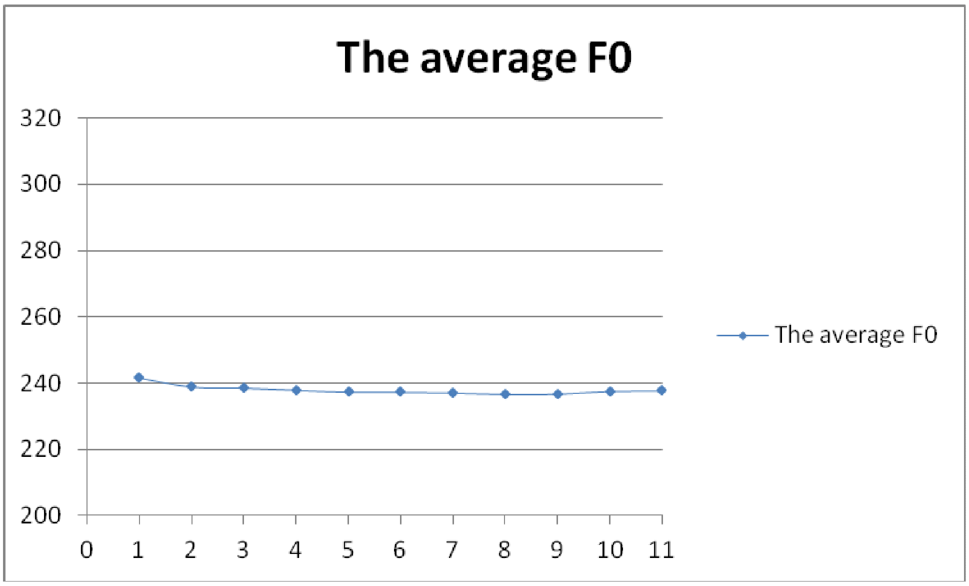


Figure 25. The average F0 contour of 15 *Yinping* syllables pronounced by F1

Based on Chart 1, the following conclusions can be made. (1) The F0 range of F1 is 225-255Hz, that is, about 30Hz. In other words, the fundamental frequencies of every monosyllable are relatively concentrated. (2) Overall, the main parts of the F0 contours of F1 are relatively straight and steady. (3) Some monosyllables have an obvious falling start, like

阿 ā, 答 dā, 低 dī, 逼 bī, 巴 bā, 搭 dā. (4) A few monosyllables have a clear rising ending, such as 都 dū, 答 dā, 巴 bā, 八 bā, while 逼 bī, 一 yī, 逋 bū have a palpable falling ending.

From Chart 2, we can see that the mean F0 contour of F1 is quite straight and steady, with a value of about 240Hz.

(2) F2

Table 33: *F2's F0 values for 15 Yinping syllables at 11 measuring points*

Syllables	1	2	3	4	5	6	7	8	9	10	11
都 dū	271	266	262	257	255	254	254	252	250	253	260
乌 wū	263	262	259	257	259	261	262	257	252	254	257
督 dū	268	262	258	256	255	255	256	255	254	252	254
阿 ā	262	264	262	258	254	252	250	251	252	250	251
答 dā	257	253	251	249	248	249	247	247	246	245	246
滴 dī	273	263	259	253	253	252	251	247	245	246	251
衣 yī	248	244	248	250	248	247	247	247	246	244	242
低 dī	270	264	255	248	246	246	247	246	247	246	246
屋 wū	259	267	267	265	265	265	266	264	263	261	267
逼 bī	263	259	254	254	255	254	252	250	250	251	251
一 yī	273	272	266	264	264	264	262	261	260	257	258
巴 bā	279	259	254	251	251	251	251	251	252	252	254
搭 dā	271	262	258	256	255	253	254	253	251	252	258
八 bā	278	267	264	261	259	257	256	254	254	255	256
逋 bū	279	269	260	256	256	256	256	256	257	257	259
Average	267.6	262.2	258.5	255.7	254.9	254.4	254.1	252.7	251.9	251.7	254

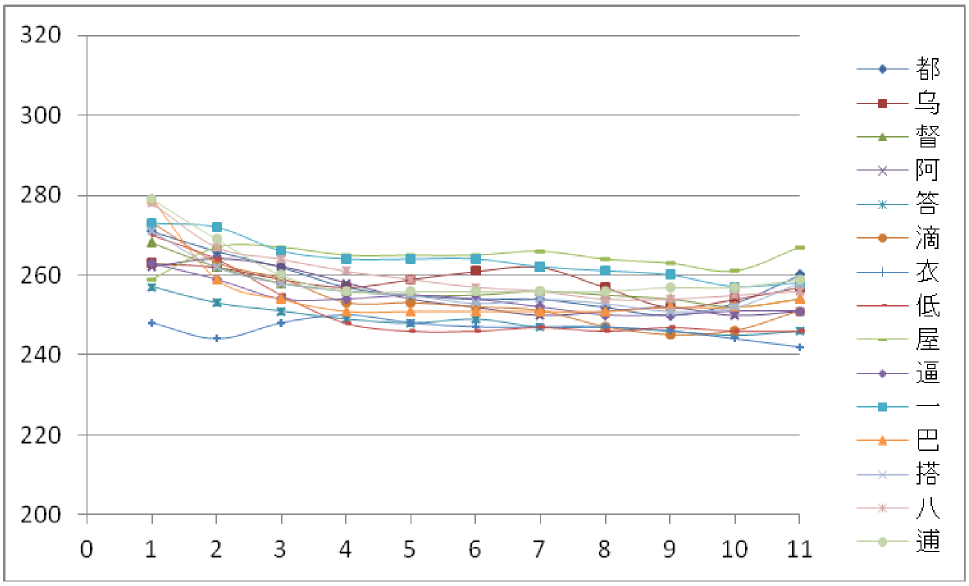


Figure 26. The F0 contour of 15 *Yinping* syllables pronounced by F2

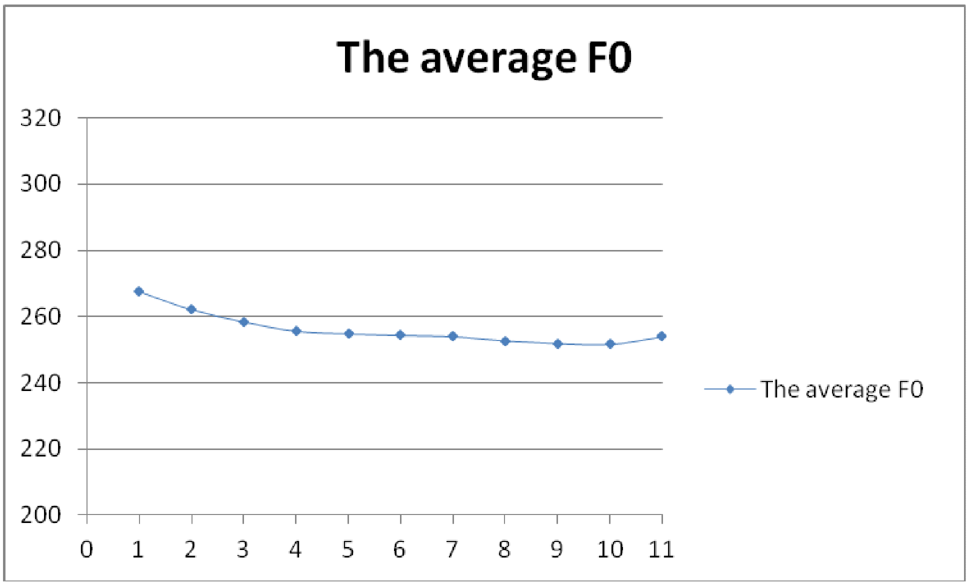


Figure 27. The average F0 contour of 15 *Yinping* syllables pronounced by F2

Based on Chart 1, the following conclusions can be made. (1) The F0 range of F2 is 240-280Hz, that is, about 40Hz. In other words, the fundamental frequencies of every monosyllable are relatively concentrated. (2) Overall, the F0 contours of F2 are not straight and steady. (3) Most monosyllables have an obvious falling start, except 衣 yīand 屋 wū, and 屋 wū has a palpable rising start. (4) Some monosyllables have a clear rising ending, such as 都 dū, 乌 wū, 屋 wū, 搭 dā.

From Chart 2, we can see that the mean F0 contour of F2 is not straight and steady, with a value of 250-270Hz.

(3) F3

Table 34: *F3's F0 values for 15 Yinping syllables at 11 measuring points*

Syllables	1	2	3	4	5	6	7	8	9	10	11
都 dū	248	249	250	250	251	252	252	253	257	258	256
乌 wū	242	245	247	246	245	246	247	248	250	252	253
督 dū	266	264	263	261	258	255	255	254	255	256	257
阿 ā	244	239	238	234	232	231	234	236	236	236	236
答 dā	233	226	226	227	224	221	222	223	224	227	229
滴 dī	241	232	232	232	230	227	226	226	229	229	230
衣 yī	238	235	235	233	230	231	231	234	234	235	232
低 dī	231	231	236	235	234	230	230	229	230	232	232
屋 wū	233	238	238	234	228	224	225	228	230	234	236
逼 bī	235	236	234	232	231	231	230	231	234	234	235
一 yī	241	241	240	239	238	239	240	241	240	240	244
巴 bā	226	224	226	223	221	223	222	220	224	223	225
搭 dā	204	202	205	204	207	213	215	215	218	222	223
八 bā	229	226	226	223	223	224	222	221	219	221	226
逋 bū	242	245	248	248	246	244	242	243	244	247	248
Average	236.9	235.5	236.3	234.7	233.2	232.7	232.9	233.5	234.9	236.4	237.5

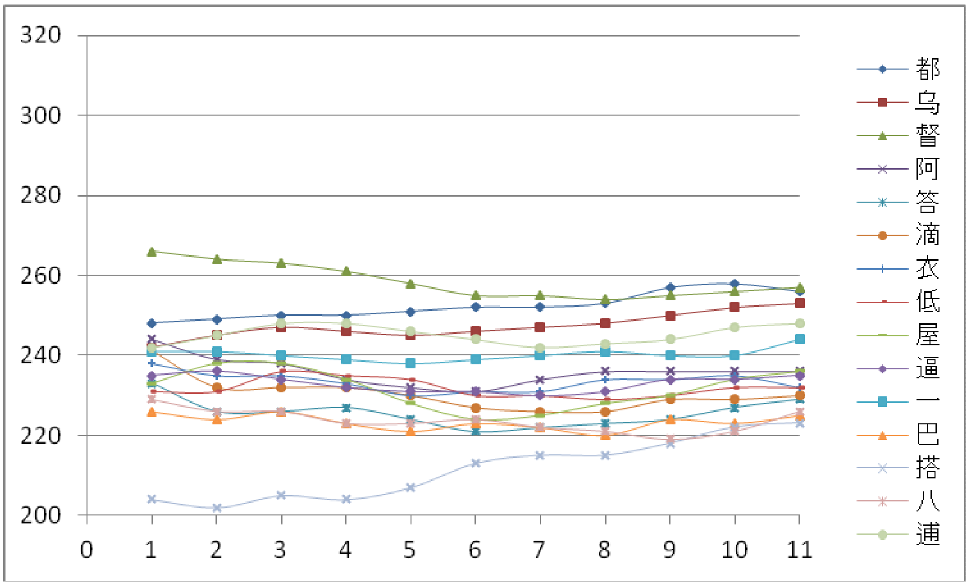


Figure 28. The F0 contour of 15 *Yinping* syllables pronounced by F3

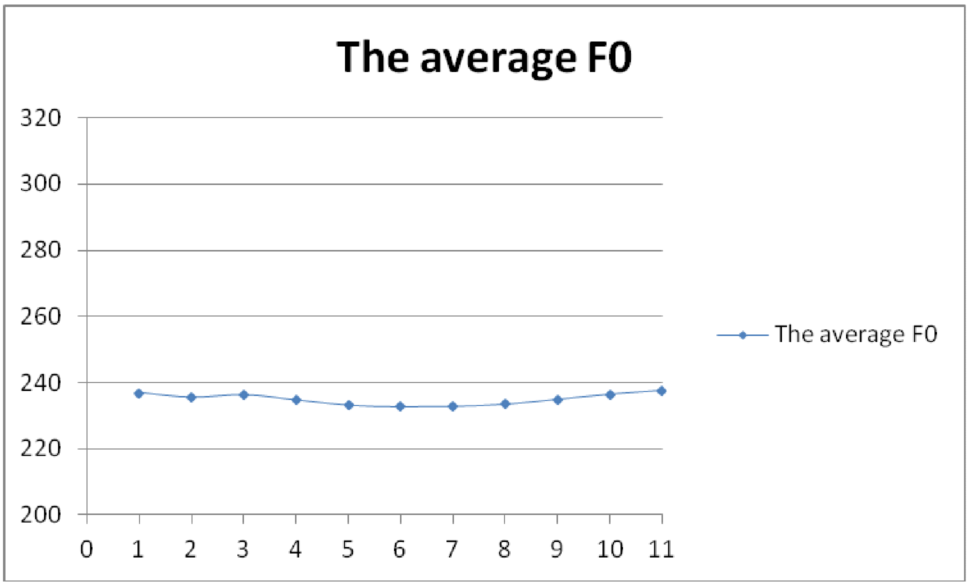


Figure 29. The average F0 contour of 15 *Yinping* syllables pronounced by F3

Based on Table X and Figure 1, the following conclusions can be made. (1) The F0 range of F3 is 200-270Hz, that is, about 70Hz. In other words, the fundamental frequencies of every monosyllable are relatively widely spread. (2) Overall, most F0 contours of F3 are not straight and steady. (3) Most monosyllables have an obvious falling start, such as 阿 ā, 答 dā, 滴 dī. (4) Some monosyllables have a clear rising ending, like 答 dā, 逼 bī, 八 bā. (5) 督 dū and 搭 dā are two exceptions. The former has a much higher start, then falls gradually; the latter has quite a low start, then rises step by step.

From Figure2, we can see that the mean F0 contour of F3 is relatively straight and steady, with a value of about 240Hz. However, in this case, the mean F0 contour represents the average of very distant values and its significance is therefore open to doubt.

(4) F4

Table 35: *F4's F0 values for 15 Yinping syllables at 11 measuring points*

Syllables	1	2	3	4	5	6	7	8	9	10	11
都 dū	315	316	315	313	310	312	312	312	311	314	317
乌 wū	315	313	310	310	307	306	307	306	308	308	307
督 dū	313	319	316	312	312	311	310	311	310	311	311
阿 ā	277	284	286	288	290	290	290	290	289	292	296
答 dā	298	287	285	289	289	290	292	291	291	294	302
滴 dī	316	308	304	306	306	304	304	305	303	303	303

衣 yī	306	303	303	303	301	302	302	304	302	306	301
低 dī	289	278	280	286	291	295	296	300	301	300	299
屋 wū	315	309	310	311	312	315	314	313	310	315	311
逼 bī	307	309	306	305	305	307	307	308	306	309	313
一 yī	294	298	297	296	298	301	299	297	304	296	303
巴 bā	310	297	289	290	292	288	286	288	292	294	294
搭 dā	280	266	266	268	274	276	274	274	274	274	281
八 bā	302	291	293	299	298	298	300	302	302	304	307
逋 bū	315	313	312	305	302	303	302	304	303	303	298
Average	303.5	299.4	298.1	298.7	299.1	299.9	299.7	300.3	300.4	301.5	302.9

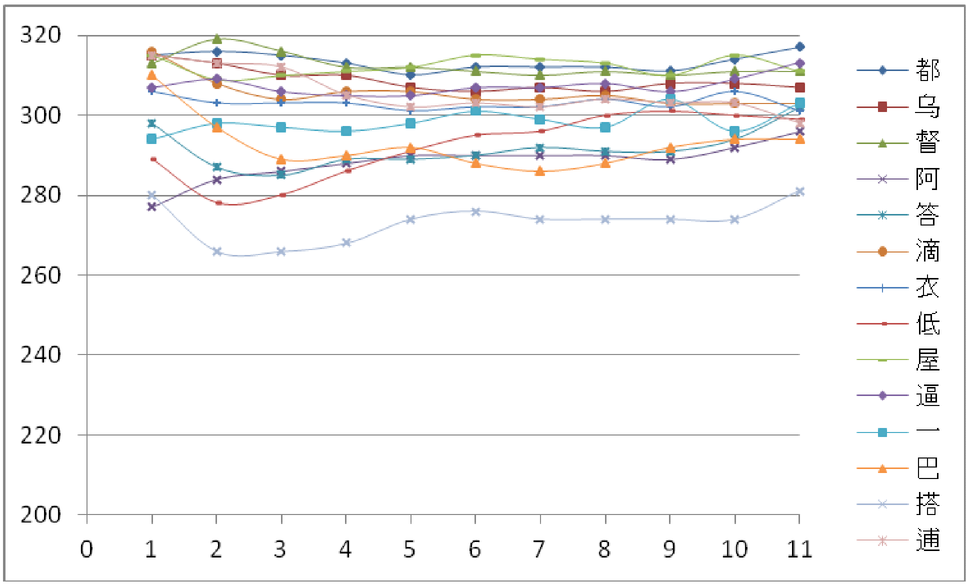


Figure 30. The F0 contour of 15 *Yinping* syllables pronounced by F4

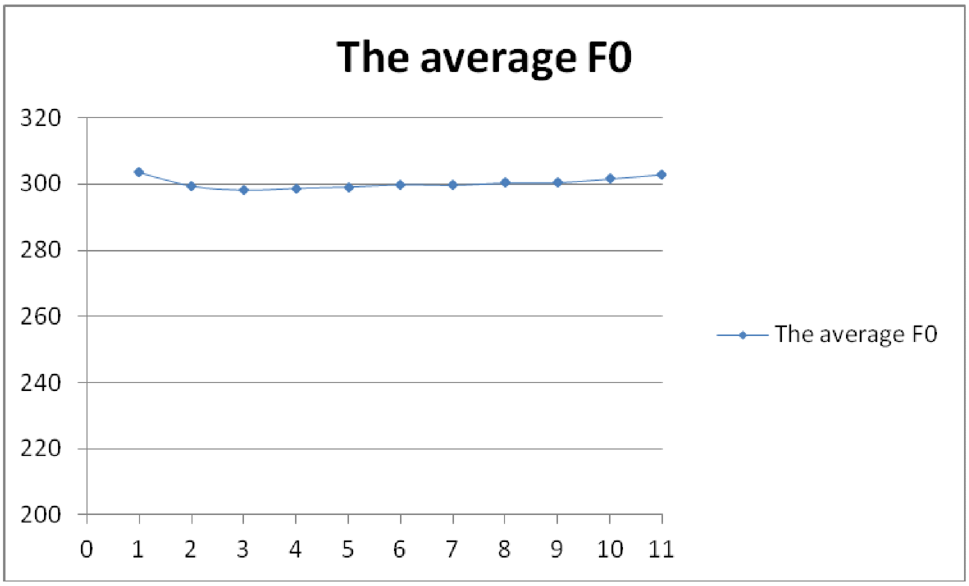


Figure 31. The average F0 contour of 15 *Yinping* syllables pronounced by F4

Based on Table X and Figure 1, the following conclusions can be made. (1) The F0 range of F4 is 260-320Hz, that is, about 60Hz. In other words, the fundamental frequencies of every monosyllable are widely spread. (2) Overall, most F0 contours of F4 are not straight and steady, but fluctuating. (3) Many monosyllables have a very obvious falling start, such as 答 dā, 滴 dī, 低 dī, 巴 bā, 屋 wū, 搭 dā, 逋 bū, while 督 dū, 阿 ā, 逼 bī, 一 yī have quite a palpable rising start. (4) Some monosyllables have a clear rising ending, such as 都 dū, 逼 bī, 答 dā, 阿 ā, 搭 dā, while 屋 wū and 逋 bū have a falling ending. 一 yīhas a fluctuating rising ending, while 衣 yīhas a fluctuant falling ending.

From Figure2, we can see that the mean F0 contour of F4 is quite straight and steady, except for a falling start, with a value of about 300Hz. As with the previous subject, in this case mean F0 contour represents the average of very distant values and its significance is therefore open to doubt.

(5) M1

Table 36: *M1's F0 values for 15 Yinping syllables at 11 measuring points*

Syllables	1	2	3	4	5	6	7	8	9	10	11
都 dū	139	137	135	135	135	135	134	135	137	139	134
乌 wū	130	133	131	132	133	136	137	137	136	133	137
督 dū	138	134	134	134	135	135	135	135	135	138	144

阿 ā	135	135	137	137	136	135	134	131	131	135	142
答 dā	135	131	132	131	131	130	132	132	136	136	134
滴 dī	133	140	137	138	140	139	138	140	141	144	143
衣 yī	144	145	145	142	141	142	144	143	142	144	141
低 dī	137	139	140	141	141	138	136	137	138	141	140
屋 wū	132	140	139	138	141	143	142	139	136	138	145
逼 bī	144	144	145	148	148	146	144	146	146	144	142
一 yī	142	142	141	143	144	143	144	146	143	141	146
巴 bā	134	138	136	138	137	135	136	136	135	135	138
搭 dā	135	134	135	133	133	132	131	133	135	139	138
八 bā	134	134	132	133	134	135	135	138	136	135	137
逋 bū	155	151	149	147	147	147	150	153	151	148	145
Average	137.8	138.5	137.9	138	138.4	138.1	138.1	138.7	138.5	139.3	140.4

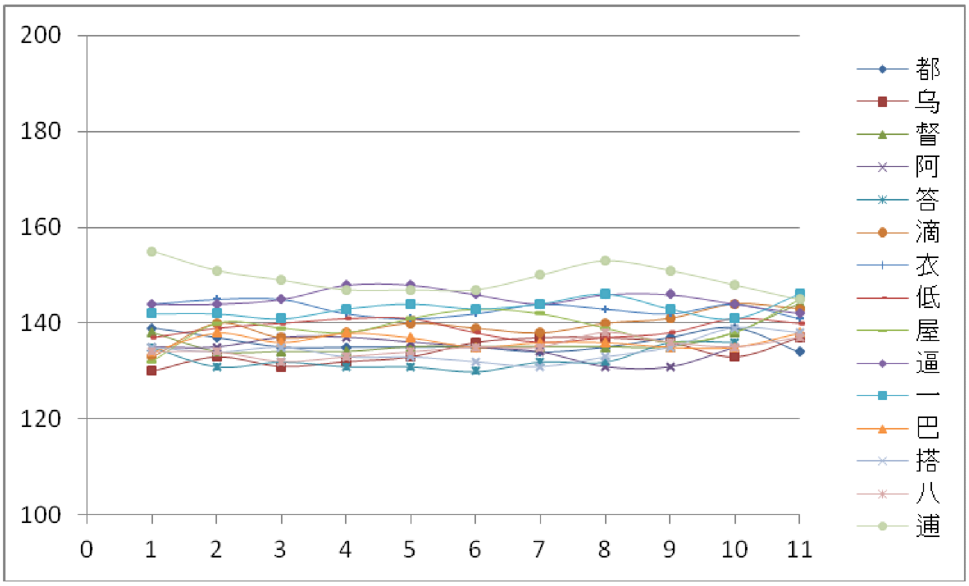


Figure 32. The F0 contour of 15 *Yinping* syllables pronounced by M1

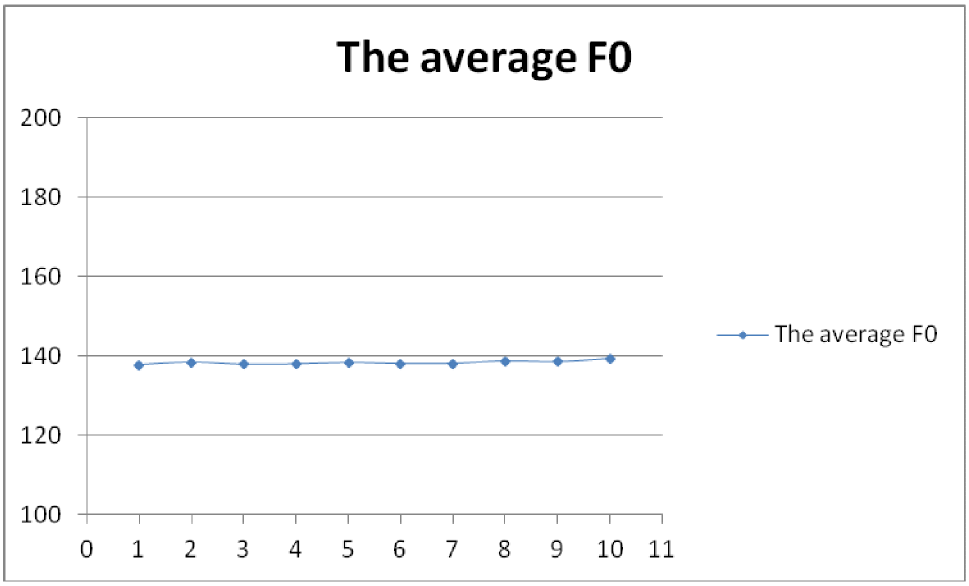


Figure 33. The average F0 contour of 15 *Yinping* syllables pronounced by M1

Based on Chart 1, the following conclusions can be made. (1) The F0 range of M1 is 130-160Hz, that is, about 30Hz. In other words, the fundamental frequencies of every monosyllable are relatively concentrated. (2) Overall, the main parts of the F0 contours of M1 are relatively steady. (3) A few monosyllables have an obvious falling start, like 滴 dīand 屋 wū. (4) A few monosyllables have a clear rising ending, such as 督 dū, 阿 ā, 屋 wū. (5) 逋 bū is an exception which has a higher F0 and a relatively fluctuating F0 contour.

From Chart 2, we can see that the mean F0 contour of M1 is very straight and steady, with a value of about 140Hz.

(6) M2

Table 37: *M2's F0 values for 15 Yinping syllables at 11 measuring points*

Syllables	1	2	3	4	5	6	7	8	9	10	11
都 dū	133	127	129	129	129	130	130	129	128	129	129
乌 wū	131	129	130	132	132	131	132	131	130	129	131
督 dū	125	125	125	124	124	124	124	125	123	125	127
阿 ā	127	122	123	122	121	123	123	125	126	127	129
答 dā	127	125	127	126	126	126	126	128	129	131	134
滴 dī	125	127	130	129	127	126	126	126	126	126	126
衣 yī	124	121	121	122	123	125	125	125	124	123	125
低 dī	120	119	121	124	125	125	123	123	124	122	122

屋 wū	130	129	131	133	132	131	130	131	131	132	133
逼 bī	123	120	123	126	128	129	130	129	128	127	129
一 yī	127	125	126	128	130	129	129	128	127	126	125
巴 bā	121	127	127	128	130	131	131	130	129	128	129
搭 dā	120	126	128	129	129	129	130	131	131	131	135
八 bā	132	125	126	126	127	127	126	126	126	127	127
逋 bū	124	126	131	133	133	131	130	130	128	129	128
Average	125.9	124.9	126.5	127.4	127.7	127.8	127.7	127.8	127.3	127.5	128.6

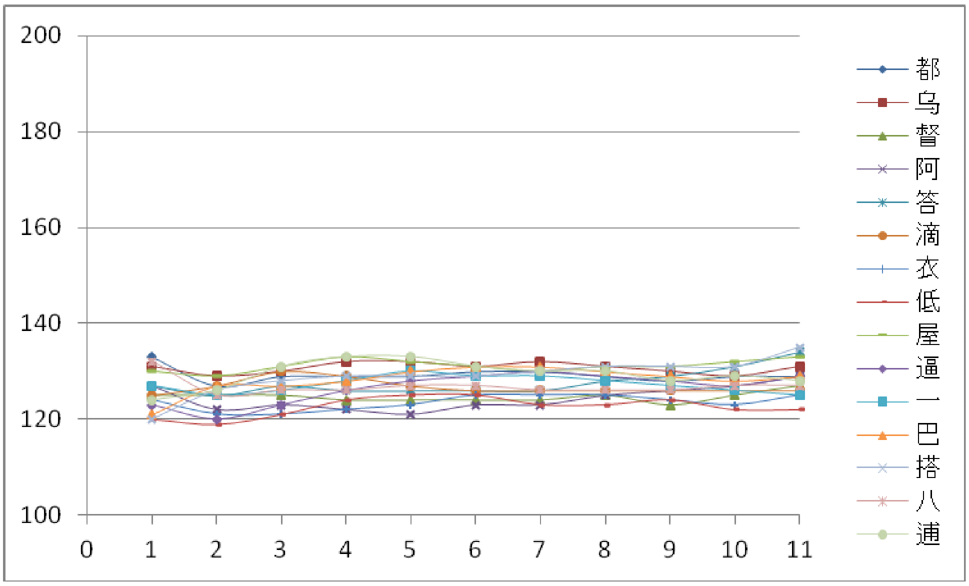


Figure 34. The F0 contour of 15 *Yinping* syllables pronounced by M2

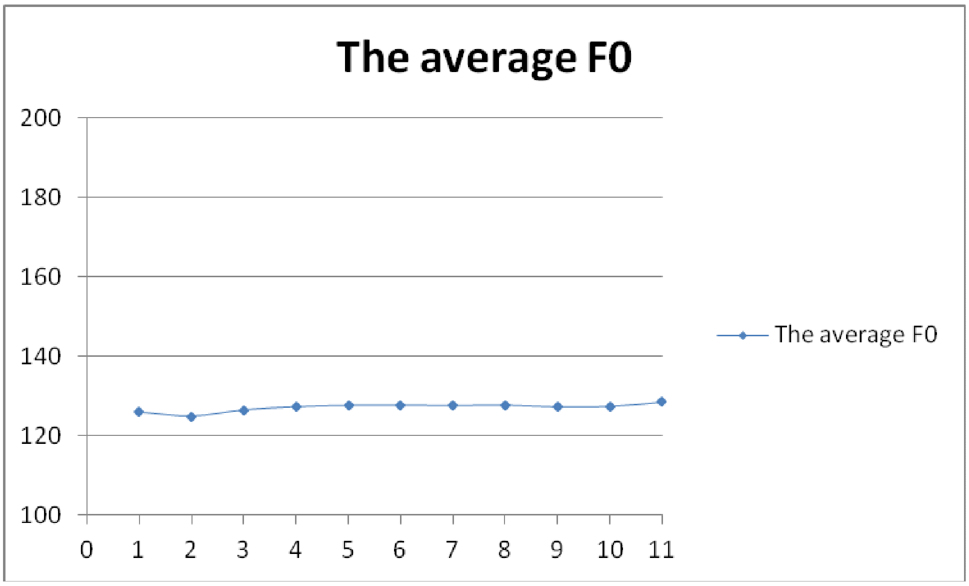


Figure 35. The average F0 contour of 15 *Yinping* syllables pronounced by M2

Based on Chart 1, the following conclusions can be made. (1) The F0 range of M2 is 120-140Hz, that is, about 20Hz. In other words, the fundamental frequencies of every monosyllable are highly concentrated. (2) Overall, the F0 contours of M2 are quite straight and steady. (3) Only a few monosyllables have an obvious falling start.

From Chart 2, we can see that the mean F0 contour of M2 is quite straight and steady, with a value of about 130Hz.

(7) M3

Table 38: *M3's F0 values for 15 Yinping syllables at 11 measuring points*

Syllables	1	2	3	4	5	6	7	8	9	10	11
都 dū	125	124	125	123	124	126	128	130	130	131	132
乌 wū	132	119	125	127	126	127	128	127	129	131	134
督 dū	123	122	123	125	125	124	123	123	123	128	130
阿 ā	125	124	123	122	123	123	125	123	123	123	120
答 dā	126	118	119	121	123	125	124	124	126	131	136
滴 dī	121	123	125	126	124	123	124	124	127	130	131
衣 yī	122	121	121	122	122	123	125	124	125	128	129
低 dī	129	126	127	126	124	123	125	123	125	128	130
屋 wū	120	124	123	124	125	126	126	125	126	129	128
逼 bī	128	129	126	126	126	125	127	126	126	130	134
一 yī	135	129	130	133	133	133	130	129	130	130	133
巴 bā	120	122	124	125	126	127	127	126	126	129	136
搭 dā	120	117	117	116	117	117	119	120	123	128	123
八 bā	119	116	114	115	114	114	116	118	117	117	116
逋 bū	127	130	129	129	129	131	133	132	133	136	136
Average	124.8	122.9	123.4	124	124.1	124.5	125.3	124.9	125.9	128.6	129.9

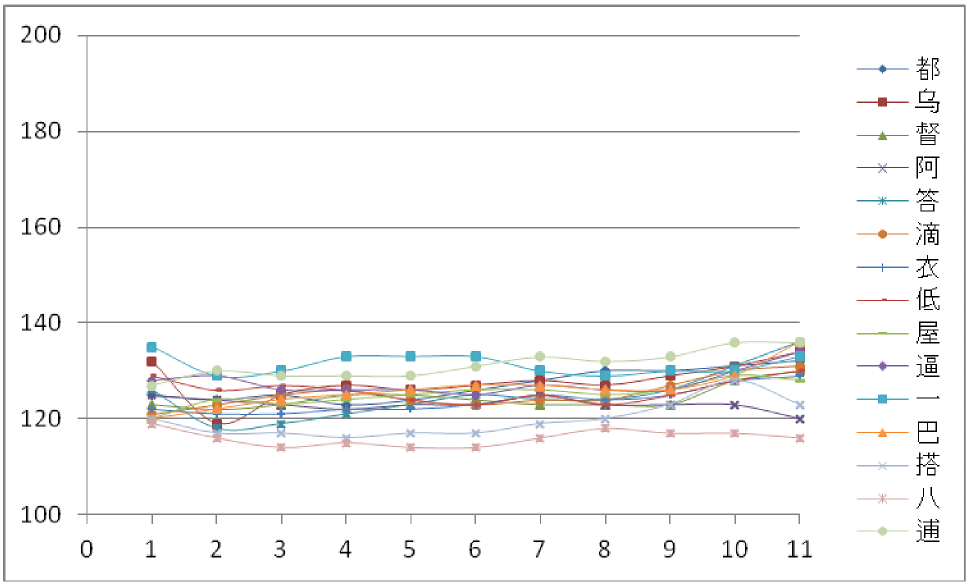


Figure 36. The F0 contour of 15 *Yinping* syllables pronounced by M3

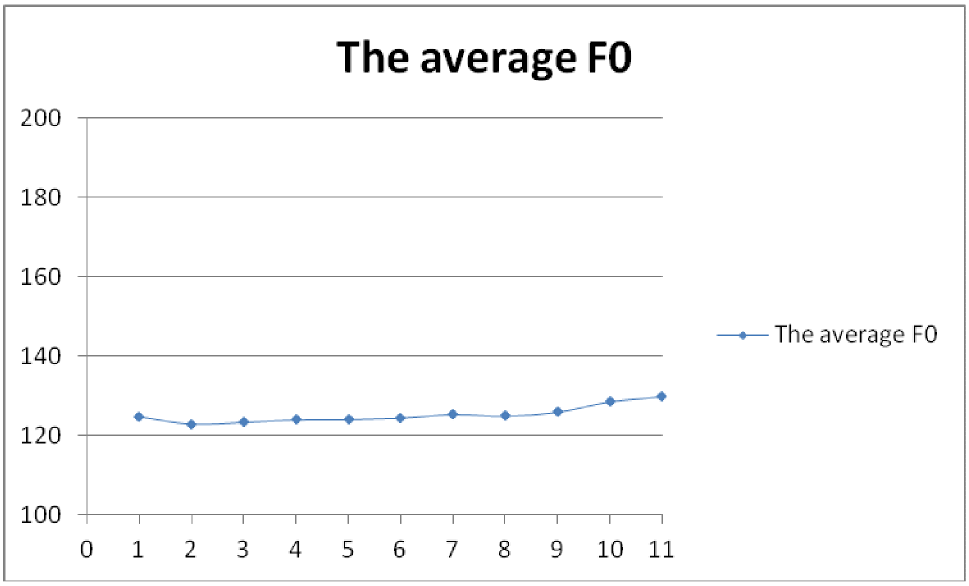


Figure 37. The average F0 contour of 15 *Yinping* syllables pronounced by M3

Based on Chart 1, the following conclusions can be made. (1) The F0 range of M3 is 115-135Hz, that is, about 20Hz. In other words, the fundamental frequencies of every monosyllable are highly concentrated. (2) Overall, the main parts of the F0 contours of M3 are relatively straight and steady. (3) Most monosyllables have a relatively "loose" start or ending.

From Chart 2, we can see that the mean F0 contour of M3 is relatively straight and steady, except for a rising ending, with a value of about 130Hz.

(8) M4

Table 39: *M4's F0 values for 15 Yinping syllables at 11 measuring points*

Syllables	1	2	3	4	5	6	7	8	9	10	11
都 dū	143	139	137	136	135	135	134	133	133	132	131
乌 wū	127	130	132	135	136	137	135	132	130	130	133
督 dū	134	133	132	132	132	132	131	131	132	131	130
阿 ā	120	125	126	126	125	125	124	123	123	124	123
答 dā	127	113	114	114	114	113	112	112	112	114	116
滴 dī	140	137	134	134	134	134	134	134	134	133	136
衣 yī	121	123	128	132	133	134	133	133	132	135	139
低 dī	131	130	130	131	131	131	129	130	130	131	134
屋 wū	129	127	128	128	129	129	130	131	131	130	132
逼 bī	150	148	145	142	141	140	139	138	137	137	138
一 yī	118	117	121	128	132	132	130	130	129	128	128
巴 bā	130	128	125	123	122	121	122	122	123	125	127
搭 dā	129	128	128	128	127	127	128	128	128	128	128
八 bā	148	136	135	135	135	134	133	134	133	133	130
逋 bū	142	140	138	137	136	135	134	134	133	133	135
Average	132.6	130.3	130.2	130.7	130.8	130.6	129.9	129.7	129.3	129.6	130.7

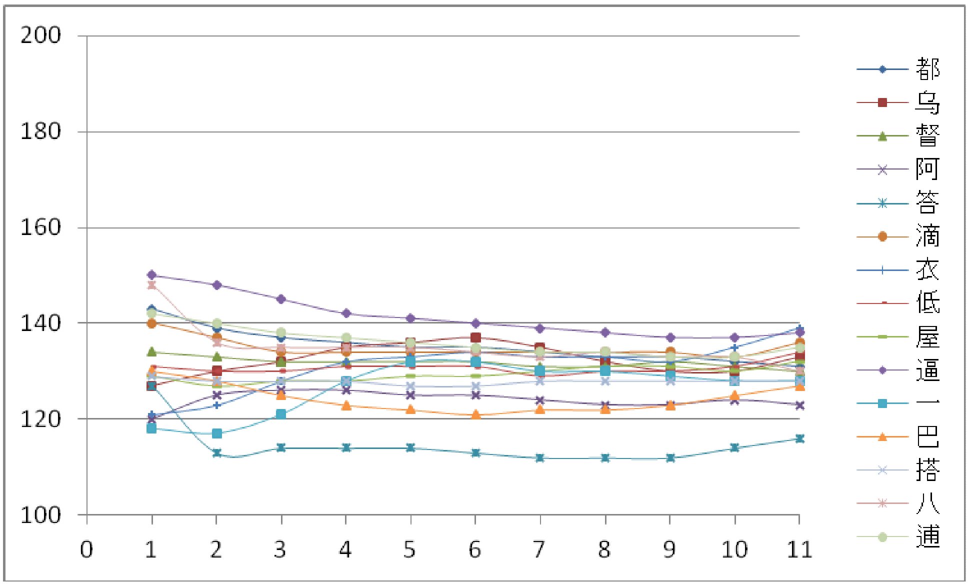


Figure 38. The F0 contour of 15 *Yinping* syllables pronounced by M4

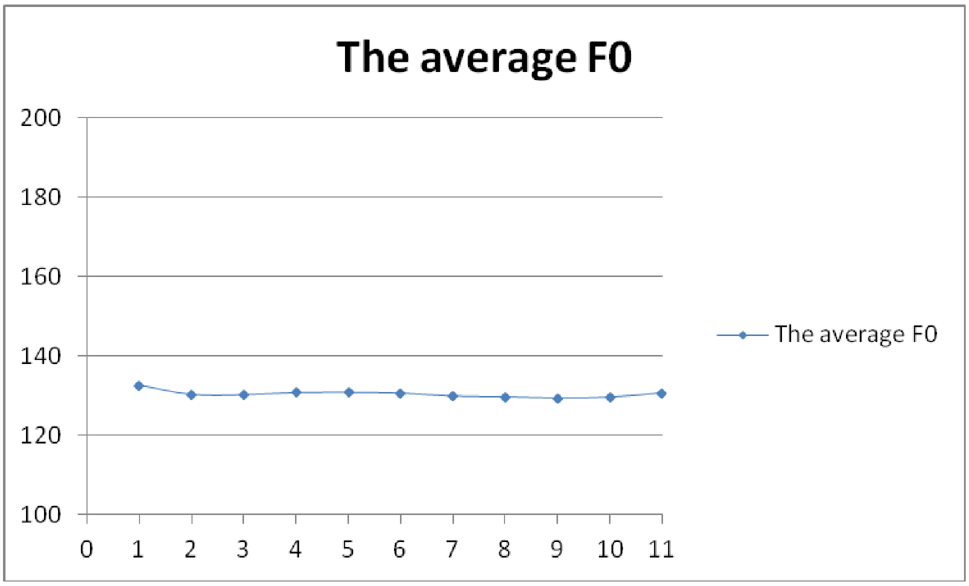


Figure 39. The average F0 contour of 15 *Yinping* syllables pronounced by M4

Based on Chart 1, the following conclusions can be made. (1) The F0 range of M4 is 110-150Hz, that is, about 40Hz. In other words, the fundamental frequencies of every monosyllable are relatively concentrated. (2) Overall, the main parts of the F0 contours of M4 are quite straight and steady. (3) A few monosyllables have a relatively loose start.

From Chart 2, we can see that the mean F0 contour of M4 is very straight and steady, with a value of about 130Hz.

(9) M5

Table 40: *M5's F0 values for 15 Yinping syllables at 11 measuring points*

Syllables	1	2	3	4	5	6	7	8	9	10	11
都 dū	171	173	173	172	172	171	172	172	174	175	172
乌 wū	129	140	153	164	170	171	172	171	172	173	173
督 dū	179	181	179	179	179	178	179	181	182	180	179
阿 ā	151	160	164	166	166	166	166	168	168	167	170
答 dā	171	168	166	167	167	167	167	167	167	168	169
滴 dī	176	179	177	176	176	176	175	175	173	176	178
衣 yī	157	165	172	174	177	178	177	177	178	179	178
低 dī	174	175	173	174	174	175	173	174	175	174	173
屋 wū	148	161	178	191	189	187	186	188	187	186	187
逼 bī	186	187	186	184	183	183	183	183	183	184	184
一 yī	149	160	169	173	175	175	177	178	178	178	179

巴 bā	161	159	163	164	165	165	165	166	167	168	168
搭 dā	173	172	171	171	173	173	172	172	173	173	174
八 bā	171	170	170	169	169	171	170	171	170	170	169
逋 bū	178	181	183	181	181	179	180	180	180	180	179
Average	164.9	168.7	171.8	173.7	174.4	174.3	174.3	174.9	175.1	175.4	175.5

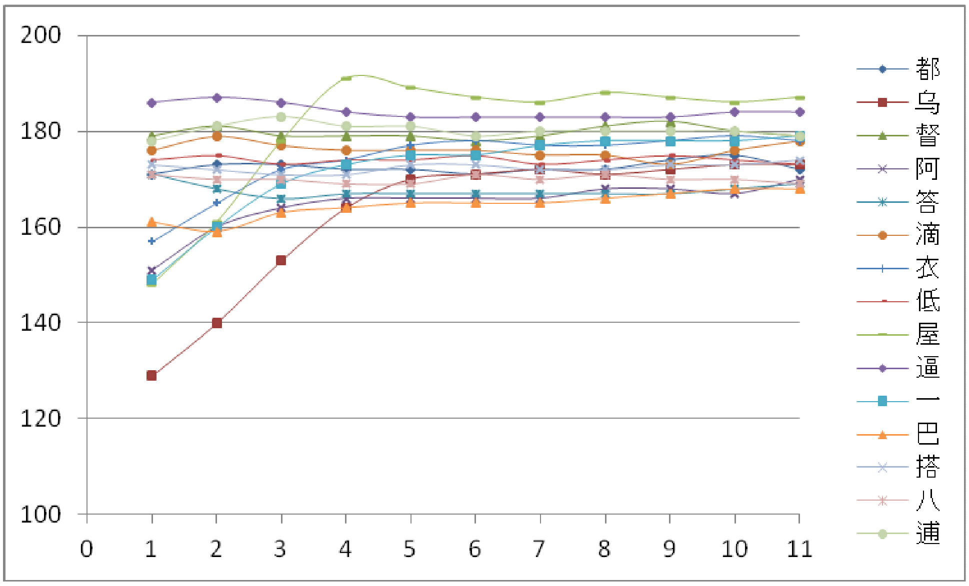


Figure 40. The F0 contour of 15 *Yinping* syllables pronounced by M5

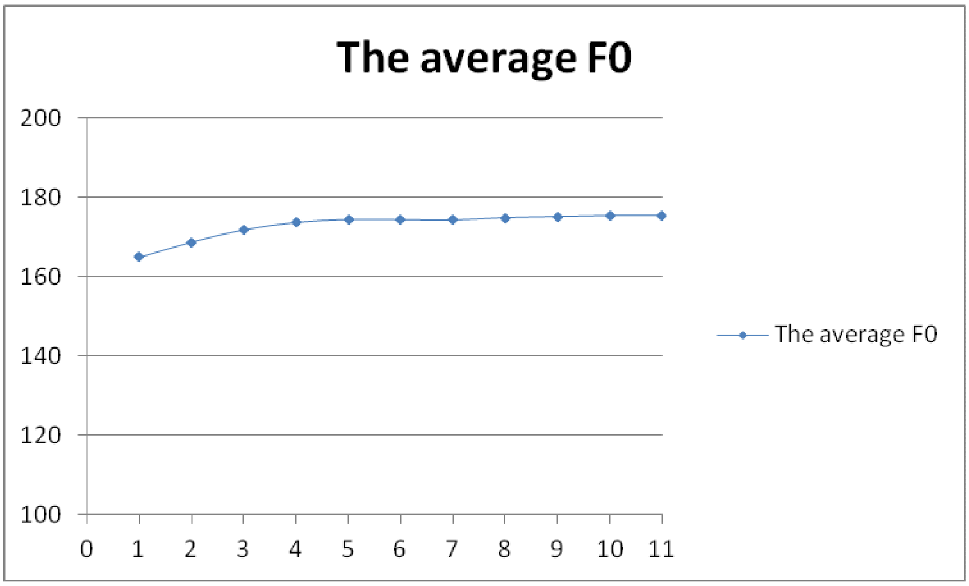


Figure 41. The average F0 contour of 15 *Yinping* syllables pronounced by M5

Based on Chart 1, the following conclusions can be made. (1) The F0 range of M5 is 130-190Hz, that is, about 60Hz. In other words, the fundamental frequencies of every monosyllable are widely spread. (2) Overall, the main parts of the F0 contours of M5 are relatively straight and steady. (3) Most monosyllables have a loose start, especially 乌 wū and 屋 wū which have a clear rising start.

From Chart 2, we can see that the mean F0 contour of M5 is not straight and steady because of the rising start, with a value range of 160-180Hz.

(10) M6

Table 41: *M6's F0 values for 15 Yinping syllables at 11 measuring points*

Syllables	1	2	3	4	5	6	7	8	9	10	11
都 dū	122	121	121	122	124	125	125	124	124	121	125
乌 wū	118	123	124	125	124	122	121	122	121	121	123
督 dū	121	119	119	120	120	121	122	122	123	125	126
阿 ā	120	117	117	113	112	113	112	109	108	108	107
答 dā	125	122	118	117	117	117	118	119	119	121	121
滴 dī	128	125	122	121	121	121	121	121	122	121	122
衣 yī	115	119	124	126	127	125	125	124	123	123	122
低 dī	115	117	116	117	118	119	120	120	121	117	124
屋 wū	116	122	126	127	127	128	127	126	125	124	126
逼 bī	120	121	122	122	123	122	122	122	120	120	117
一 yī	109	113	116	119	121	123	124	124	123	123	123
巴 bā	113	115	113	114	114	114	114	114	115	117	117
搭 dā	118	117	115	115	115	115	116	116	117	116	115
八 bā	114	116	114	113	114	114	115	115	114	116	117
逋 bū	126	129	128	129	130	130	129	127	126	124	124
Average	118.7	119.7	119.7	120	120.5	120.6	120.7	120.3	120.1	119.8	120.6

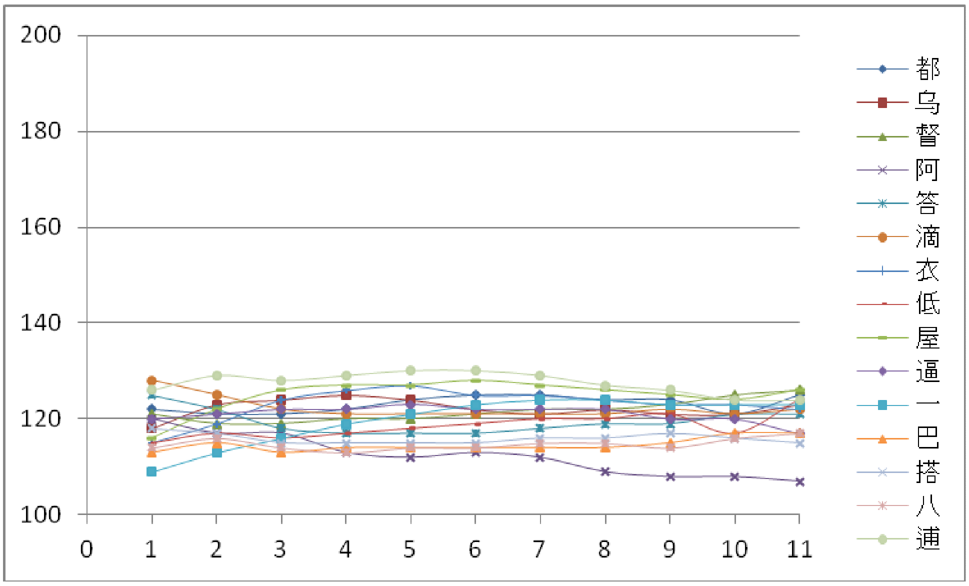


Figure 42. The F0 contour of 15 *Yinping* syllables pronounced by M6

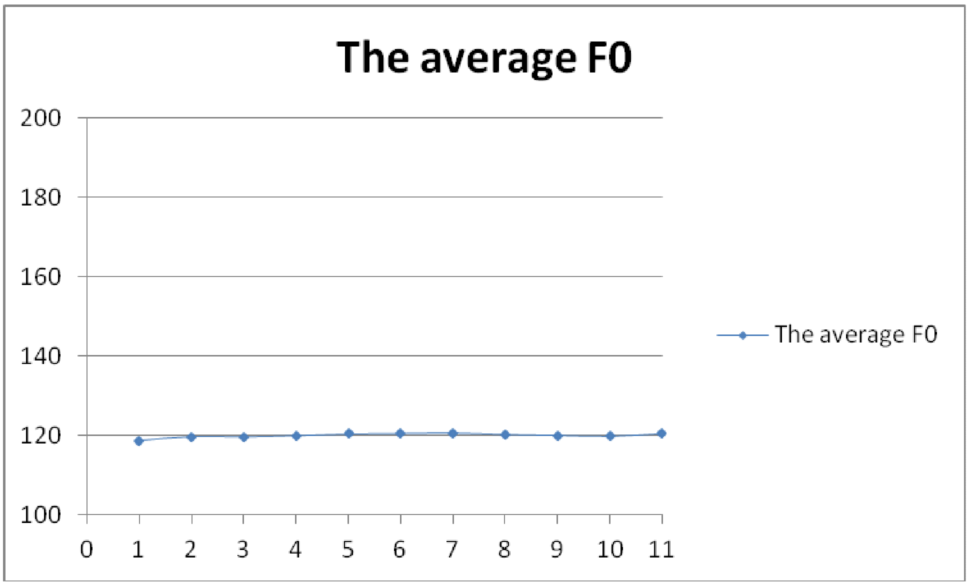


Figure 43. The average F0 contour of 15 *Yinping* syllables pronounced by M6

Based on Chart 1, the following conclusions can be made. (1) The F0 range of M6 is 110-130Hz, that is, about 20Hz. In other words, the fundamental frequencies of every monosyllable are highly concentrated. (2) Overall, the main parts of the F0 contours of M6 are quite straight and steady. (3) Some monosyllables have a relatively loose start or ending.

From Chart2, we can see that the mean F0 contour of M6 is very straight and steady, with a value of about 120Hz.

(11) FB

Table 42 :*FB's F0 values for 15 Yinping syllables at 11 measuring points*

Syllables	1	2	3	4	5	6	7	8	9	10	11
都 dū	216	237	234	234	233	234	233	229	235	236	242
乌 wū	226	225	227	226	223	224	224	228	229	228	226
督 dū	246	232	228	225	224	224	224	225	231	231	231
阿 ā	205	206	211	210	214	212	209	210	209	210	214
答 dā	219	228	221	225	223	220	221	220	223	224	224
滴 dī	236	230	229	227	226	230	227	225	225	226	224
衣 yī	235	229	231	229	229	228	225	225	228	225	219
低 dī	235	236	232	229	227	226	226	225	225	224	225
屋 wū	228	229	232	232	231	231	231	228	228	235	236
逼 bī	222	225	224	224	221	220	220	224	220	218	216
一 yī	214	228	235	237	239	239	238	239	239	240	234
巴 bā	234	232	230	231	229	227	229	228	230	231	233
搭 dā	236	234	231	230	226	225	228	228	226	228	234
八 bā	225	223	221	219	219	218	215	217	215	219	223
逋 bū	235	236	235	235	229	228	228	228	231	229	228
Average	227.5	228.7	228.1	227.5	226.2	225.7	225.2	225.3	226.3	226.9	227.3

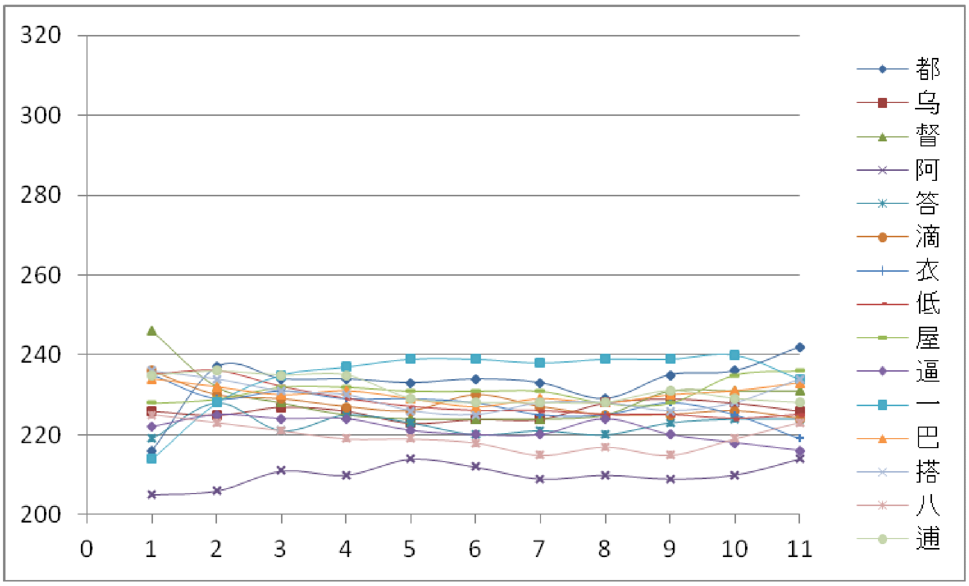


Figure 44. The F0 contour of 15 *Yinping* syllables pronounced by FB

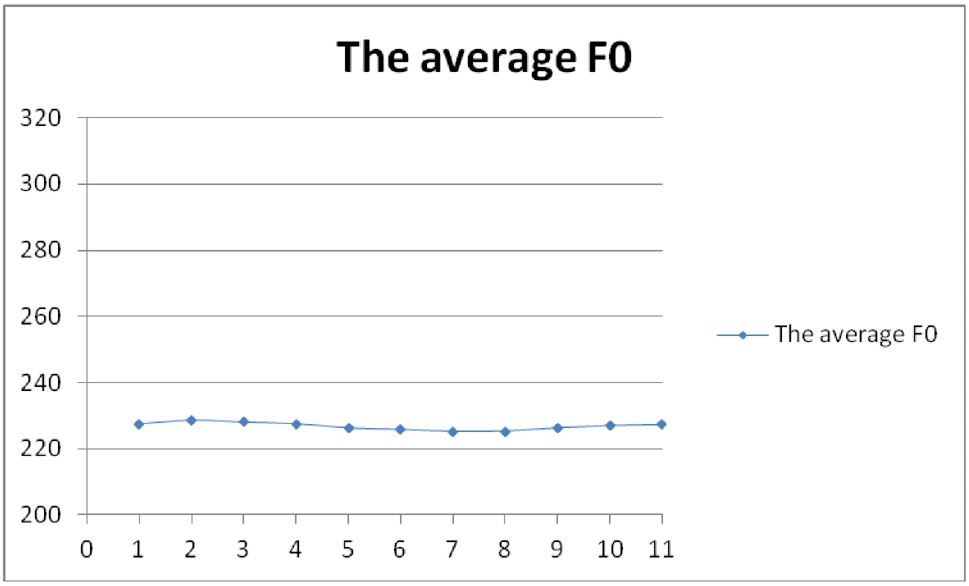


Figure 44, the average F0 contour of 15 *Yinping* syllables pronounced by FB

Based on Chart 1, the following conclusions can be made. (1) The F0 range of FB is 205-250Hz, that is, about 45Hz. In other words, the fundamental frequencies of every monosyllable are relatively concentrated. (2) Overall, the main parts of the F0 contours of FB are quite straight and steady. (3) A few monosyllables have a relatively loose start or ending.

From Chart 2, we can see that the mean F0 contour of FB is very straight and steady, with a value of about 230Hz.

(12) MB

Table 43: *MB's F0 values for 15 Yinping syllables at 11 measuring points*

Syllables	1	2	3	4	5	6	7	8	9	10	11
都 dū	124	123	123	120	118	118	117	117	116	117	118
乌 wū	112	112	112	112	113	113	113	113	113	114	116
督 dū	111	110	110	111	112	112	111	110	109	108	105
阿 ā	121	123	123	121	122	121	120	120	119	120	119
答 dā	133	134	134	132	132	133	132	132	128	126	127
滴 dī	137	135	132	134	134	133	131	131	130	130	135
衣 yī	128	126	129	131	131	130	131	132	130	128	128
低 dī	132	132	132	132	132	131	131	130	129	127	128
屋 wū	134	133	133	135	135	134	134	136	137	136	131
逼 bī	137	139	138	138	138	137	136	139	141	139	134
一 yī	133	135	137	136	137	138	136	135	136	140	133

巴 bā	130	129	128	128	129	127	127	127	127	127	125
搭 dā	146	143	142	142	142	143	142	140	140	139	141
八 bā	144	142	141	141	142	140	140	138	137	136	137
逋 bū	147	151	151	151	152	149	148	149	149	148	144
Average	131.3	131.1	131	130.9	131.3	130.6	129.9	129.9	129.4	129	128.1

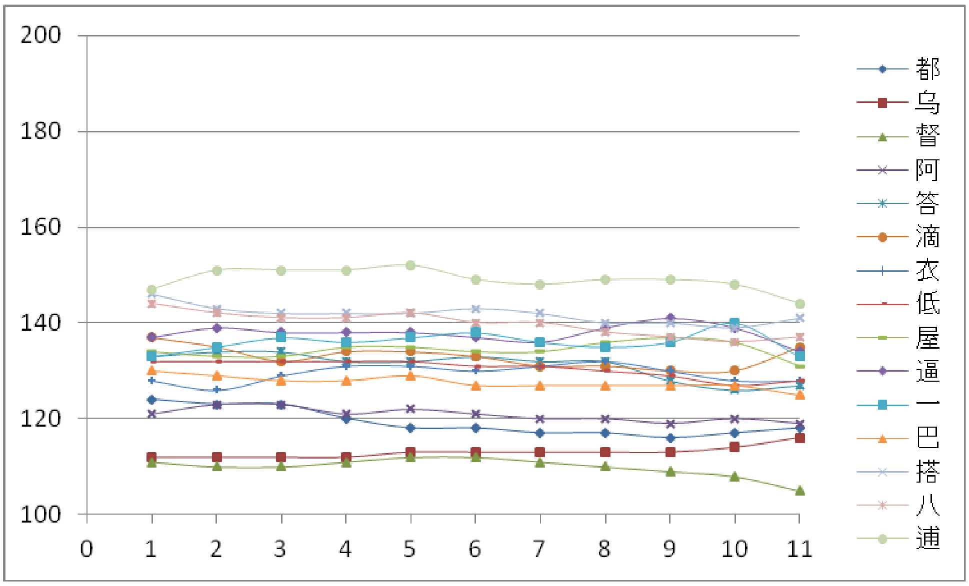


Figure 46. The F0 contour of 15 *Yinping* syllables pronounced by MB

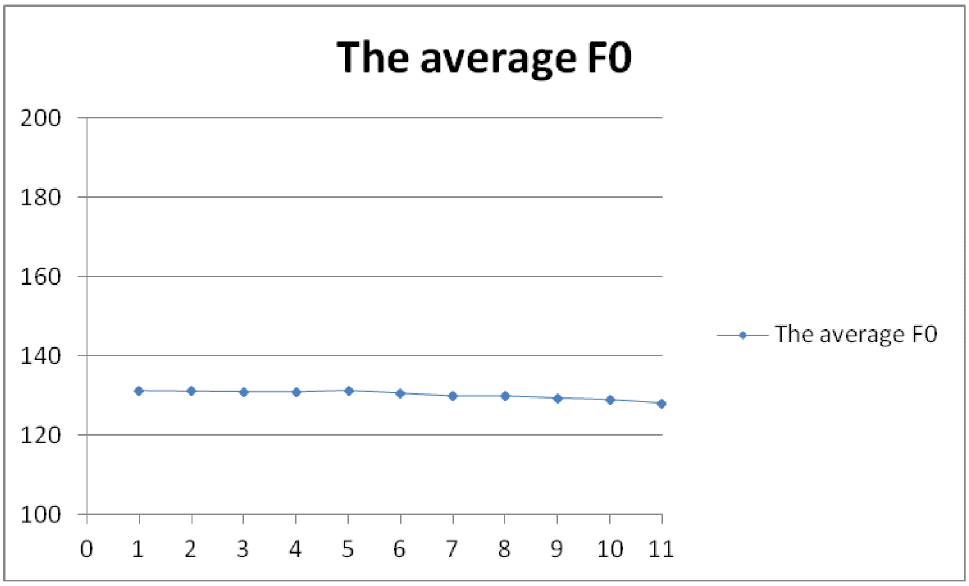


Figure 47. The average F0 contour of 15 *Yinping* syllables pronounced by MB

Based on Chart 1, the following conclusions can be made. (1) The F0 range of MB is 105-150Hz, that is, about 45Hz. In other words, the fundamental frequencies of every monosyllable are relatively concentrated. (2) Overall, the main parts of the F0 contours of MB are quite straight and steady.

From Chart 2, we can see that the mean F0 contour of MB is very straight and steady, with a value of about 130Hz. However, in this case, mean F0 contour represents the average of very distant values and its significance is therefore open to doubt.

5.1.2.2. General comments

From the tables and figures presented above, we can see that 80% of the Hungarian subjects pronounce *Yinping* syllables with an appropriately flat contour, especially M1, M4 and M6, who can pronounce them perfectly. However, 20% of the subjects produce deviant contours, most often as a result of a loose start or ending.

In order to better compare the 10 Hungarian subjects with the 2 native speakers in general terms, we made a chart for the average F0 contours of *Yinping* syllables pronounced by 12 subjects.

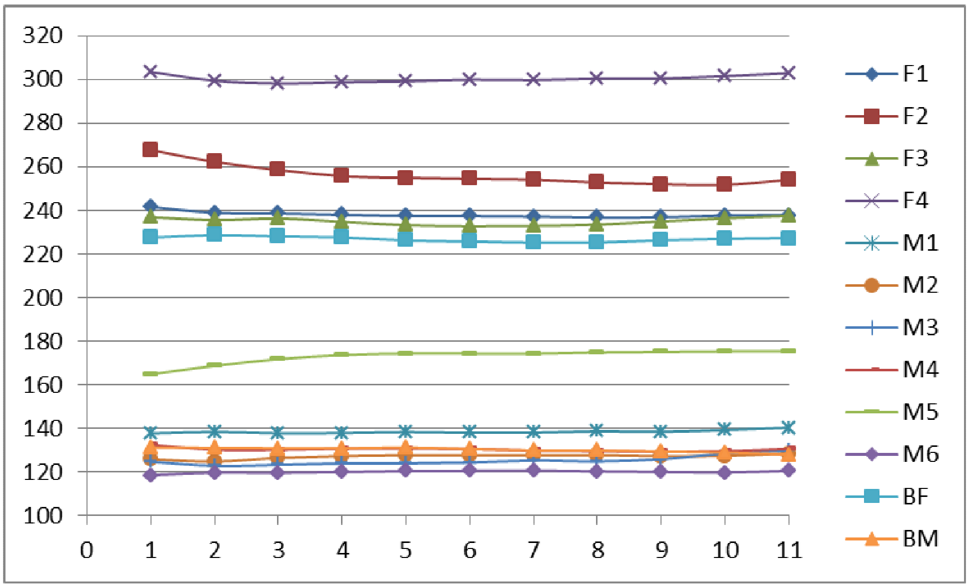


Figure 48. The average F0 contours of *Yinping* syllables pronounced by 12 subjects

On the basis of the average F0 contours, the following conclusions can be made:

(1) 80% of the Hungarian university students have a straight and steady average F0 contour when pronounce *Yinping* syllables;

(2) 20% of the Hungarian subjects do not pronounce *Yinping* syllables correctly; in other words, their average F0 contours are not straight and steady;

(3) The values of the average F0 contour of the Chinese subjects FB and MB are 230Hz, 130Hz, respectively. F1's, F2's and F3's are close to FB's; M1's, M2's, M3's, M4's, and M6's are close to MB's. While F1's is much higher than FB's, and M5's is much higher than MB's. In other words, 20% of Hungarian university learners produce considerably higher F0 contours when pronouncing *Yinping* syllables.

(4) The F0 values produced by females are considerably higher than males': about twice as high;

(5) The average F0 contours of males are flatter than females', which suggests that males have less problems in producing appropriate *Yinping* syllables than females.

5.1.3. Intrinsic vowel fundamental frequency

The intrinsic vowel fundamental frequency, usually abbreviated to IF0, refers to the fact that the F0 of a high vowel such as "i", is normally higher in terms of pitch than the F0 of a low one such as "a" when the two vowels are pronounced under the same conditions, for example when they are preceded by the same type of consonant, and with the same tone. (Lehiste, 1970)[174]. Zhu Xiaonong (2005) believes that the IF0 in a high area of the individual's own frequency domain is more distinct than that in a low area, and that IF0 appears more obviously in female than in male speakers[175].

For the purposes of defining the relationship between the height of the F0 contours and vowels, we classified all *Yinping* tones produced by each subject when pronouncing syllables with /i/, /u/ and /a/, and average the F0 values of each category. The following table shows the relevant details and the average F0 of all Hungarian female subjects and Hungarian male subjects (labeled F and M, respectively).

[174] Lehiste, Ilse 1970: 68.

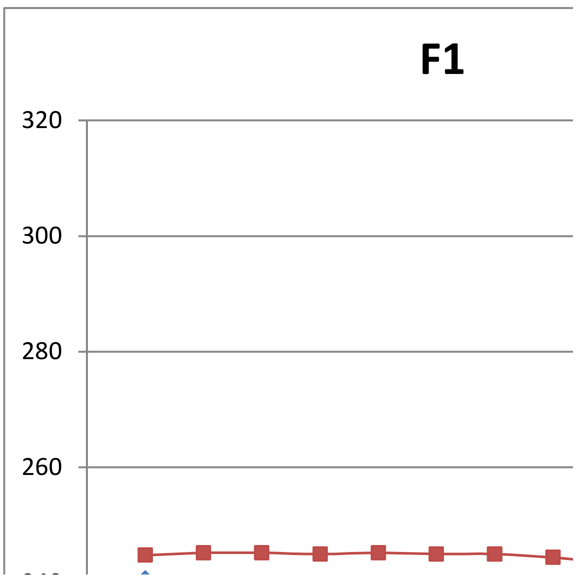
[175] Zhu 2005: 79.

Table 44: *Yinping F0 of each Hungarian female subject and Hungarian male subject, with averages for all female and all male subjects.*

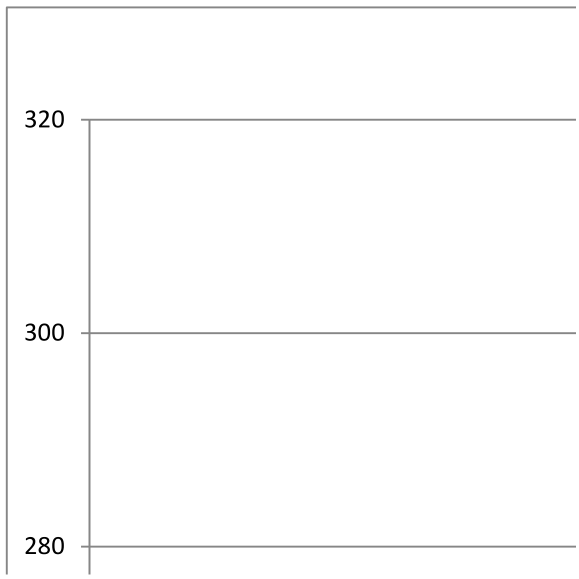
		1	2	3	4	5	6	7	8	9	10	11	ms
F1	/i/	240.8	238.4	236.6	235.6	235	235.4	234.8	234.6	234.8	235.2	233.2	410
	/u/	244.8	245.2	245.2	245	245.2	245	245	244.4	243.6	244.4	243.8	420
	/a/	239	233	233.8	233.2	232	231.6	231.6	231.2	232	233.2	236.4	416
F2	/i/	265.4	260.4	256.4	253.8	253.2	252.6	251.8	250.2	249.6	248.8	249.6	529
	/u/	268	265.2	261.2	258.2	258	258.2	258.8	256.8	255.2	255.4	259.4	499
	/a/	269.4	261	257.8	255	253.4	252.4	251.6	251.2	251	250.8	253	447
F3	/i/	237.2	235	235.4	234.2	232.6	231.6	231.4	232.2	233.4	234	234.6	405
	/u/	246.2	248.2	249.2	247.8	245.6	244.2	244.2	245.2	247.2	249.4	250	344
	/a/	227.2	223.4	224.2	222.2	221.4	222.4	223	223	224.2	225.8	227.8	395
F4	/i/	302.4	299.2	298	299.2	300.2	301.8	301.6	302.8	303.2	302.8	303.8	523
	/u/	314.6	314	312.6	310.2	308.6	309.4	309	309.2	308.4	310.2	308.8	535
	/a/	293.4	285	283.8	286.8	288.6	288.4	288.4	289	289.6	291.6	296	486
M1	/i/	140	142	141.6	142.4	142.8	141.6	141.2	142.4	142	142.8	142.4	289
	/u/	138.8	139	137.6	137.2	138.2	139.2	139.6	139.8	139	139.2	141	277
	/a/	134.6	134.4	134.4	134.4	134.2	133.4	133.6	134	134.6	136	137.8	283
M2	/i/	123.8	122.4	124.2	125.8	126.6	126.8	126.6	126.2	125.8	124.8	125.4	342
	/u/	128.6	127.2	129.2	130.2	130	129.4	129.2	129.2	128	128.8	129.6	362
	/a/	125.4	125	126.2	126.2	126.6	127.2	127.2	128	128.2	128.8	130.8	340
M3	/i/	127	125.6	125.8	126.6	125.8	125.4	126.2	125.2	126.6	129.2	131.4	283
	/u/	125.4	123.8	125	125.6	125.8	126.8	127.6	127.4	128.2	131	132	282
	/a/	122	119.4	119.4	119.8	120.6	121.2	122.2	122.2	123	125.6	126.2	277
M4	/i/	132	131	131.6	133.4	134.2	134.2	133	133	132.4	132.8	135	407
	/u/	135	133.8	133.4	133.6	133.6	133.6	132.8	132.2	131.8	131.2	132.2	336
	/a/	130.8	126	125.6	125.2	124.6	124	123.8	123.8	123.8	124.8	124.8	315
M5	/i/	168.4	173.2	175.4	176.2	177	177.4	177	177.4	177.4	178.2	178.4	562
	/u/	161	167.2	173.2	177.4	178.2	177.2	177.8	178.4	179	178.8	178	600
	/a/	165.4	165.8	166.8	167.4	168	168.4	168	168.8	169	169.2	170	629
M6	/i/	117.4	119	120	121	122	122	122.4	122.2	121.8	120.8	121.6	368
	/u/	120.6	122.8	123.6	124.6	125	125.2	124.8	124.2	123.8	123	124.8	323
	/a/	118	117.4	115.4	114.4	114.4	114.6	115	114.6	114.6	115.6	115.4	344
F	/i/	261.5	258.3	256.6	255.7	255.3	255.4	254.9	255	255.3	255.2	255.3	468
	/u/	246.2	248.2	249.2	247.8	245.6	244.2	244.2	245.2	247.2	249.4	250	447
	/a/	227.2	223.4	224.2	222.2	221.4	222.4	223	223	224.2	225.8	227.8	437
M	/i/	134.8	135.5	136.4	137.6	138.1	137.9	137.7	137.7	137.7	138.1	139	402

	/u/	134.9	135.6	137	138.1	138.5	138.6	138.6	138.5	138.3	138.7	139.6	380
	/a/	132.7	131.3	131.3	131.2	131.4	131.5	131.6	131.9	132.2	133.3	134.2	380
FB	/i/	228.4	229.6	230.2	229.2	228.4	228.6	227.2	227.6	227.4	226.6	223.6	447
	/u/	230.2	231.8	231.2	230.4	228	228.2	228	227.6	230.8	231.8	232.6	419
	/a/	223.8	224.6	222.8	223	222.2	220.4	220.4	220.6	220.6	222.4	225.6	398
MB	/i/	133.4	133.4	133.6	134.2	134.4	133.8	133	133.4	133.2	132.8	131.6	518
	/u/	125.6	125.8	125.8	125.8	126	125.2	124.6	125	124.8	124.6	122.8	462
	/a/	134.8	134.2	133.6	132.8	133.4	132.8	132.2	131.4	130.2	129.6	129.8	458

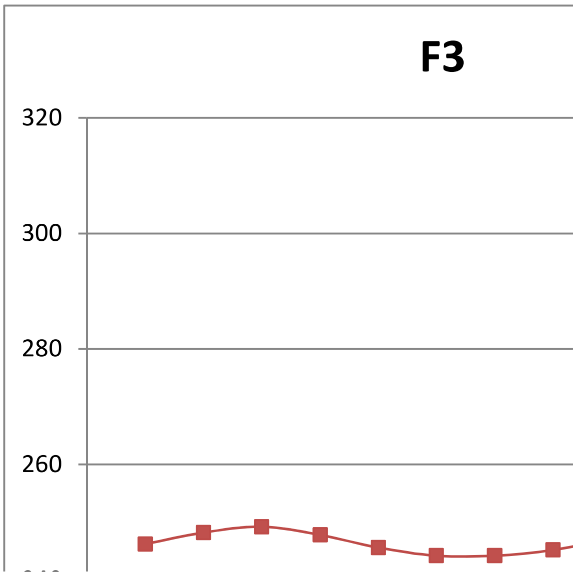
Based on the above table, we made charts of the F0 contours of /i/, /u/ and /a/ for each subject. In each chart the vertical (Y) axis shows the value of fundamental frequency, with Hz as the basic unit of measurement. The horizontal (X) axis presents the 11 measuring points. For the purposes of better comparison, the vertical axis covers a fixed value range based on the maximum and minimum frequencies recorded by female and male subjects respectively. For female subjects, the range is from 200Hz to 320Hz; for male subjects, from 80Hz to 200Hz.



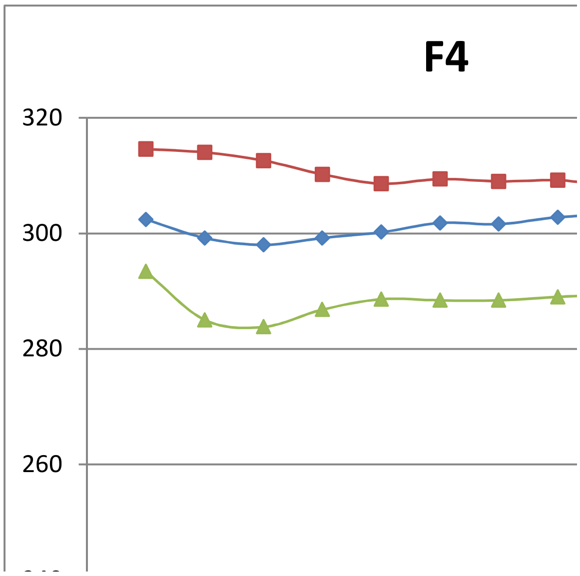
F1



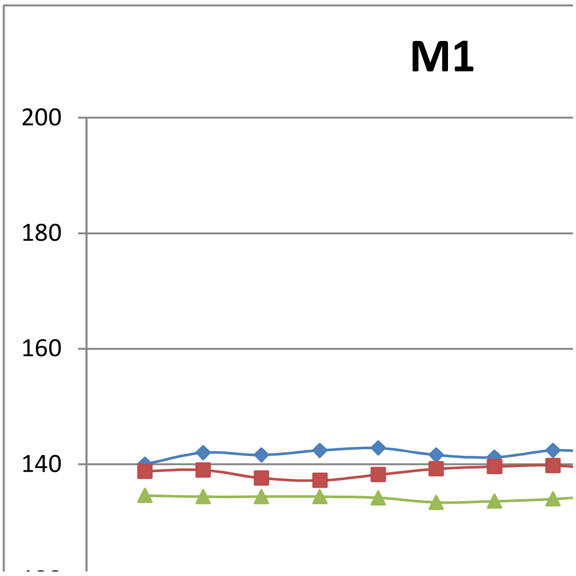
F2


F3
320
300
280
260

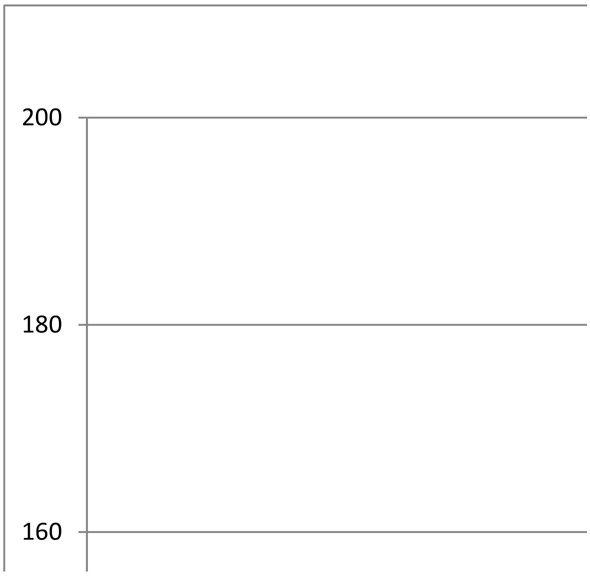
F3


F4
320
300
280
260

F4


M1
200
180
160
140

M1


200
180
160

M2

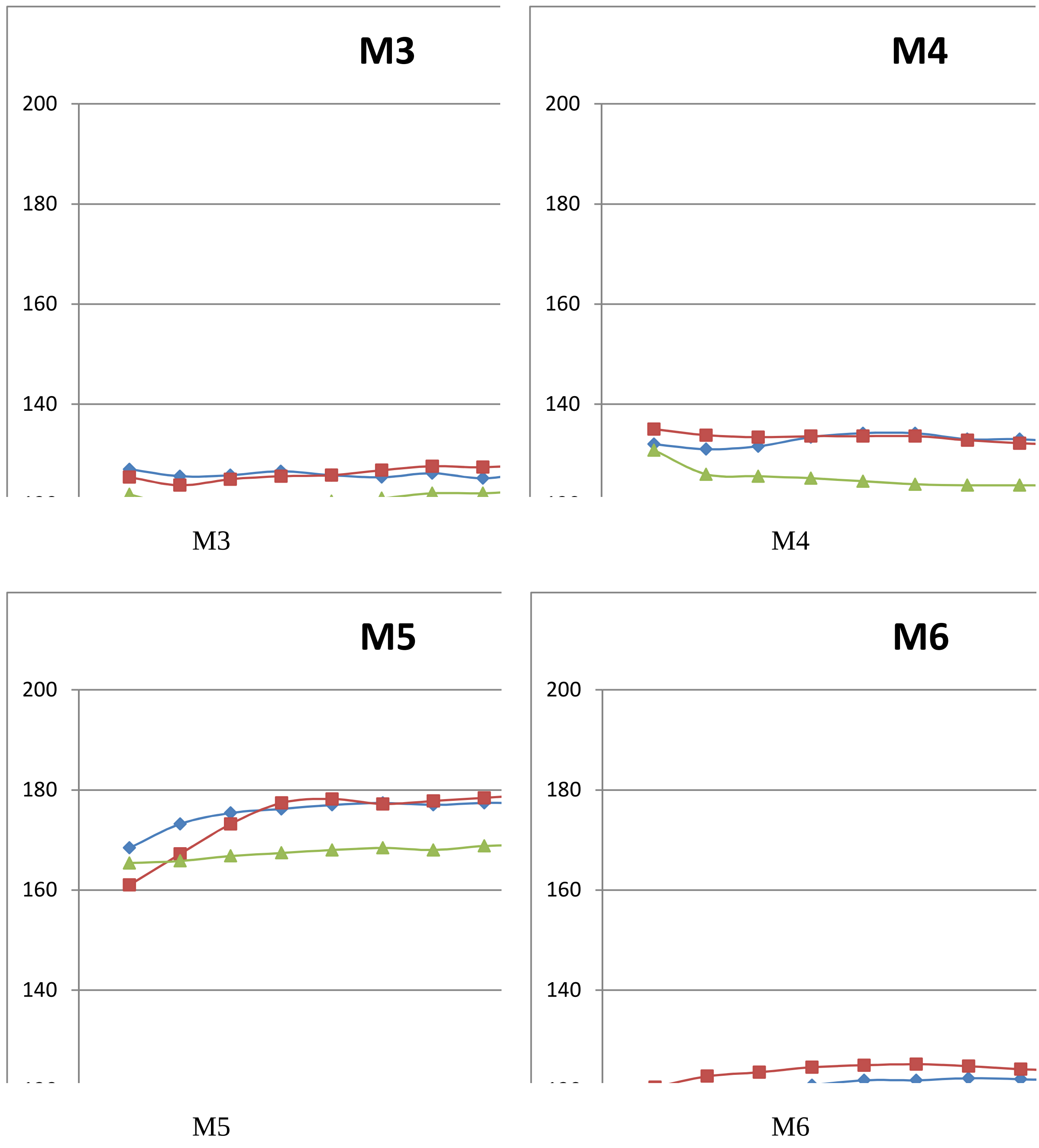


Figure 49. F0 *Yinping* contours of /i/, /u/ and /a/ by each Hungarian subject

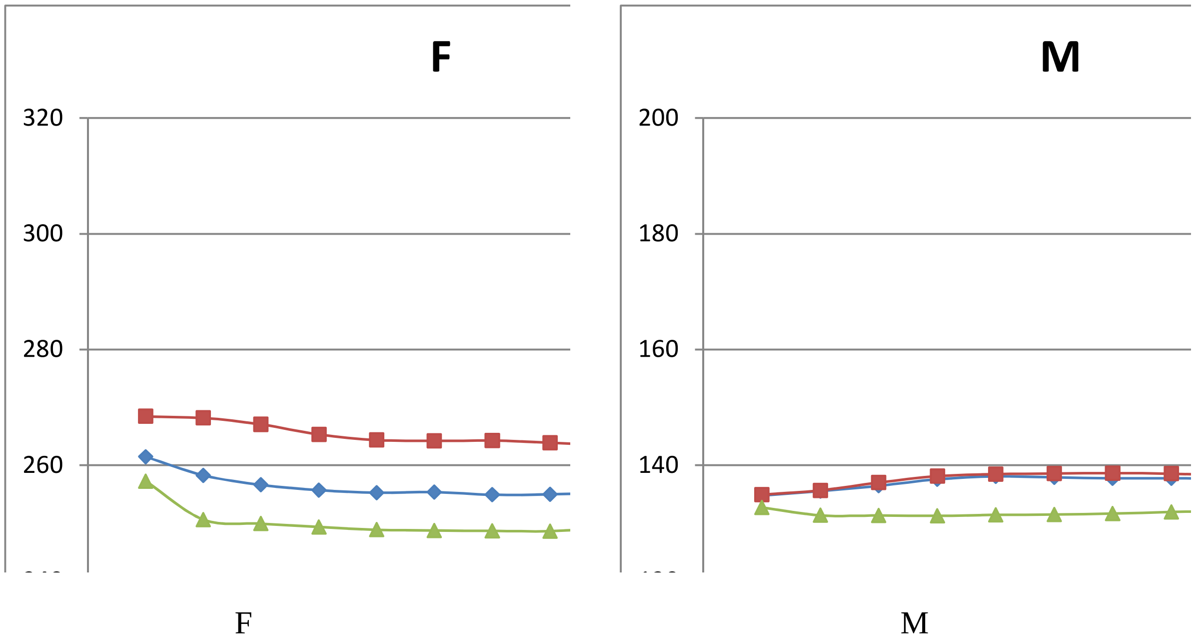


Figure 50. Average *Yinping* F0 contours of /i/, /u/ and /a/ for all Hungarian male (M) and female (F) subjects

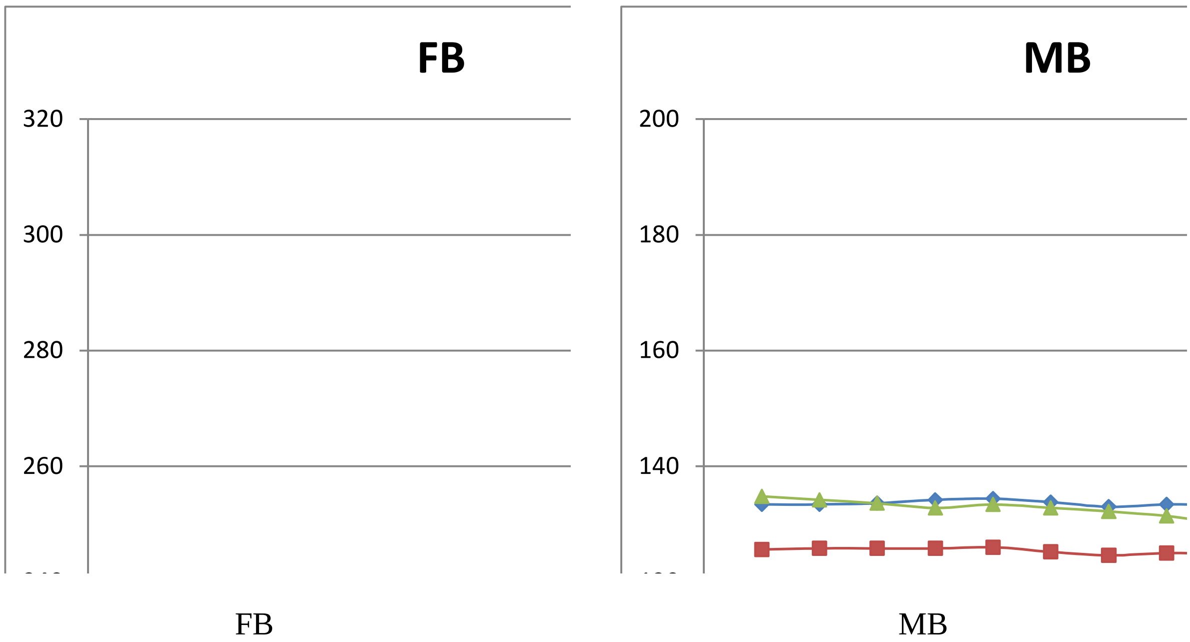


Figure 51. Average *Yinping* F0 contours of /i/, /u/ and /a/ for all both Chinese subjects

From the charts presented above, we can see: (1) The manifestations of the IF0 in *Yinping* syllables produced by FB and MB are not very obvious, since their average F0 contours of /i/,

/u/ and /a/ are quite concentrated. (2) The *Yinping* syllables produced by all the female subjects contain distinctly different F0 levels in terms of /i/, /u/ and /a/, with the order of /u/＞/i/＞/a/, except for F2, whose F0 contours of /i/ and /a/ are almost overlapping, and both are considerably close to the F0 contour of /u/. (3) The IF0 is reflected not very clearly in *Yinping* syllables produced by M4, M5, and M6, since their average F0 contours of /i/, /u/ and /a/ are considerably concentrated. (4) The performances of the IF0 in M1, M2, M3, and F2 as well, are by no means obvious, since their average F0 contours of /i/, /u/ and /a/ are highly concentrated. (5) The average F0 contours of /i/, and /u/ produced by all Hungarian female subjects are almost overlapping and relatively higher that of /a/ produced by them, which indicates the considerably obvious performances of the IF0. However, the average F0 contours of /i/, /u/ and /a/ produced by all Hungarian male subjects are quite concentrated, which means that these subjects do not manifest obvious performances of the IF0.

All in all, the manifestations of the IF0 in *Yinping* tone produced by both native speakers are not very significant, and the same can be said of *Yinping* tone produced by all Hungarian male subjects. However, only one of four female Hungarian subjects produced the *Yinping* tone without showing the obvious manifestations of the IF0, which suggests that 75% of Hungarian female learners are likely to produce *Yinping* tone accompanied by the quite obvious existence of the IF0. In other words, the IF0 appears more obviously in these female than in tese male subjects when they are producing *Yinping* tone, which corresponds to Zhu Xiaonong's statement about native speakers of Shanghai dialect.

5.1.4. Range

Range, the area of level or pitch at which sounds are produced, is an indispensable factor in analysis of learners' acquisition of tones. Chao Yuenren (1980) believes that it is in fact the range, and not the tone pattern, that causes the problems when learners try to produce tones[176]. Zhao Jinming (1988) agrees with Chao on this point, as he also found that the most common error concerns range[177]. In other words, learners may produce tone patterns which are appropriate in themselves, but which occur at an inappropriately high or low pitch-range in

[176] Chao 1983.

[177] Zhao 1988: 10-14.

comparison with some or all of the other three tone patterns. Shen Xiaonan (1989) also showed that for American learners, errors tended to affect range rather than tone patterns. He further points out that errors of range can consist of the entire sound area of a tone occurring in too high or too low a pitch-range[178]. We will discuss this issue within the whole tone system later in this chapter.

For *Yinping*, the range also indicates how straight and steady the contour is. At this point, the range means the difference between the maximal and minimum value of F0. Therefore, the smaller the range, the straighter and steadier the contour of *Yinping*.

Table 45: *The average range of 15 Yinping syllables pronounced by 12 subjects*

Subjects	Maximal value of average F0	Minimum value of average F0	Range
F1	241.5	236.7	4.8
F2	267.6	251.7	15.9
F3	226	220	4.8
F4	303.5	298.1	5.4
M1	140.4	137.8	2.6
M2	128.6	124.9	3.7
M3	129.9	122.9	7
M4	132.6	129.3	3.3
M5	175.5	164.9	10.6
M6	120.7	118.7	2
FB	228.7	225.2	3.5
MB	131.3	128.1	3.2

178 Shen 1989.

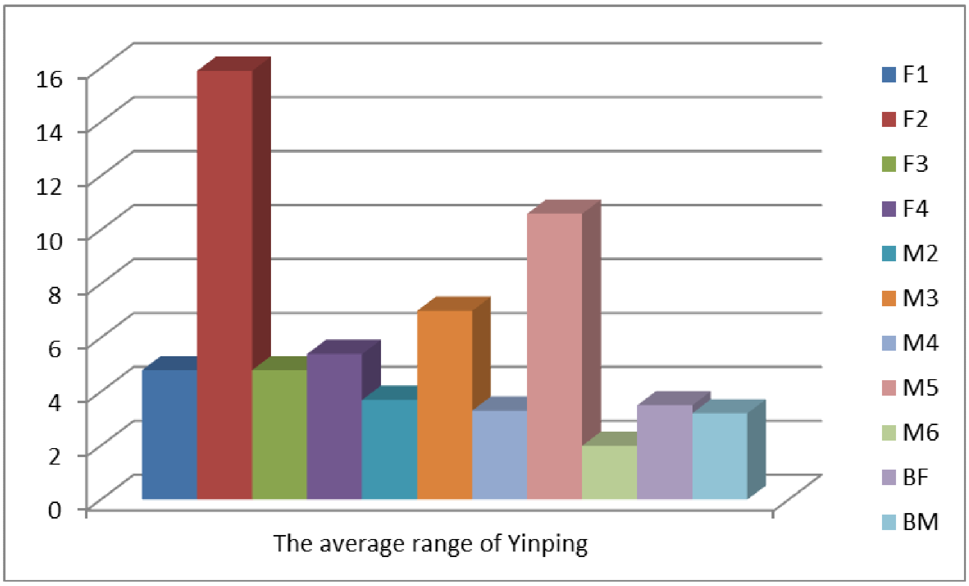


Figure 52. The average range of 15 *Yinping* syllables pronounced by 12 subjects

From the table and the figure presented above, we can see: (1) the average ranges of FB and MB are 3.5Hz and 3.2Hz; (2) the average ranges of F1, F3, F4, M1, M2, M4, M6 are very close to FB's and MB's, and the average ranges of M1 and M6 are 2.6 and 2, which are even smaller than FB's and MB's. In a word, the straight and steady level of *Yinping* pronounced by these 6 subjects are very high; (3) the average ranges of M3 are almost twice those of FB and MB; however, informal reports from their Chinese teachers indicate that the latter do not find this "deviance" at all disturbing, or even perceptible; (4) the average ranges of F2 and M5 are considerably (and, this case perceptibly and disturbingly) higher than FB's and MB's, which indicates that these two learners do not produce acceptably straight and steady levels in *Yinping*. All in all, 80% of Hungarian university learners do not appear to experience problems in pronouncing *Yinping*, but 20% of them do find it difficult.

5.1.4. Summary

In this section, the recordings of 12 subjects are examined in terms of the duration, the fundamental frequency (F0) and the tone range of their *Yinping* syllables. On the basis of measurement, extraction, analysis and comparison, the following conclusions can be drawn.

(1) In terms of the duration, 60% of Hungarian university learners – all the male subjects in the sample - differ considerably from the "model" native speakers, most often in

producing much shorter durations. In other words, the male learners in this sample are more likely have problems with duration. What is more, 50% of Hungarian learners produce deviant orders of durations of /i/, /u/ and /a/ which indicates that the inappropriately produced vowels probably cause problems in duration.

(2) With regard to fundamental frequency (F0), the F0 contour of the majority of these Hungarian university learners' are quite straight and steady, only 20% of them have difficulty producing appropriately flat contours. As to the pitch value of F0, 20% of them (one female and one male in the sample) have much higher values, compared to female and male "model" native speakers, respectively. Apart from these two, it emerges that the male Hungarian subjects are closely grouped in terms of pitch around the native male, while all the female Hungarian subjects were higher in pitch than the female native. Concerning the intrinsic vowel fundamental frequency, it turned out that the effect of the IF0 is more obvious in Hungarian female learners than in male learners when they are producing the *Yinping* tone.

(3) Summarising the findings about range of pitch, we can say that most of these Hungarian university learners have a fairly appropriate range in *Yinping*, that is, their ranges are quite close to those of native speakers. However, 20% of them (one female and one male in the sample) can be considered deviant in the sense that they produce much wider ranges.

5.2. *Yangping*

The second tone, called *Yángpíng*, is at mid-to-high level. The value of tone in the Chao system is 35, which means that it begins at mid pitch 3, and then rises to the highest pitch 5. In *Xiandai Hanyu* Feng Zhichun (2008) defines *Yángpíng* as a high rising tone or mid rising tone with its voice from middle to high[179].

The linguistic goal of the rising tone lies in the peak point of the end tail. In general, there is a

[179] Feng 2008: 100.

falling start in the fundamental frequency (F0) contour of the rising tone, so the whole tone shape is concave, but it does not count as a dipping tone like *Shǎngshēng*. The literature suggests that difference between the rising tone and the dipping tone is the inflexion point. It should be considered as a rising tone if the inflexion point lies at a point around 20% along the total length and it should be considered as a dipping tone if the inflexion lies at a point around 40% of the total length[180].

The main aim of this section is to present the voice samples of the second tone, Yangping, which were recorded by 12 subjects, including the measurement, extraction, analysis and comparison. The recordings are examined in terms of the duration, the fundamental frequency (F0), the tone range and the tone value of the monosyllabic tone produced by the recorded subjects, thus creating a detailed profile of the *Yangping* pronunciation of this sample of Hungarian university students (six male, labeled M1, M2 etc., four female, labeled F1, F2 etc.), juxtaposed with the pronunciation of one male and one female native speaker of Chinese (labeled MB and FB respectively). The actual methodology used in the measurement of duration, F0 and tone range was detailed in 4.3 above; note that the initial or "onset" of each syllable was discarded; the 11 measuring points all fall within the nucleus or rime. Thus the word "start" used in the comments about the subjects' performance refers to the beginning of the nucleus, and not to the beginning of the whole syllable.

The 14 *Yangping* syllables used in the recordings, as explained in 4.3 above, fit into the following matrix:

Table 46: *Matrix of Yangping syllables*

	á	í	ú
b	拔	鼻	
d	达	迪、敌、笛	读 、独、毒
Zero-initials		姨、宜、咦	吴、无

[180] Zhu 2010: 278.

5.2.1. Duration

The method used for determining the duration of the syllable was been presented above. The following table shows the durations of each *Yangping* syllable pronounced by 12 subjects, as well as average durations. The unit of duration is the millisecond (ms).

Table 47: *Duration of 12 different Yangping syllables pronounced by 12 different subjects.*

Syllables	F1	F2	F3	F4	M1	M2	M3	M4	M5	M6	FB	MB
姨 yí	535	420	297	365	283	468	241	387	775	360	455	538
读 dú	504	486	285	387	318	491	233	285	792	319	456	461
独 dú	358	460	227	340	298	336	216	306	600	366	406	515
拔 bá	434	364	299	347	217	334	216	349	307	245	370	466
鼻 bí	432	434	318	369	334	431	182	293	509	285	525	514
吴 wú	456	520	346	389	342	415	222	398	342	284	533	475
达 dá	397	419	239	359	318	382	237	351	501	301	413	495
宜 yí	523	454	325	375	285	434	198	374	568	289	422	608
咦 yí	488	433	232	448	331	490	225	384	501	323	466	568
迪 dí	481	455	302	412	332	531	258	320	478	315	457	519
敌 dí	432	527	260	278	349	362	222	263	580	375	465	498
无 wú	450	502	382	401	391	517	293	312	638	342	450	537
笛 dí	437	469	204	308	333	488	263	309	546	361	457	657
毒 dú	398	529	295	363	246	441	223	293	491	341	502	441
Average	452	462	285	367	313	437	231	330	545	322	451	521

On the basis of the averages displayed in the bottom row of the above table, we made a chart for further comparison.

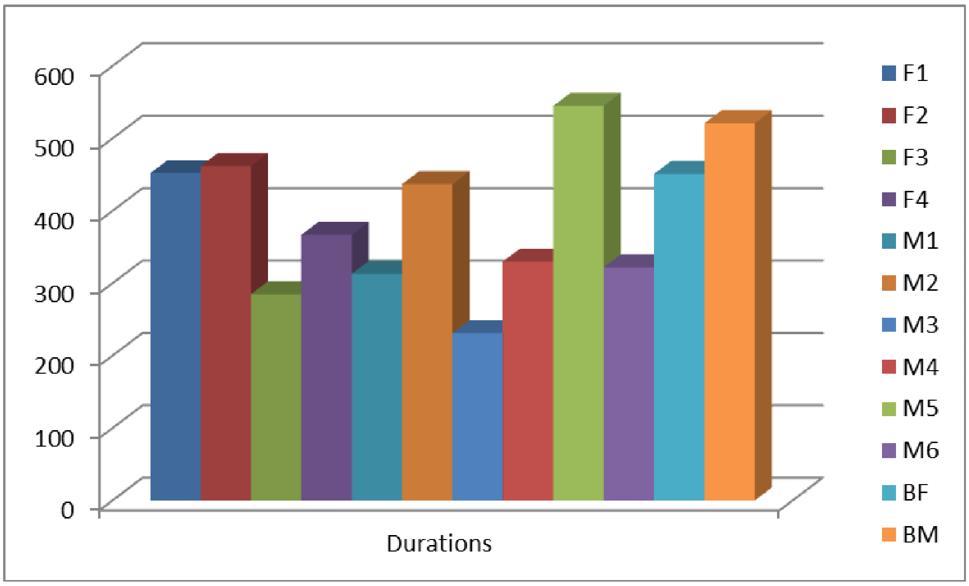


Figure 53. Average duration of *Yangping* syllables pronounced by 12 different subjects

The chart clearly shows that the average duration of M5 is the longest, while the average duration of M3 is the shortest, at less than half of M5's. The average durations of FB and MB are 456ms and 521ms, respectively. The average durations of F1, F2, M2, and M5 are close to FB's and MB's, the average durations of F4, M1, M4, and M6 are considerably shorter than FB's and MB's, and the average duration of F3 and M3 are much short than FB's and MB's, especially M3's. In other words, 60% of Hungarian university learners – two female and two male subjects in the sample - differ considerably from the "model" native speakers in terms of the duration of Yangping syllables, all in producing much shorter durations.

As we have seen, producing a tone in too long or too short a manner is a common causes of problems in duration, but it is probably not the only reason. The inappropriate pronounciation the vowel bearing a tone also can lead to duration problems. Generally speaking, the duration of a given vowel is associated with its height, and each vowel has an intrinsic vowel duration (referred to as IVD)[181]. Catford (1977, 45) believes that the duration of a high vowel is longer than that of a low one; his statement is described as IVD "i>a"[182].

The following table shows the average durations (m) of *Yangping* produced by 12 subjects, in

[181] Zhu 2005: 72.

[182] Catford 1977: 45.

terms of different vowels /i/, /u/ and /a/ which are bearing the *Yangping* tone, as well as the average durations of all female subjects (F) and male subjects (M). As before, the unit of duration is the millisecond (ms), and “n” refers to the number of sample syllables. The last column in the table shows the subject’s “duration order” of /i/, /u/ and /a/, with the longest vowel first and the shortest vowel last. For example, F1’s average durations of *Yangping* tone which are contained by vowel /i/, /u/ and /a/ are 475ms, 433ms and 416ms, respectively. The “duration order” of /i/, /u/ and /a/ is /i/>/u/>/a/.

Table 48: *Average durations (m) of vowels /i/, /u/ and /a/ with Yangping tone produced by 12 subjects*

	/i/		/u/		/a/		the order
	m	n	m	n	m	n	
F1	475	7	433	5	416	2	/i/>/u/>/a/
F2	456	7	499	5	392	2	/u/>/i/>/a/
F3	277	7	307	5	269	2	/u/>/i/>/a/
F4	365	7	376	5	353	2	/u/>/i/>/a/
M1	321	7	319	5	268	2	/i/>/u/>/a/
M2	458	7	440	5	358	2	/i/>/u/>/a/
M3	227	7	237	5	227	2	/u/>/i/=/a/
M4	333	7	319	5	350	2	/a/>/i/>/u/
M5	565	7	573	5	404	2	/u/>/i/>/a/
M6	330	7	330	5	273	2	/u/=/i/>/a/
F	393	7	404	5	357	2	/u/>/i/>/a/
M	372	7	370	5	313	2	/i/>/u/>/a/
FB	464	7	469	5	392	2	/u/>/i/>/a/
MB	557	7	486	5	481	2	/i/>/u/>/a/

From the above table, we can see that: (1) FB’s average durations of /i/, /u/ and /a/ are 464ms, 469ms and 392ms, and MB’s durations of these three vowels are 557ms, 486ms and 481ms, respectively. In other words, the orders of both FB’s and MB’s durations of these three vowels are /i/>/a/ (for FB, is /u/>/i/>/a/; for MB, is /i/>/u/>/a/), which corresponds to the “i>a” principle. (2) For the majority of the subjects, F1, F2, F3, F4, M1, M2 M5 and M6, their “duration order” of /i/, /u/ and /a/ are the same as “model” native speakers’, which is /i/>/a/. Since both /i/ and /u/ are high vowels, both the order of /u/>/i/>/a/ or /i/>/u/>/a/ make sense. (4) 20% of Hungarian subjects produce deviant “duration orders” of /i/, /u/ and /a/. The average durations produced by M4 is in the order of /u/>/i/=/a/, and M5’s in /a/>/i/>/u/. (4) The average durations of /i/, /u/ and /a/ produced by all Hungarian female subjects are sorted by /u/>/i/>/a/, which is identical to “model” native speakers’, and the average duration order of /i/, /u/ and /a/ produced by all Hungarian male subjects is /i/>/u/>/a/. Both F’s and M’s duration orders of vowels are consistent with the “i>a” principle, that is, /i/>/a/. All in all, the deviant duration orders of /i/, /u/ and /a/ indicate that it is the inappropriately produced vowels that are likely to cause problems in duration.

It is worth mentioning that the average durations of /i/, /u/ and /a/ produced by all Hungarian female subjects (F) and male subjects (M) are 385ms [(393＋404＋357) ÷ 3], 352ms [(372＋370＋313) ÷ 3], respectively. In other words, while the sample is too small to allow generalized statements, there is reason to believe that a connection may exist between duration and gender. This conclusion is in contrast with Zhu Xiaonong’s view: the latter believes there is no connection between duration and gender[183].

However, as we have seen duration is a relative concept. It is hard to define a standard or universally “desirable” value of duration, because it depends so strongly on the individual’s general rate of speech, not to mention the nature of each specific utterance and the conditions under which it is made. Discussing the duration of specific syllables within the parameters of an individual’s own tone system probably makes better sense. This issue will be discussed later in this chapter.

5.2.2. Fundamental frequency

In this section, firstly, each subject’s F0 values for every *Yangping* monosyllable, together with the means for all the syllables, are presented in a combined table. Secondly, each subject’s F0 contours for all the *Yangping* monosyllables are plotted in one chart for further

[183] Zhu 2005: 77.

analysis and comparison. In each chart the vertical (Y) axis shows the value of fundamental frequency, with Hz as the basic unit of measurement. The horizontal (X) axis presents the 11 measuring points. For better comparison purposes, the vertical axis covers a fixed value range based on the maximum and minimum frequencies recorded by female and male subjects respectively. For female subjects, two ranges are set based on the quite spread F0 contours, from 130Hz to 300Hz and from 170Hz to 420Hz, respectively; for male subjects, from 90Hz to 210Hz. Thirdly, the average contour of all the subject's *Yangping* syllables is plotted in another chart. (For details of the methods used for measuring fundamental frequency (F0), see 3.2.1)

5.2.2.1. Individual subjects' F0 data

The F0 data table and charts of every subject will now be provided.

(1) F1

The following table shows the F0 data for subject F1, including 14 *Yangping* monosyllables. Each monosyllable has 11 F0 values which correspond to the 11 measuring points. The means of the measurements made at each point are presented in the final row of the table.

Table 49: *F1's F0 values for 14 Yangping syllables at 11 measuring points*

Syllables	1	2	3	4	5	6	7	8	9	10	11
姨 yí	197	175	176	178	176	175	179	187	200	219	223
读 dú	196	186	187	188	189	195	201	208	216	229	232
独 dú	185	177	173	173	173	175	180	190	194	198	201
拔 bá	183	171	167	167	169	170	172	177	189	204	217
鼻 bí	185	175	167	164	163	166	169	175	183	192	210
吴 wú	182	180	179	179	178	180	182	188	198	205	215
达 dá	193	185	180	178	179	190	195	200	209	220	228
宜 yí	208	185	184	186	190	193	201	210	222	229	237
咦 yí	200	193	190	190	192	199	206	215	220	226	229
迪 dí	208	196	192	189	191	195	202	210	219	222	224
敌 dí	189	179	173	174	176	181	188	198	206	212	205
无 wú	197	189	184	183	180	181	183	191	201	209	213
笛 dí	193	183	178	177	180	187	196	206	217	230	233

毒 dú	196	189	186	188	192	197	205	211	219	228	236
Average	193.7	183.1	179.7	179.6	180.6	184.6	189.9	197.6	206.6	215.9	221.6

On the basis of the above table, we made the following two F0 contour charts of F1's *Yangping* pronunciation. The first chart shows the contour for every monosyllable, the second shows the mean of all the contours. The two charts are followed by brief comments on the data.

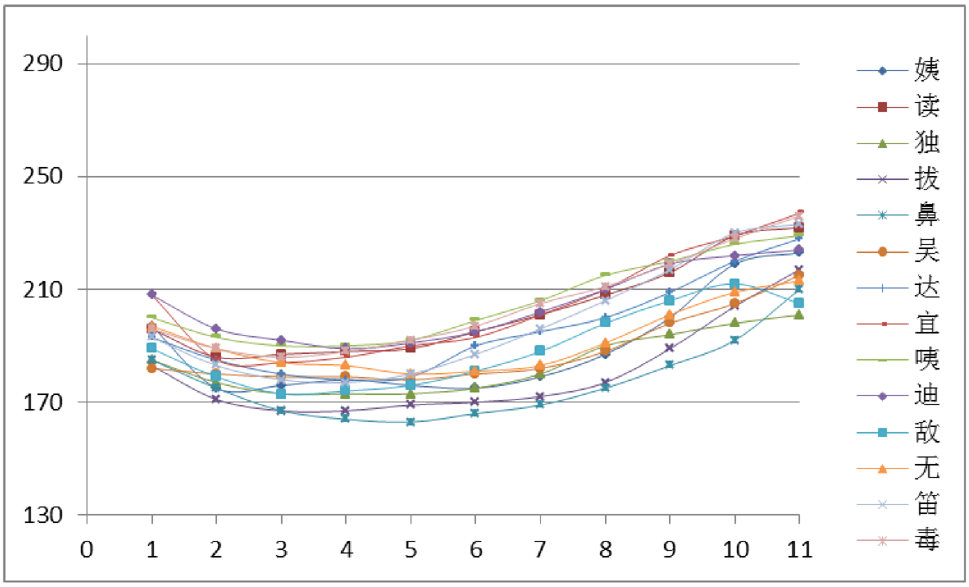


Figure 54. The F0 contour of 14 *Yangping* syllables pronounced by F1

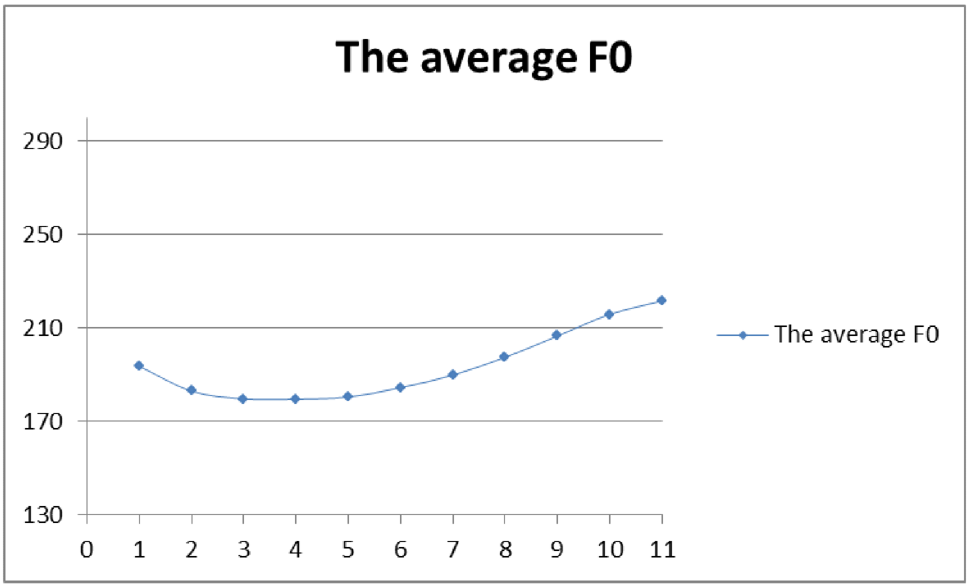


Figure 55. The average F0 contour of 14 *Yangping* syllables pronounced by F1

Based on Chart 1, the following conclusions can be made. (1) The F0 range of F1 is 160-240Hz, that is, about 80Hz. (2) The F0 contours of F1 are relatively concentrated. (3) Many F0 contours have a clear falling start. (4) Overall, the rising parts of most F0 contours are relatively gentle, but because of the falling start, the F0 contours of F1 turn out as fall-rise patterns. (5) Most F0 contours have an inflection point of rising at the first half of its contour.

From Chart 2, we can see that the rising part of the mean F0 contour of F1 is quite gentle and does not begin until the fifth measuring point, in other words, at a point about 45% through the whole F0 contour. Indeed the mean F0 contour of F1 actually presents a fall-rise pattern. The value of the mean F0 contour ranges between 175 and 225Hz.

(2) F2

Table 50: *F2's F0 values for 14 Yangping syllables at 11 measuring points*

Syllables	1	2	3	4	5	6	7	8	9	10	11
姨 yí	216	205	201	205	209	213	218	229	239	263	282
读 dú	242	222	208	203	201	203	207	214	224	244	267
独 dú	231	212	205	201	202	207	214	220	234	251	271
拔 bá	220	212	211	212	214	217	224	231	246	258	265
鼻 bí	240	214	201	197	198	204	213	220	231	243	257
吴 wú	207	206	201	201	202	206	210	217	232	261	291
达 dá	222	210	206	206	208	213	219	233	250	265	278
宜 yí	221	204	204	210	215	220	226	237	248	255	269
咦 yí	227	205	202	204	206	210	214	220	233	248	260
迪 dí	234	213	204	203	207	213	219	226	235	253	270
敌 dí	225	202	197	200	205	213	219	230	241	257	279
无 wú	211	210	213	210	211	216	224	237	252	268	292
笛 dí	237	209	201	202	205	211	218	226	243	266	291
毒 dú	231	212	207	208	212	219	230	243	259	284	299
Average	226	209.7	204.4	204.4	206.8	211.8	218.2	227.4	240.5	258.3	276.5

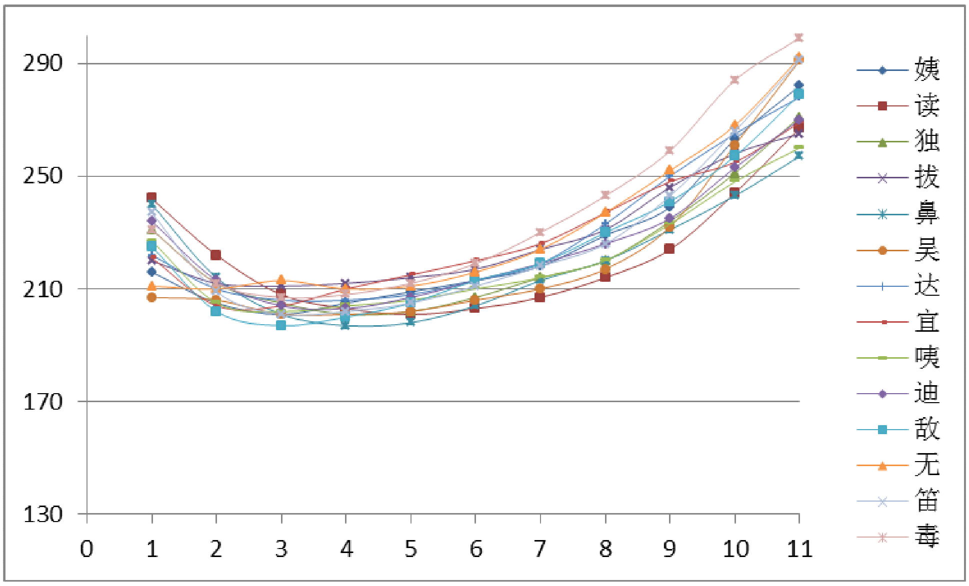


Figure 56. The F0 contour of 14 *Yangping* syllables pronounced by F2

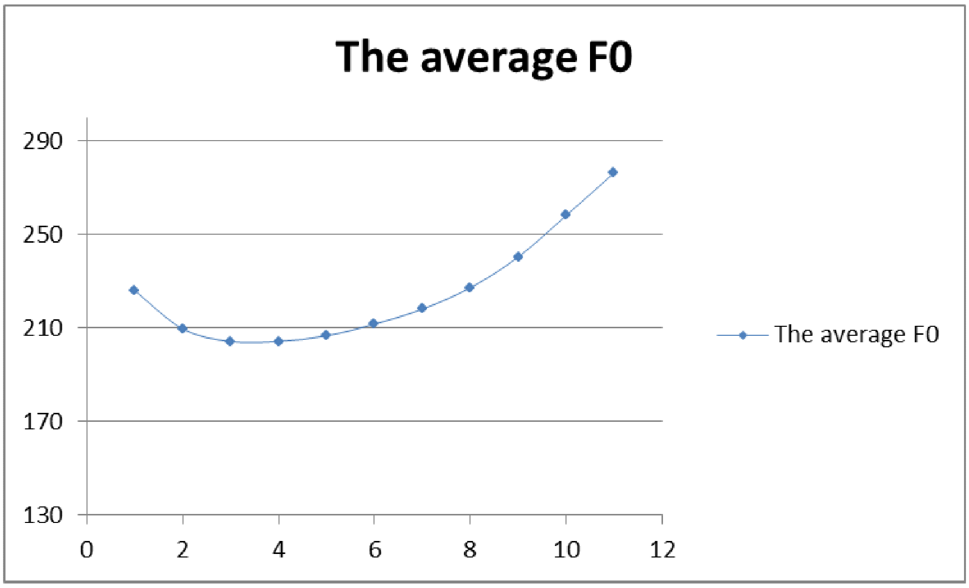


Figure 57. The average F0 contour of 14 *Yangping* syllables pronounced by F2

Based on Chart 1, the following conclusions can be made. (1) The F0 range of F2 is 190-300Hz, that is, about 110Hz. (2) The F0 contours of F2 are highly concentrated. (3) Several of the F0 contours have a very clear falling start. (4) Overall, the rising parts of most F0 contours are appropriately steep, but because of the falling start, the F0 contours of F2 turn out as fall-rise patterns. (5) Most F0 contours change direction and start rising after the first half of the contour.

From Chart 2, we can see that the rising part of the mean F0 contour of F2 is appropriately steep and begins at the fourth measuring point, in other words, at a point about 35% through the whole F0 contour. The mean F0 contour of F2 presents a fall-rise pattern. The value of the mean F0 contour ranges between 200 and 280Hz.

(3) F3

Table 51: *F3's F0 values for 14 Yangping syllables at 11 measuring points*

Syllables	1	2	3	4	5	6	7	8	9	10	11
姨 yí	205	206	201	198	197	195	205	223	255	301	337
读 dú	223	220	220	219	220	225	240	270	310	338	357
独 dú	201	194	194	194	198	208	228	260	302	336	362
拔 bá	225	222	220	220	218	215	214	212	214	218	223
鼻 bí	210	207	206	205	204	207	212	228	266	313	361
吴 wú	212	211	210	209	209	208	207	213	232	260	278
达 dá	231	218	213	213	217	219	222	225	230	232	236
宜 yí	240	246	249	250	249	248	248	248	249	250	248
咦 yí	227	208	205	208	216	224	233	247	272	306	341
迪 dí	202	204	209	214	216	220	233	250	266	285	300
敌 dí	209	205	207	208	213	218	226	249	268	298	330
无 wú	210	210	206	207	207	202	201	203	210	239	267
笛 dí	221	221	225	233	244	256	270	286	301	311	323
毒 dú	200	201	210	219	236	252	272	297	322	363	385
Average	215.4	212.4	212.5	214.1	217.4	221.2	229.4	243.6	264.1	289.3	310.6

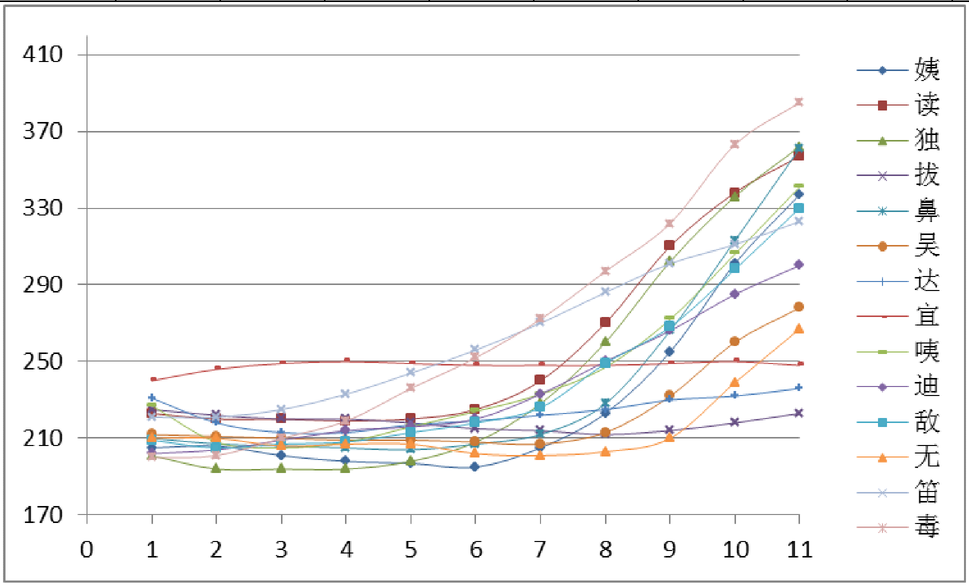


Figure 58. The F0 contour of 14 *Yangping* syllables pronounced by F3

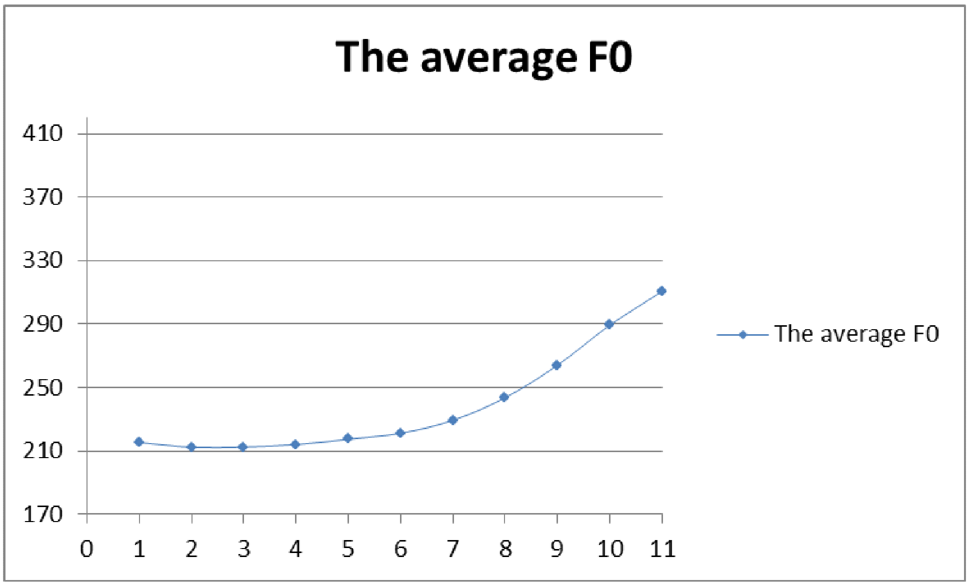


Figure 59. The average F0 contour of 14 *Yangping* syllables pronounced by F3

Based on Chart 1, the following conclusions can be made. (1) The F0 range of F3 is 190-390Hz, that is, about 200Hz. (2) The F0 contours of F3 are quite spread out. (3) Most F0 contours have a very smooth start. (4) Overall, the rising parts of most F0 contours are very steep. (5) However, some F0 contours appear as flat patterns, without an apparent rising trend, such as 拔 bá, 达 dá and 宜 yí. (6) Most F0 contours feature a rising tone which begins after the first half of the contour.

From Chart 2, we can see that the rising part of the mean F0 contour of F3 is quite steep and starts rising at the third measuring point, in other words, at a point about 25-30% through the whole F0 contour. The mean F0 contour of F3 appears as a rising pattern. The value of the mean F0 contour ranges between 210 and 320Hz.

(4) F4

Table 52: *F4's F0 values for 14 Yangping syllables at 11 measuring points*

Syllables	1	2	3	4	5	6	7	8	9	10	11
姨 yí	247	225	218	212	205	204	208	216	242	317	369
读 dú	228	201	172	173	188	190	204	226	258	306	350
独 dú	221	206	191	184	185	189	195	205	247	351	412
拔 bá	232	202	187	184	185	183	182	190	204	240	294
鼻 bí	251	205	174	176	181	188	196	209	242	323	377

吴 wú	230	228	217	207	196	194	195	218	257	326	372
达 dá	221	200	183	182	185	193	197	204	216	251	302
宜 yí	204	191	189	187	187	190	201	214	250	319	350
咦 yí	211	203	192	185	181	182	182	194	213	266	343
迪 dí	211	187	181	179	180	186	201	215	236	288	349
敌 dí	202	185	185	182	186	193	204	220	247	287	327
无 wú	219	202	181	175	175	178	182	196	226	291	349
笛 dí	212	184	180	179	181	186	193	216	261	306	342
毒 dú	210	194	185	183	184	190	205	226	259	330	376
Average	221.4	200.9	188.2	184.9	185.6	189	196.1	210.6	239.9	300.1	350.9

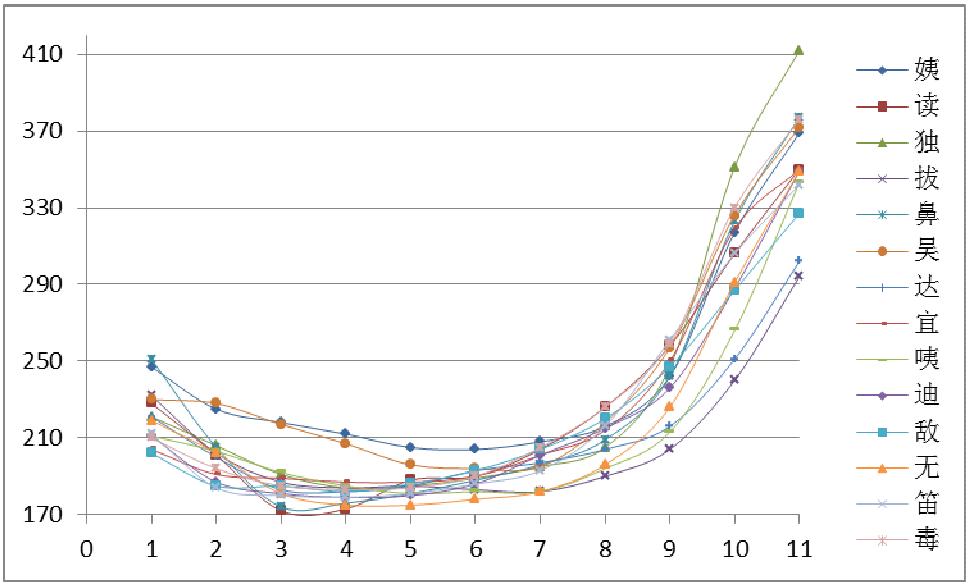


Figure 60. The F0 contour of 14 *Yangping* syllables pronounced by F4

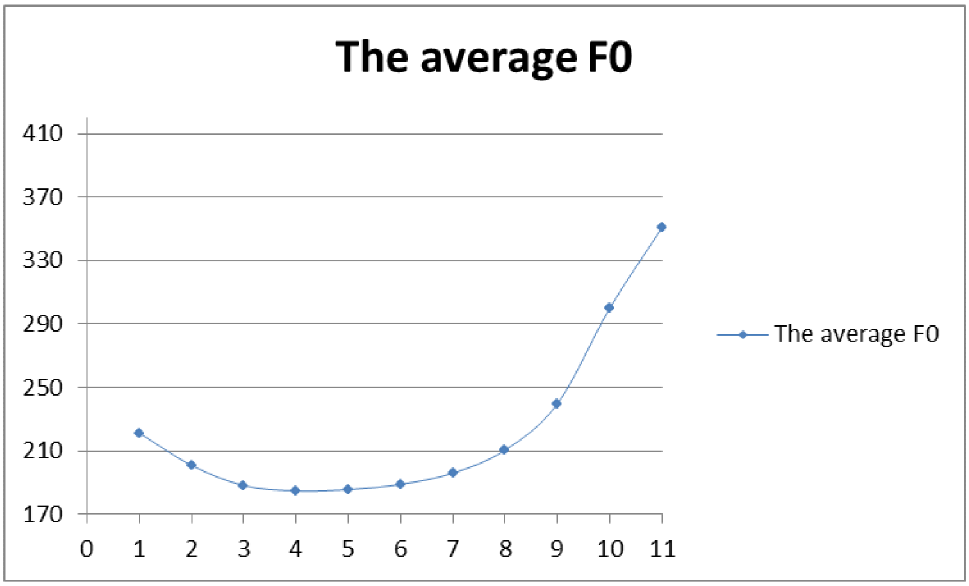


Figure 61. The average F0 contour of 14 *Yangping* syllables pronounced by F4

Based on Chart 1, the following conclusions can be made. (1) The F0 range of F4 is 170-420Hz, that is, about 250Hz. (2) The F0 contours of F4 are quite concentrated. (3) Most F0 contours have a very clear falling start. (4) Overall, the rising parts of most F0 contours are very steep, and because of the falling start, the F0 contours of F2 turn out as fall-rise pattern. (5) Most F0 contours start rising after the first half of the contour.

From Chart 2, we can see that the rising part of the mean F0 contour of F4 is very steep, starting at the fifth measuring point, in other words, at a point a about a 45% through the whole F0 contour. The mean F0 contour of F4 presents a fall-rise pattern. The value of the mean F0 contour ranges between 180 and 360Hz.

(5) M1

Table 53: *M1's F0 values for 14 Yangping syllables at 11 measuring points*

Syllables	1	2	3	4	5	6	7	8	9	10	11
姨 yí	129	117	111	107	107	108	110	114	120	137	156
读 dú	119	109	107	107	110	113	117	123	129	138	153
独 dú	113	110	104	102	99	101	101	115	128	139	162
拔 bá	93	103	107	105	103	107	110	121	129	148	143
鼻 bí	111	108	103	102	107	111	113	119	125	127	143
吴 wú	109	108	103	105	103	105	110	121	132	139	148
达 dá	110	106	105	103	105	116	108	113	122	131	138

宜 yí	113	113	107	107	109	112	115	123	131	147	154
咦 yí	114	110	104	102	103	105	109	117	136	146	159
迪 dí	113	109	106	105	103	104	109	115	124	139	147
敌 dí	115	106	104	102	103	109	115	122	129	138	150
无 wú	108	107	104	102	100	102	106	114	123	135	150
笛 dí	124	111	108	105	104	109	112	118	124	136	138
毒 dú	107	106	104	105	106	106	111	115	120	131	141
Average	112.7	108.8	105.5	104.2	104.4	107.7	110.4	117.9	126.6	137.9	148.7

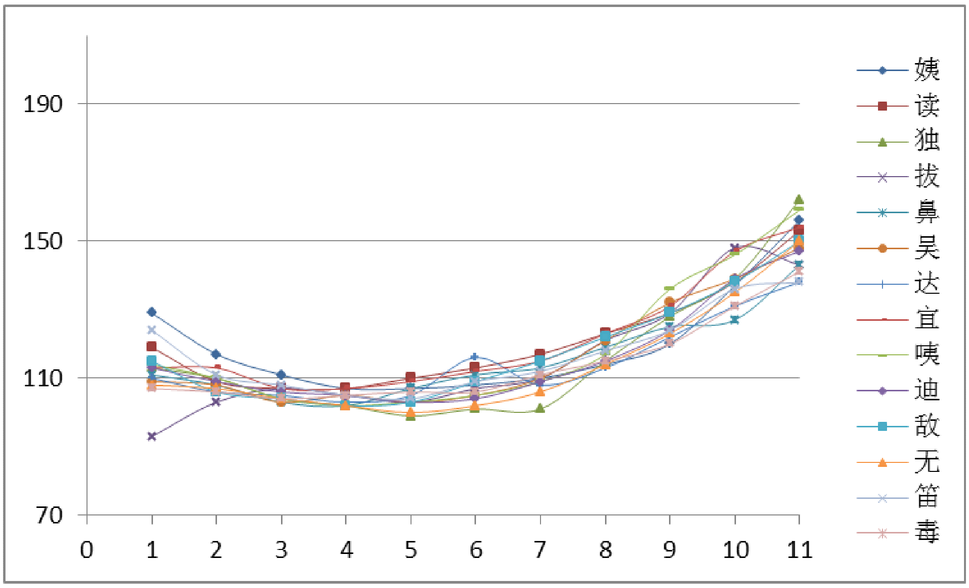


Figure 62. The F0 contour of 14 *Yangping* syllables pronounced by M1

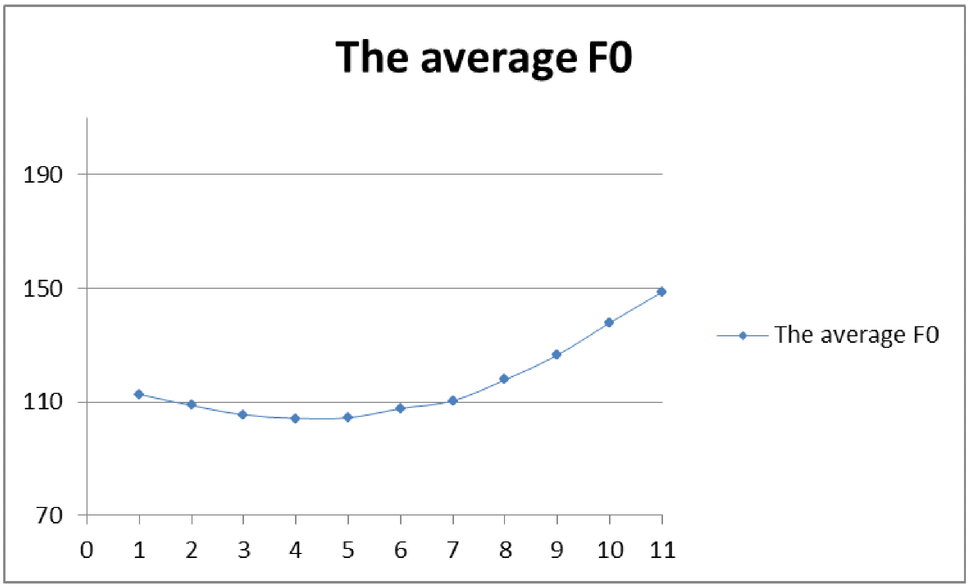


Figure 63. The average F0 contour of 14 *Yangping* syllables pronounced by M1

Based on Chart 1, the following conclusions can be made. (1) The F0 range of M1 is 90-170Hz, that is, about 80Hz. (2) The F0 contours of M1 are highly concentrated. (3) Most F0 contours have a quite smooth start, except for 拔 bá which contains a rising start. (4) Overall, the rising parts of most F0 contours are relatively gentle. (5) Most F0 contours appear as a rising tone which starts rising during the first half of the contour, 宜 yí is an exception which is fluctuant.

From Chart 2, we can see that the rising part of the mean F0 contour of M1 is relatively gentle with an inflection point of rising point at the fifth measuring point, in other words, at a point about 45% through of the whole F0 contour. The mean F0 contour of M1 appears as a rising shape. The value of the mean F0 contour ranges between 100 and 150Hz.

(6) M2

Table 54: *M2's F0 values for 14 Yangping syllables at 11 measuring points*

Syllables	1	2	3	4	5	6	7	8	9	10	11
姨 yí	117	109	107	108	110	113	117	119	125	133	140
读 dú	114	109	106	107	108	113	119	124	128	130	131
独 dú	116	106	104	106	108	109	111	114	116	120	131
拔 bá	109	105	106	107	109	111	116	121	126	129	126
鼻 bí	106	110	103	104	105	108	110	113	116	126	129
吴 wú	109	104	105	105	107	112	116	117	121	123	128
达 dá	104	106	108	109	108	111	111	116	120	124	135
宜 yí	120	116	112	109	111	113	118	122	127	133	140
咦 yí	101	104	106	106	110	117	124	129	134	139	142
迪 dí	107	110	108	110	114	117	122	125	128	133	133
敌 dí	111	111	107	108	108	108	112	114	117	120	138
无 wú	128	111	108	107	108	110	113	121	124	127	138
笛 dí	112	101	104	106	106	112	116	118	122	123	130
毒 dú	103	105	107	108	109	112	118	124	126	127	133
Average	111.2	107.6	106.5	107.1	108.6	111.9	115.9	119.8	123.6	127.6	133.9

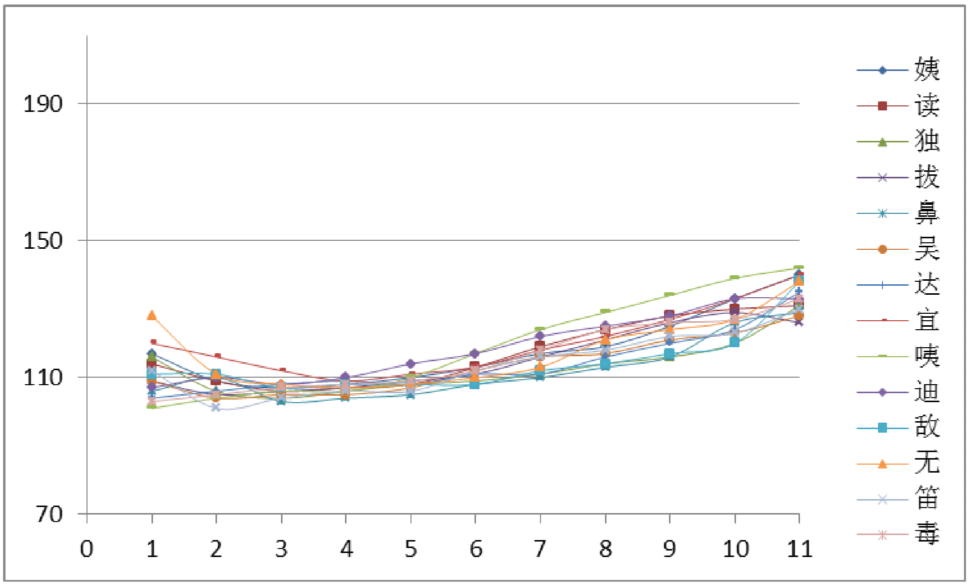


Figure 64. The F0 contour of 14 *Yangping* syllables pronounced by M2

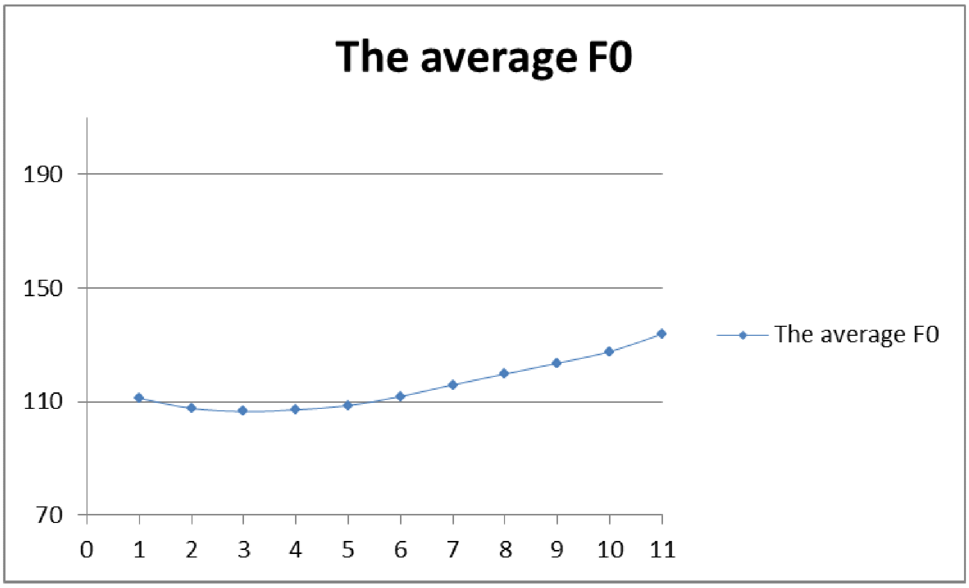


Figure 65. The average F0 contour of 14 Yangping syllables pronounced by M2

Based on Chart 1, the following conclusions can be made. (1) The F0 range of M2 is 100-150Hz, that is, about 50Hz. (2) The F0 contours of M2 are highly concentrated. (3) Most F0 contours have a smooth start. (4) Overall, the rising parts of most F0 contours are fairly gentle. (5) Most F0 contours appear as a rising tone which starts rising during the first half of the contour.

From Chart 2, we can see that the rising part of the mean F0 contour of M2 is quite mild with an inflection point of rising point at the fourth measuring point, in other words, at the point of about the 35% of the whole F0 contour. The mean F0 contour of M2 appears as a rising shape. The value of the mean F0 contour ranges between 105 and 135Hz.

(7) M3

Table 55: *M3's F0 values for 14 Yangping syllables at 11 measuring points*

Syllables	1	2	3	4	5	6	7	8	9	10	11
姨 yí	122	116	114	112	112	114	121	132	151	168	180
读 dú	120	114	112	114	117	119	122	134	152	169	173
独 dú	128	124	119	117	119	123	133	147	164	183	196
拔 bá	110	107	106	107	108	112	123	122	129	151	159
鼻 bí	111	111	113	116	116	119	124	131	145	159	169
吴 wú	113	111	113	115	115	118	124	133	148	167	186
达 dá	119	110	113	111	111	112	114	120	135	160	162
宜 yí	108	108	109	110	112	119	126	137	152	165	175
咦 yí	119	118	119	118	121	123	124	126	128	131	131
迪 dí	120	115	113	114	112	113	118	131	150	171	192
敌 dí	120	119	115	114	114	116	124	136	153	172	191
无 wú	110	106	108	109	110	113	120	133	153	179	200
笛 dí	118	118	114	113	111	116	128	140	156	173	186
毒 dú	120	119	117	116	117	119	126	140	159	182	202
Average	117	114	113.2	113.3	113.9	116.9	123.4	133	148.2	166.4	178.7

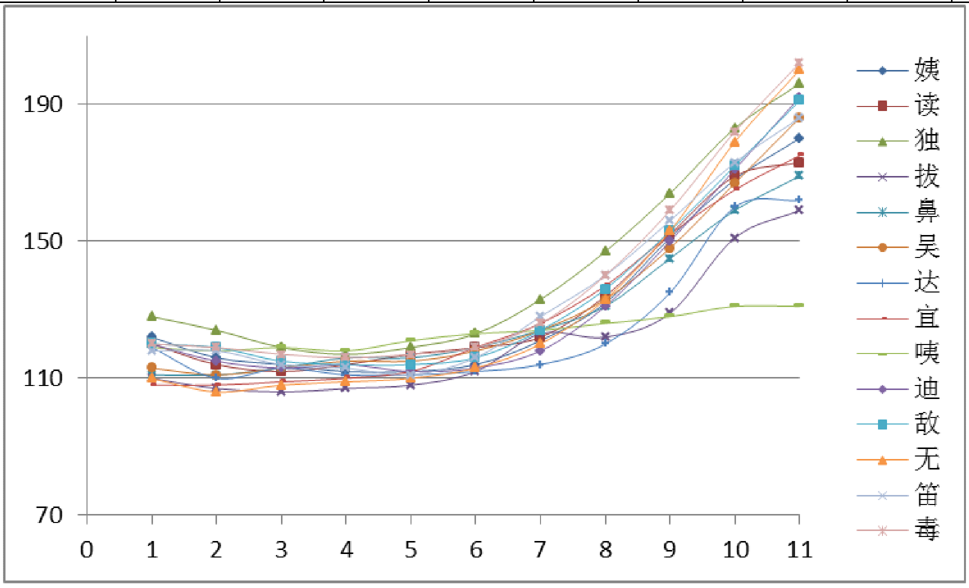


Figure 66. The F0 contour of 14 *Yangping* syllables pronounced by M3

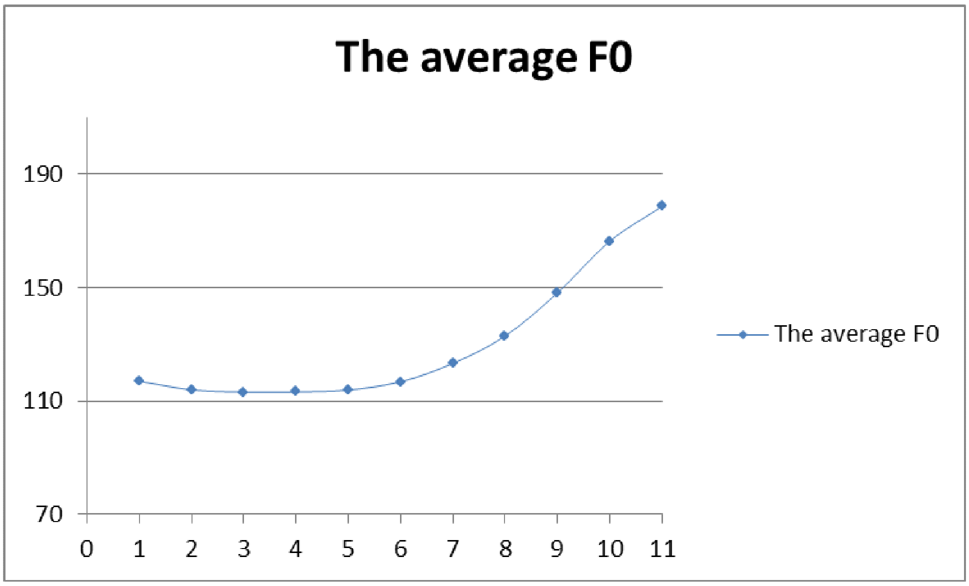


Figure 67. The average F0 contour of 14 *Yangping* syllables pronounced by M3

Based on Chart 1, the following conclusions can be made. (1) The F0 range of M3 is 100-210Hz, that is, about 110Hz. (2) The F0 contours of M3 are highly concentrated. (3) Most F0 contours have a very smooth start. (4) Overall, the rising parts of most F0 contours are appropriately steep, except for 咦 yí which does not present a rising trend. (5) Most F0 contours appear as rising tones which start rising during the first half of its contour.

From Chart 2, we can see that the rising part of the mean F0 contour of M3 is appropriately steep and starts at the fourth measuring point, in other words, at a point about the 35% through the whole F0 contour. The mean F0 contour of M3 takes as a rising shape. The value of the mean F0 contour ranges between 110 and 180Hz.

(8) M4

Table 56: *M4's F0 values for 14 Yangping syllables at 11 measuring points*

Syllables	1	2	3	4	5	6	7	8	9	10	11
姨 yí	120	121	122	123	122	122	122	122	122	123	126
读 dú	128	127	126	125	124	123	121	121	121	121	124
独 dú	134	131	129	129	129	129	129	129	130	131	135
拔 bá	123	120	119	119	120	120	121	120	120	121	122
鼻 bí	134	131	128	129	130	131	131	131	131	130	130
吴 wú	117	114	115	116	118	122	122	120	118	117	118
达 dá	120	117	115	113	115	118	117	117	117	118	119

宜 yí	125	127	128	129	130	131	130	128	128	128	131
咦 yí	116	118	119	123	127	129	128	125	123	123	125
迪 dí	133	129	129	130	129	129	128	128	128	128	127
敌 dí	128	129	130	131	131	130	130	130	129	129	131
无 wú	125	125	125	127	129	128	128	129	127	127	129
笛 dí	131	127	126	126	127	126	126	127	127	129	130
毒 dú	137	137	135	134	134	134	134	134	133	132	132
Average	126.5	125.2	124.7	125.3	126.1	126.6	126.2	125.8	125.3	125.5	127.1

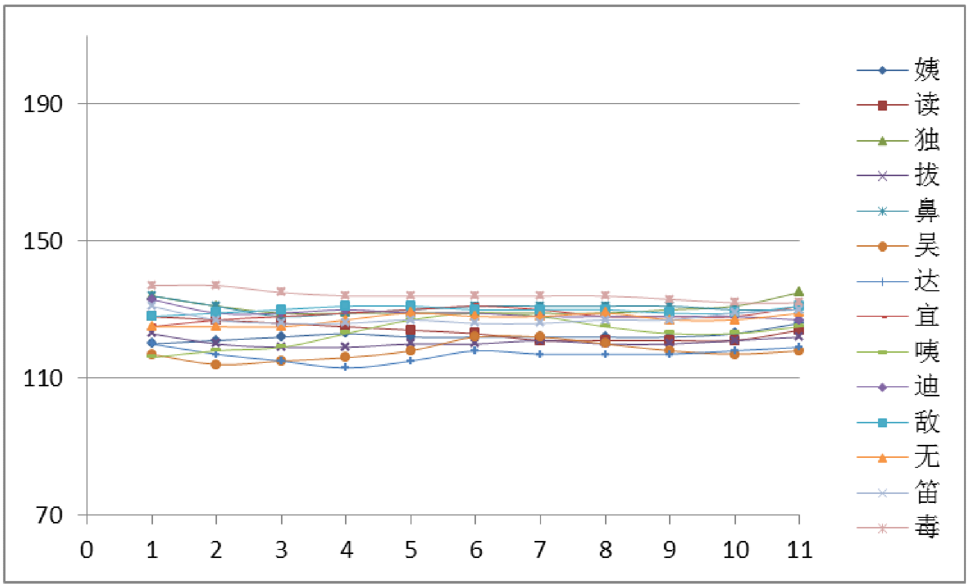


Figure 68. The F0 contour of 14 *Yangping* syllables pronounced by M4

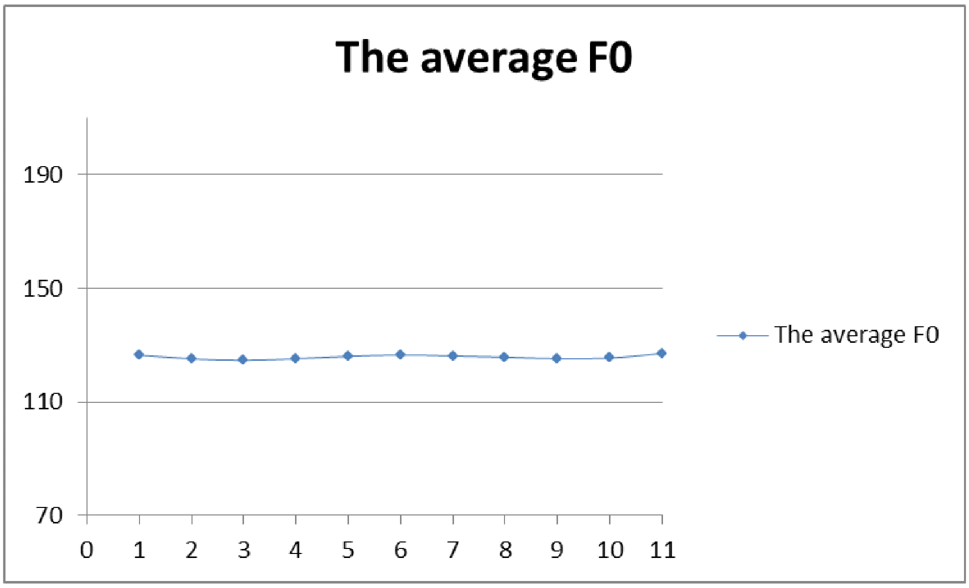


Figure 69. The average F0 contour of 14 *Yangping* syllables pronounced by M4

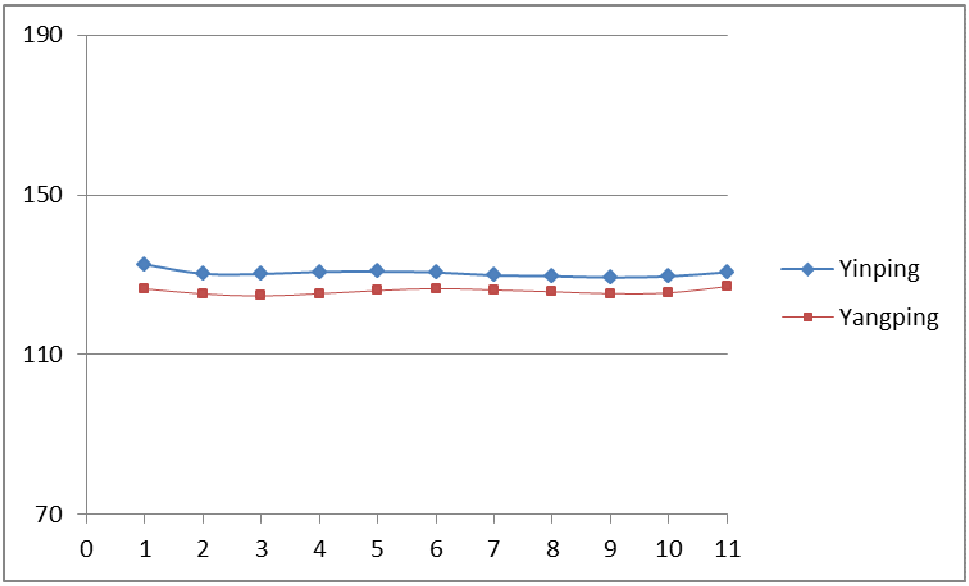


Figure 70. The average F0 contours of Yinping and Yangping syllables pronounced by M4

Based on Chart 1, the following conclusions can be made. (1) The F0 range of M4 is 110-140Hz, that is, about 30Hz. (2) The F0 contours of M4 are highly concentrated. (3) The F0 contours of M4 appear quite flat, without any obvious rising trend.

From Chart 2, we can see that the mean F0 contour of M4 is also quite flat, with no evident rise. The value of the mean F0 contour falls within the narrow range of 120-130Hz. For the purpose of comparison, we plotted M4's average F0 contours for *Yinping* and *Yangping* together on the same Chart (above). It clearly shows that two F0 contours are virtually parallel. In other words, the *Yangping* syllables pronounced by M4 appear as *Yinping*.

(9) M5

Table 57: *M5's F0 values for 14 Yangping syllables at 11 measuring points*

Syllables	1	2	3	4	5	6	7	8	9	10	11
姨 yí	113	101	98	96	95	94	97	100	112	136	157
读 dú	110	100	97	95	96	95	97	102	115	131	152
独 dú	112	103	99	97	95	96	99	109	125	147	160
拔 bá	107	100	98	97	99	104	113	123	139	151	161
鼻 bí	102	99	95	93	92	93	101	114	134	152	163
吴 wú	104	104	104	105	108	115	127	144	161	175	181
达 dá	109	98	95	95	95	94	96	104	118	137	156
宜 yí	112	102	98	98	97	97	97	106	120	144	163

咦 yí	117	109	103	100	98	98	100	104	115	138	160
迪 dí	120	105	101	97	96	94	97	107	123	144	162
敌 dí	109	101	98	96	96	97	98	101	112	128	148
无 wú	120	110	106	103	100	99	97	95	97	111	135
笛 dí	101	97	95	93	93	93	94	98	111	129	148
毒 dú	106	100	97	96	95	95	95	102	114	132	149
Average	110.1	102.1	98.9	97.2	96.8	97.4	100.6	107.8	121.1	139.6	156.8

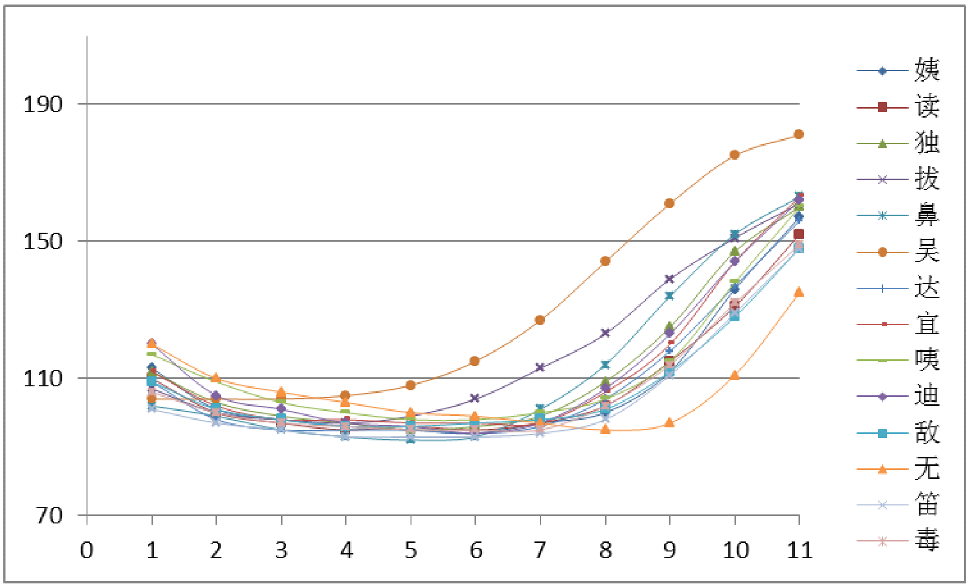


Figure 71. The F0 contour of 14 Yangping syllables pronounced by M5

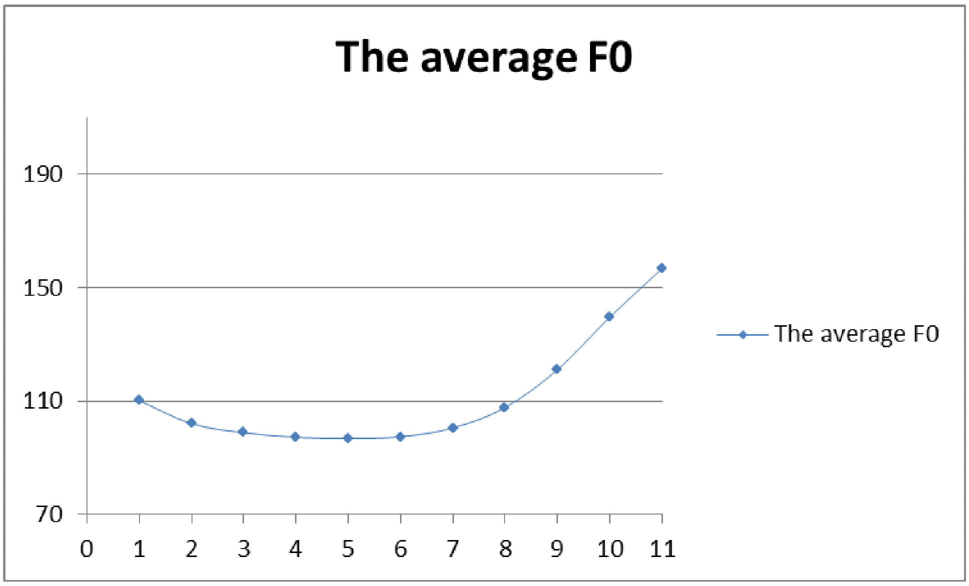


Figure 72. The average F0 contour of 14 Yangping syllables pronounced by M5

Based on Chart 1, the following conclusions can be made. (1) The F0 range of M5 is 90-180Hz, that is, about 90Hz. (2) The F0 contours of M5 are highly concentrated, except for 吴 wú and 无 wú. (3) Most F0 contours have a relatively obvious falling start. (4) Overall, the rising parts of most F0 contours are appropriately steep, but because of the falling start, the F0 contours of M5 follow fall-rise patterns. (5) Most F0 contours start to rise at the mid-point of the contour.

From Chart 2, we can see that the rising part of the mean F0 contour of M5 is appropriately steep, but only starts at the sixth measuring point, in other words, at a point of about 55% through the whole F0 contour. The mean F0 contour of M5 follows a fall-rise pattern. The value of the mean F0 contour is between 95-160Hz.

(10) M6

Table 58: *M6's F0 values for 14 Yangping syllables at 11 measuring points*

Syllables	1	2	3	4	5	6	7	8	9	10	11
姨 yí	100	100	98	93	97	105	111	116	121	127	142
读 dú	89	94	91	94	97	102	107	111	114	117	122
独 dú	92	89	85	87	92	95	99	103	111	120	122
拔 bá	83	90	88	83	85	89	90	94	99	101	103
鼻 bí	83	93	90	89	91	94	99	107	113	118	120
吴 wú	84	91	91	94	96	98	99	102	105	109	114
达 dá	99	93	88	85	85	84	88	92	94	99	103
宜 yí	96	94	93	91	90	94	96	102	104	108	111
咦 yí	95	95	94	96	97	97	100	103	106	108	111
迪 dí	93	97	96	95	96	98	100	104	107	110	111
敌 dí	87	86	79	78	85	87	98	104	108	112	112
无 wú	91	89	89	91	92	94	97	99	101	105	110
笛 dí	97	95	90	89	89	89	91	90	103	111	111
毒 dú	94	87	86	89	90	92	95	102	109	116	117
Average	91.6	92.4	89.9	89.6	91.6	94.1	97.9	102.1	106.8	111.5	114.9

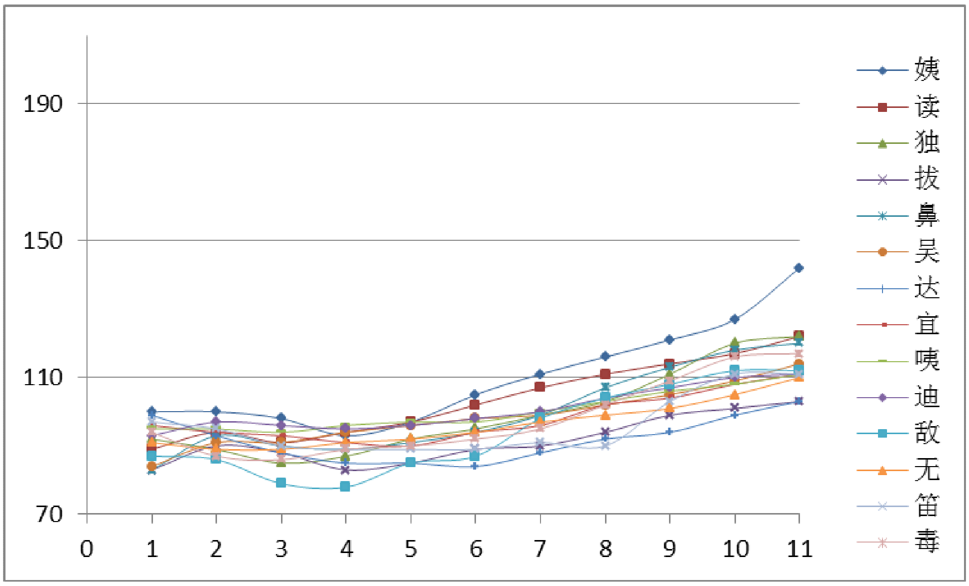


Figure 73. The F0 contour of 14 *Yangping* syllables pronounced by M6

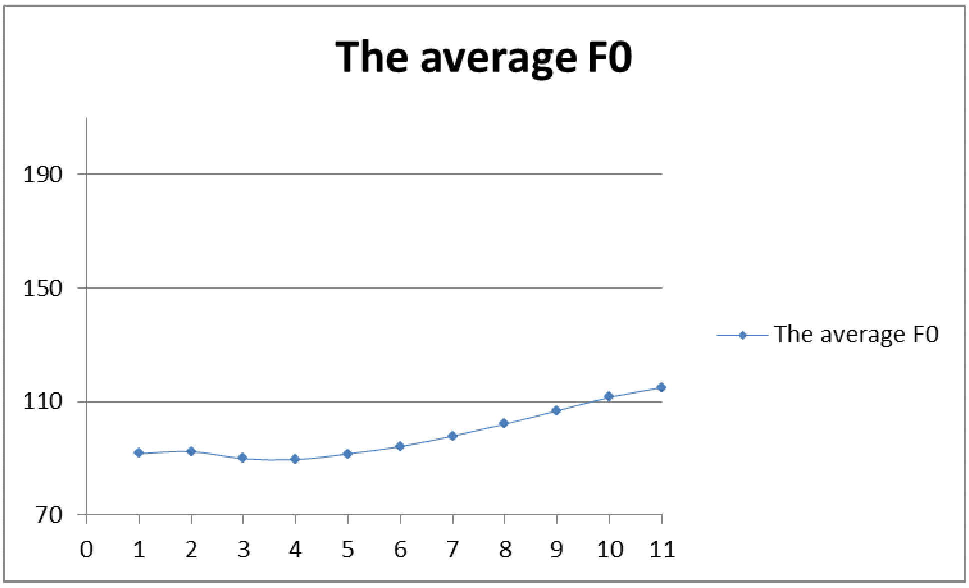


Figure 74. The average F0 contour of 14 *Yangping* syllables pronounced by M6

Based on Chart 1, the following conclusions can be made. (1) The F0 range of M6 is 75-145Hz, that is, about 70Hz. (2) The F0 contours of M6 are highly concentrated, except for 姨 yí. (3) Most F0 contours have a very smooth start. (4) Overall, the rising parts of most F0 contours are quite gentle. (5) Most F0 contours appear as a rising tone which begins during the first half of the contour.

From Chart 2, we can see that the rising part of the mean F0 contour of M6 is quite gentle with and start at the fourth measuring point, in other words, at a point about 45% through the whole F0 contour. The mean F0 contour of M6 follows a rising pattern. The value of the mean F0 contour ranges between 85 and 115Hz.

(11) FB

Table 59: *FB's F0 values for 14 Yangping syllables at 11 measuring points*

Syllables	1	2	3	4	5	6	7	8	9	10	11
姨 yí	162	164	164	164	165	171	176	186	202	219	237
读 dú	171	164	168	159	159	171	184	194	215	234	245
独 dú	177	160	152	151	154	159	170	183	200	214	225
拔 bá	159	142	138	138	137	144	146	159	178	196	210
鼻 bí	183	179	175	172	173	176	185	201	222	237	240
吴 wú	169	177	178	173	170	175	181	191	212	218	240
达 dá	167	151	141	143	145	148	162	177	191	202	219
宜 yí	172	168	166	166	167	167	175	184	197	214	224
咦 yí	163	165	162	160	160	170	185	201	215	224	224
迪 dí	163	165	162	160	160	168	183	199	214	222	226
敌 dí	177	172	165	164	168	172	191	209	221	234	238
无 wú	170	173	178	176	174	178	188	205	232	239	244
笛 dí	183	183	180	177	179	184	197	214	230	247	253
毒 dú	180	174	168	170	171	175	191	209	224	237	247
Average	171.1	166.9	164.1	162.4	163	168.4	179.6	193.7	210.9	224.1	233.7

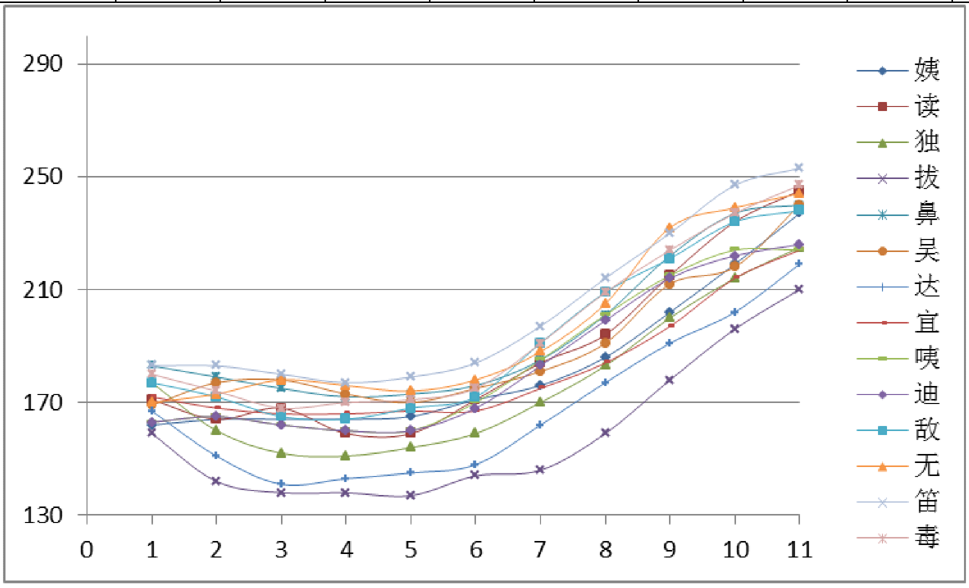


Figure 75. The F0 contour of 14 *Yangping* syllables pronounced by FB

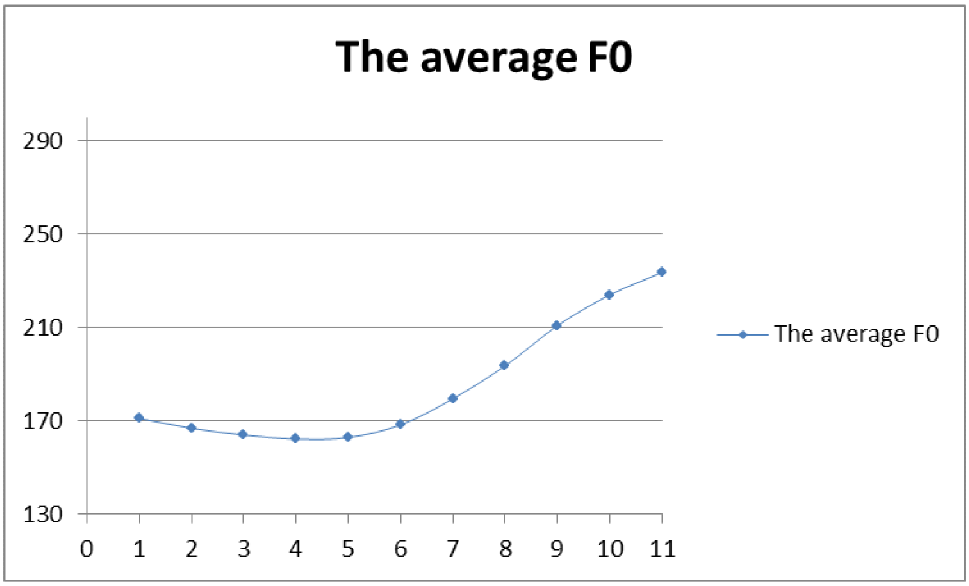


Figure 76. The average F0 contour of 14 Yangping syllables pronounced by FB

Based on Chart 1, the following conclusions can be made. (1) The F0 range of FB is 135-255Hz, that is, about 120Hz. (2) The F0 contours of FB are relatively concentrated. (3) Most F0 contours have a quite start. (4) Overall, the rising parts of most F0 contours are appropriately steep. (5) Most F0 contours appear as a rising tone which begins during the first half of the contour.

From Chart 2, we can see that the rising part of the mean F0 contour of FB is appropriately steep with an inflection point of rising point at the fifth measuring point, in other words, at a point about 45% through the whole F0 contour. The mean F0 contour of FB follows a rising pattern. The value of the mean F0 contour ranges between 160 and 240Hz.

(12) MB

Table 60: *MB's F0 values for 14 Yangping syllables at 11 measuring points*

Syllables	1	2	3	4	5	6	7	8	9	10	11
姨 yí	103	102	101	101	103	105	110	115	124	125	127
读 dú	105	103	102	101	103	107	112	116	126	138	142
独 dú	103	100	101	102	104	108	112	117	125	130	139
拔 bá	103	98	97	97	100	104	111	118	126	134	136
鼻 bí	103	99	98	99	102	105	110	118	126	135	144
吴 wú	103	103	104	107	111	115	120	126	136	148	157
达 dá	108	99	98	99	100	102	108	114	121	128	136

宜 yí	105	103	104	105	108	113	121	121	142	152	164
咦 yí	102	101	102	105	108	112	120	133	147	158	166
迪 dí	114	112	111	113	117	125	137	149	160	170	177
敌 dí	106	101	100	101	105	110	118	130	143	151	156
无 wú	107	108	107	108	112	116	126	138	149	158	177
笛 dí	107	106	106	108	114	121	131	139	147	158	170
毒 dú	110	105	105	106	108	111	118	128	139	147	156
Average	105.6	102.9	102.6	103.7	106.8	111.1	118.4	126.9	136.8	145.8	153.7

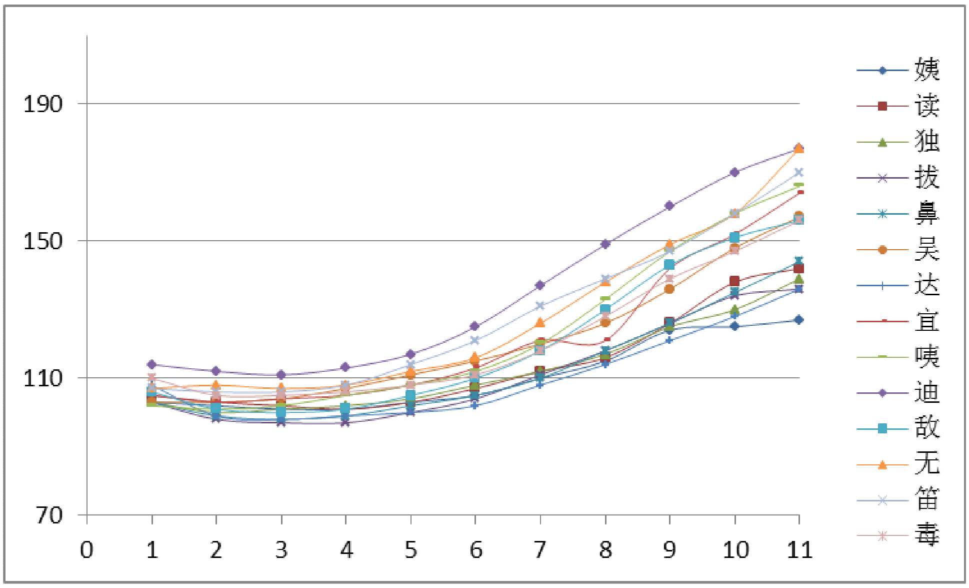


Figure 77. The F0 contour of 14 Yangping syllables pronounced by MB

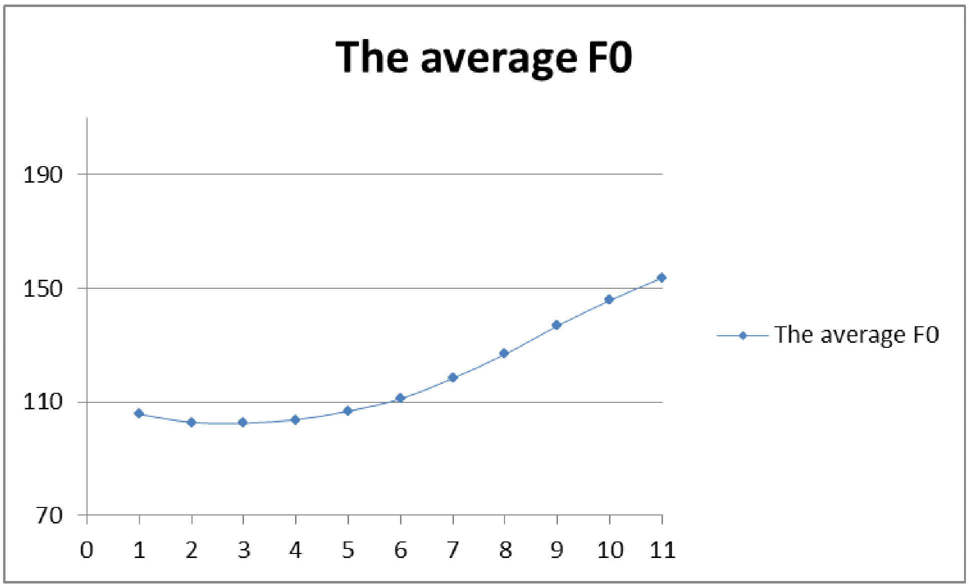


Figure 78. The average F0 contour of 14 *Yangping* syllables pronounced by MB

Based on Chart 1, the following conclusions can be made. (1) The F0 range of MB is 90-210Hz, that is, about 120Hz. (2) The F0 contours of MB are relatively concentrated. (3) Most F0 contours have a very smooth start. (4) Overall, the rising parts of most F0 contours are appropriately steep. (5) Most F0 contours appear as a rising tone which starts rising during the first half of the contour.

From Chart 2, we can see that the rising part of the mean F0 contour of MB is appropriately steep and starts rising at the fourth measuring point, in other words, at a point about 35% through the whole F0 contour. The mean F0 contour of FB follows a rising pattern. The value of the mean F0 contour ranges between 100 and 160Hz.

5.2.2.2. General comments

From the tables and figures presented above, we can see that only 20% of the Hungarian subjects (M1 and M3) pronounce *Yangping* syllables with an appropriate rising contour. 80% of the subjects produce deviant contours, mainly due to a falling start, an excessively steep or gentle rise, a fall-rise or flat pattern, or because the rise begins at an inappropriate point.

In order to better compare the 10 Hungarian subjects with the 2 native speakers in general terms, we made a chart for the average F0 contours of *Yangping* syllables pronounced by 12 subjects. The vertical axis covers a fixed value range which is from 50Hz to 400Hz.

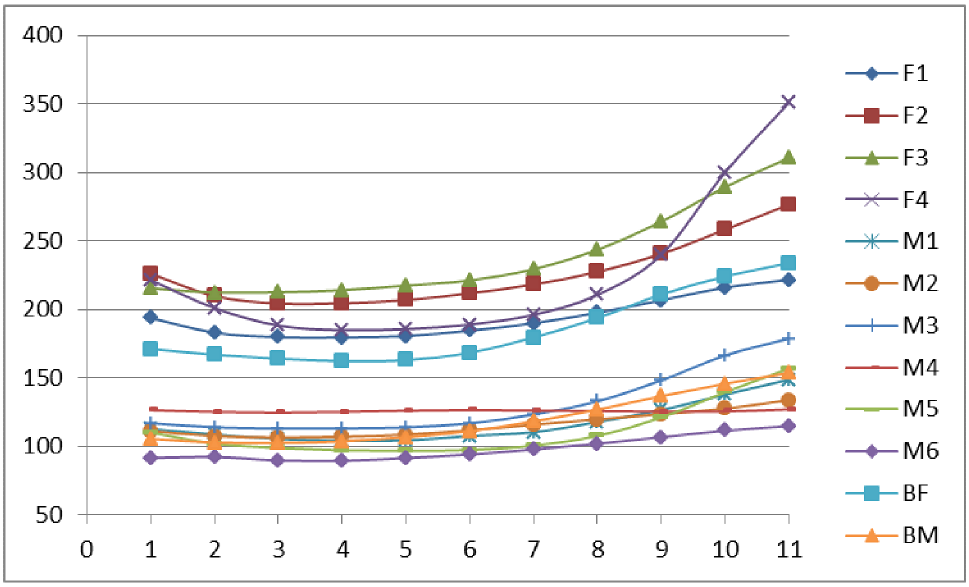


Figure 79. The average F0 contours of *Yangping* syllables pronounced by 12 subjects

On the basis of the average F0 contours, the following conclusions can be made:

(1) 90% of the Hungarian university students have a rising average F0 contour when pronouncing *Yangping* syllables;

(2) 20% of the Hungarian subjects pronounce *Yangping* syllables correctly; in other words, their average F0 contours are appropriately lean with a smooth start;

(3) 80% of the Hungarian subjects do not pronounce *Yangping* syllables appropriately, even though most of them do manage to pronounce *Yangping* syllables with a rising pattern. The deviant contours are as follows: F1, F2, F4 and M5 pronounce *Yangping* syllables with a quite apparent falling start, therefore, their F0 contours, especially F2's and F4's, appear as fall-rise patterns which are close to *Shangsheng*; F3's F0 contour starts rising relatively early and follows too steep a rising pattern, while M5's starts to rise relatively late; M2 and M6's F0 contours follow too gentle a rising pattern; M4's F0 contour is very flat and actually runs virtually parallel with his *Yinping* F0 contour.

(4) The F0 values produced by females are considerably higher than males': about twice as high;

(5) The average F0 contours of males are better than females', which suggests that males have fewer problems in producing appropriate *Yangping* syllables than females.

5.2.3. Intrinsic vowel fundamental frequency

The intrinsic vowel fundamental frequency, shortened for IF0, refers to the fact that the F0 of a high vowel is generally higher than that of a low one when the two vowels are produced under the same conditions[184]. Zhu Xiaonong believes that the IF0 in the high area of the individual's own frequency domain is more distinct than that in the low area, and the IF0 appears more obviously in female than in male speakers (Zhu Xiaonong, 2005,79).

In order to define the relationship between the height of the F0 contours and vowels, we classified all *Yangping* syllables produced by each subject in terms of /i/, /u/ and /a/, and averaged the F0 values of each category. The following table shows the relevant details for

[184] Lehiste, Ilse 1970: 68.

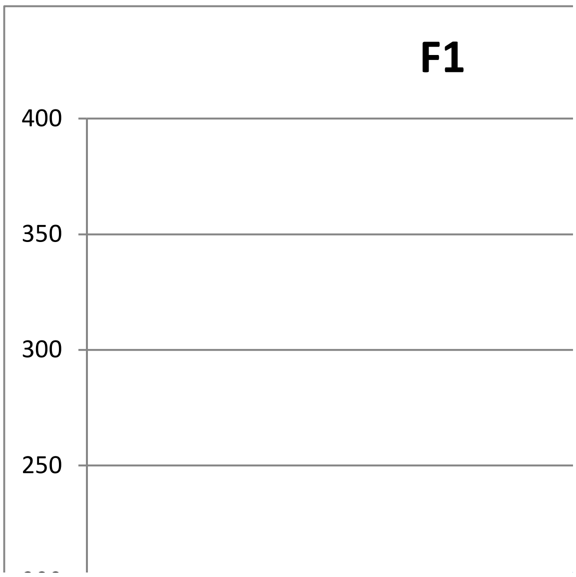
each subject, as well as the average F0 of all the Hungarian female subjects and the Hungarian male subjects (labeled F, M, respectively).

Table 61: *Yangping F0 of each Hungarian female subject and Hungarian male subject, with averages for all female and all male subjects.*

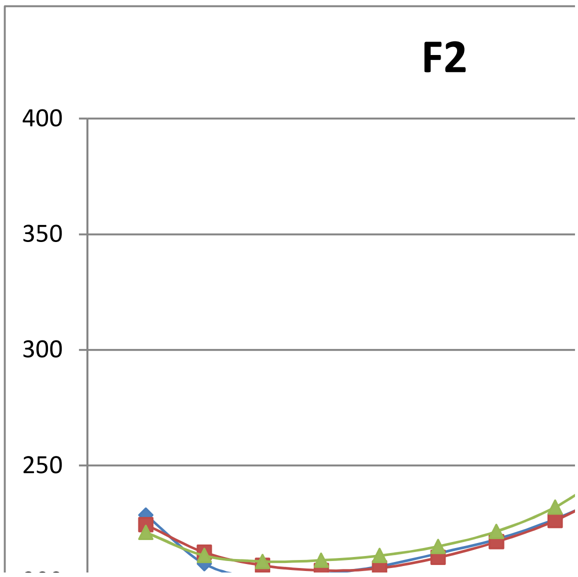
		1	2	3	4	5	6	7	8	9	10	11	ms
F1	/i/	197.1	183.7	180	179.7	181.1	185.1	191.6	200.1	209.6	218.6	223	475
	/u/	191.2	184.2	181.8	182.2	182.4	185.6	190.2	197.6	205.6	213.8	219.4	433
	/a/	188	178	173.5	172.5	174	180	183.5	188.5	199	212	222.5	416
F2	/i/	228.6	207.4	201.4	203	206.4	212	218.1	226.9	238.6	255	272.6	456
	/u/	224.4	212.4	206.8	204.6	205.6	210.2	217	226.2	240.2	261.6	284	499
	/a/	221	211	208.5	209	211	215	221.5	232	248	261.5	271.5	392
F3	/i/	216.3	213.9	214.6	216.6	219.9	224	232.4	247.3	268.1	294.9	320	277
	/u/	209.2	207.2	208	209.6	214	219	229.6	248.6	275.2	307.2	329.8	307
	/a/	228	220	216.5	216.5	217.5	217	218	218.5	222	225	229.5	269
F4	/i/	219.7	197.1	188.4	185.7	185.9	189.9	197.9	212	241.6	300.9	351	365
	/u/	221.6	206.2	189.2	184.4	185.6	188.2	196.2	214.2	249.4	320.8	371.8	376
	/a/	226.5	201	185	183	185	188	189.5	197	210	245.5	298	353
M1	/i/	117	110.6	106.1	104.3	105.1	108.3	111.9	118.3	127	138.6	149.6	321
	/u/	111.2	108	104.4	104.2	103.6	105.4	109	117.6	126.4	136.4	150.8	319
	/a/	101.5	104.5	106	104	104	111.5	109	117	125.5	139.5	140.5	268
M2	/i/	110.6	108.7	106.7	107.3	109.1	112.6	117	120	124.1	129.6	136	458
	/u/	114	107	106	106.6	108	111.2	115.4	120	123	125.4	132.2	440
	/a/	106.5	105.5	107	108	108.5	111	113.5	118.5	123	126.5	130.5	358
M3	/i/	116.9	115	113.9	113.9	114	117.1	123.6	133.3	147.9	162.7	174.9	227
	/u/	118.2	114.8	113.8	114.2	115.6	118.4	125	137.4	155.2	176	191.4	237
	/a/	114.5	108.5	109.5	109	109.5	112	118.5	121	132	155.5	160.5	227
M4	/i/	126.7	126	126	127.3	128	128.3	127.9	127.3	126.9	127.1	128.6	333
	/u/	128.2	126.8	126	126.2	126.8	127.2	126.8	126.6	125.8	125.6	127.6	319
	/a/	121.5	118.5	117	116	117.5	119	119	118.5	118.5	119.5	120.5	350
M5	/i/	110.6	102	98.3	96.1	95.3	95.1	97.7	104.3	118.1	138.7	157.3	565
	/u/	110.4	103.4	100.6	99.2	98.8	100	103	110.4	122.4	139.2	155.4	572.6
	/a/	108	99	96.5	96	97	99	104.5	113.5	128.5	144	158.5	404
M6	/i/	93	94.3	91.4	90.1	92.1	94.9	99.3	103.7	108.9	113.4	116.9	330
	/u/	90	90	88.4	91	93.4	96.2	99.4	103.4	108	113.4	117	330.4
	/a/	91	91.5	88	84	85	86.5	89	93	96.5	100	103	273
F	/i/	215.4	200.5	196.1	196.3	198.3	202.8	210	221.6	239.5	267.3	291.6	393
	/u/	211.6	202.5	196.5	195.2	196.9	200.8	208.3	221.7	242.6	275.9	301.3	404

	/a/	215.9	202.5	195.9	195.3	196.9	200	203.1	209	219.8	236	255.4	357
M	/i/	112.5	109.4	107.1	106.5	107.3	109.4	112.9	117.8	125.5	135.0	143.9	372
	/u/	112	108.3	106.5	106.9	107.7	109.7	113.1	119.2	126.8	136	145.7	370
	/a/	107.2	104.6	104	102.8	103.5	106.5	108.9	113.6	120.7	130.8	135.6	313
FB	/i/	171.9	170.9	167.7	166.1	167.4	172.6	184.6	199.1	214.4	228.1	234.6	464
	/u/	173.4	169.6	168.8	165.8	165.6	171.6	182.8	196.4	216.6	228.4	240.2	469
	/a/	163	146.5	139.5	140.5	141	146	154	168	184.5	199	214.5	392
MB	/i/	105.7	103.4	103.1	104.6	108.1	113	121	129.3	141.3	149.9	157.7	557
	/u/	105.6	103.8	103.8	104.8	107.6	111.4	117.6	125	135	144.2	154.2	485.8
	/a/	105.5	98.5	97.5	98	100	103	109.5	116	123.5	131	136	480.5

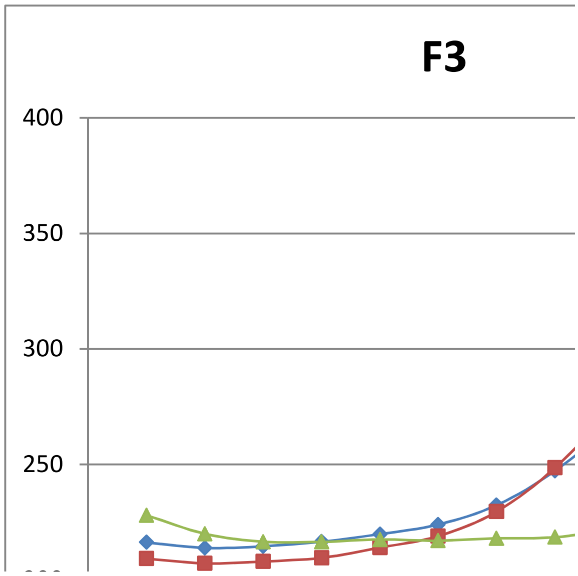
Based on the above table, we made charts of the F0 contours of /i/, /u/ and /a/ for each subject. In each chart the vertical (Y) axis shows the value of fundamental frequency F0, with Hz as the basic unit of measurement. The horizontal (X) axis presents the 11 measuring points. For better comparison purposes, the vertical axis covers a fixed value range based on the maximum and minimum frequencies recorded by female and male subjects respectively. For female subjects, the range is from 100Hz to 400Hz; for male subjects, from 50Hz to 250Hz.

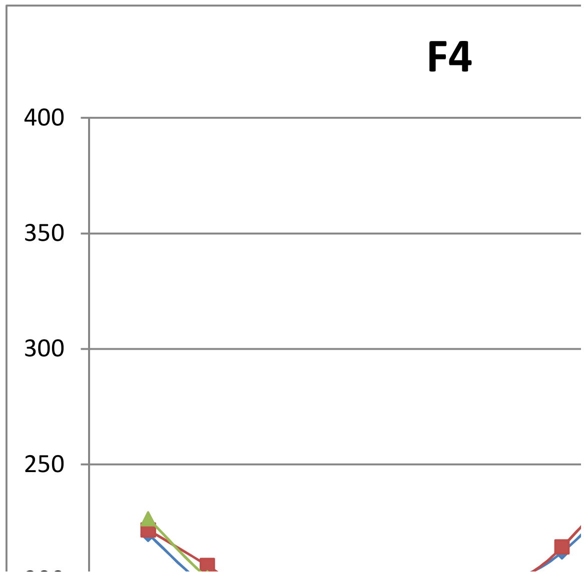


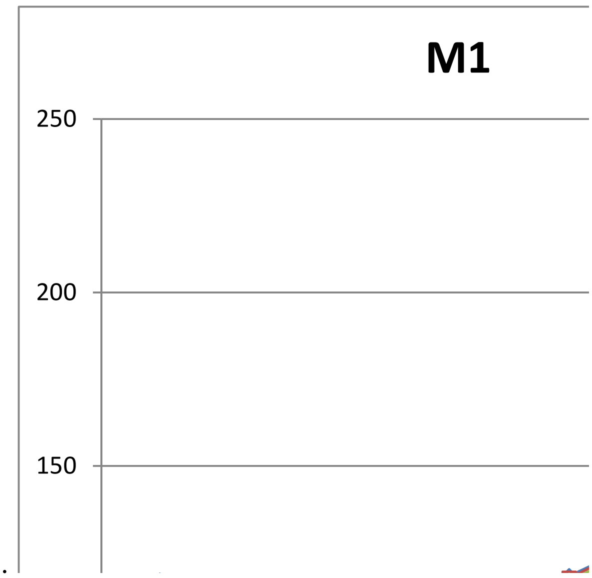
F1

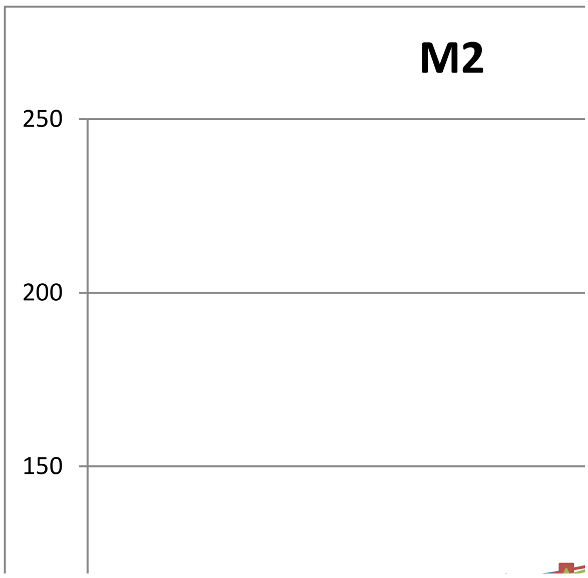


F2


F3
400
350
300
250
F3


F4
400
350
300
250
F4


M1
250
200
150
M1


M2
250
200
150
M2

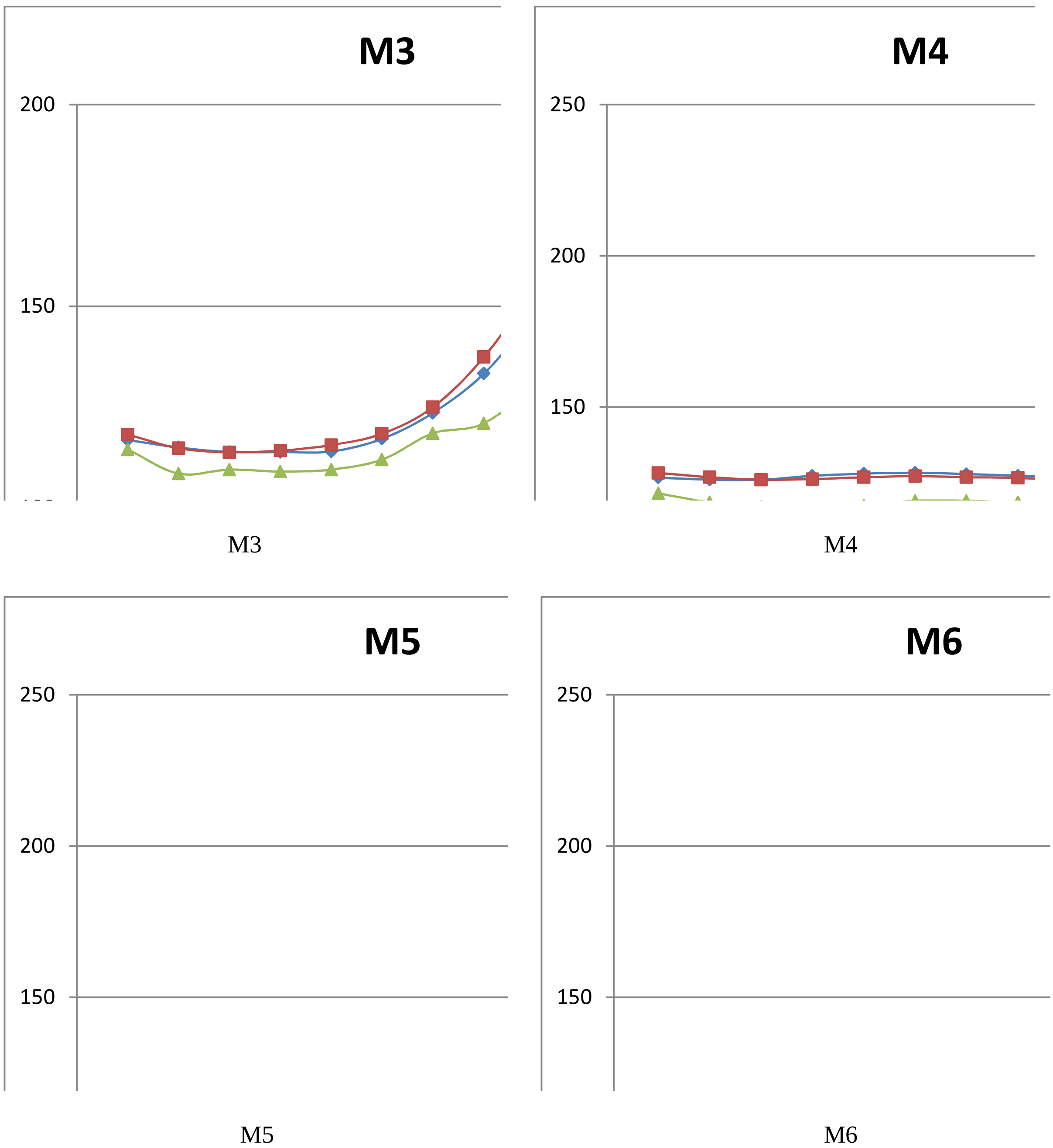


Figure 80. Average *Yangping* F0 contours of /i/, /u/ and /a/ for each Hungarian subject

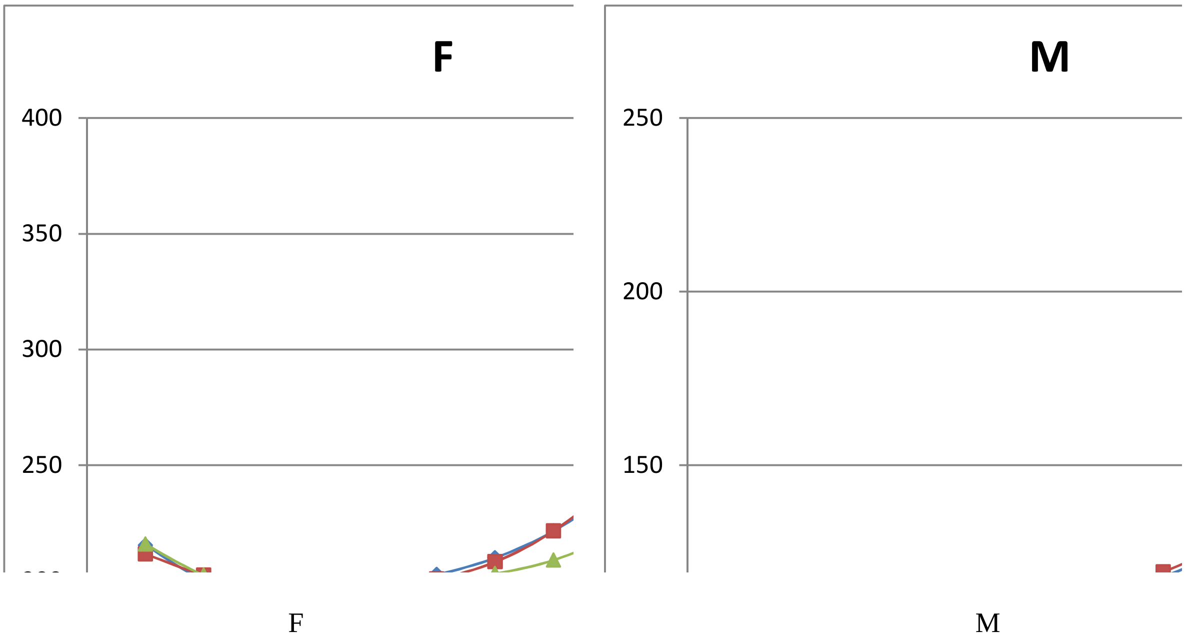


Figure 81. Average *Yangping* F0 contours of /i/, /u/ and /a/ for all all Hungarian subjects

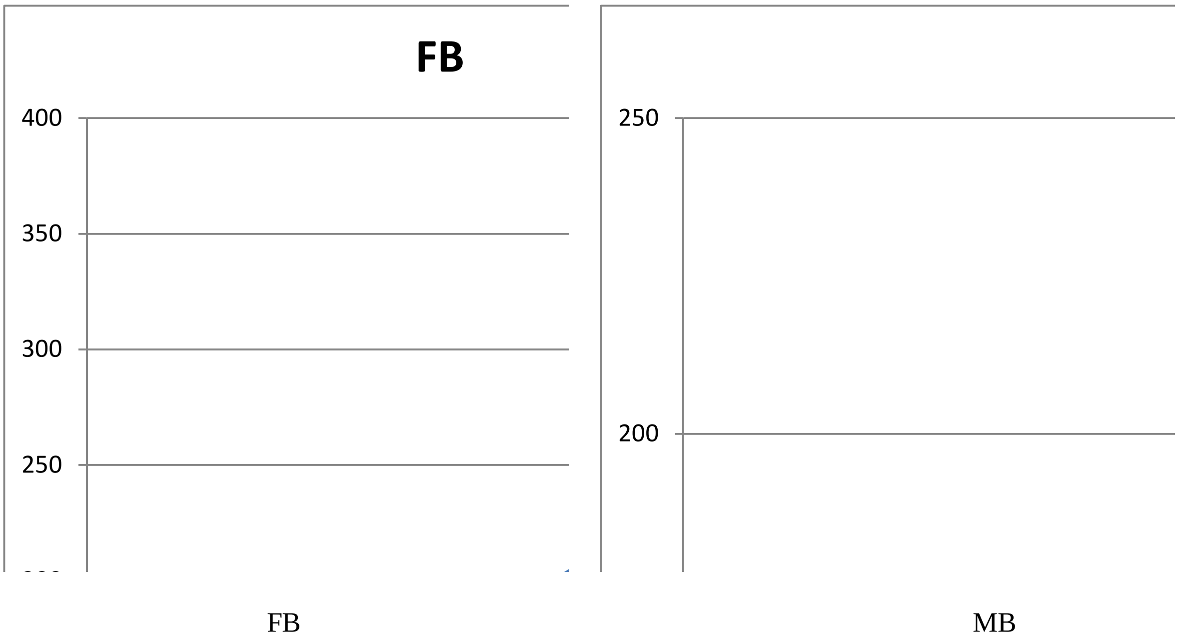


Figure 82. Average *Yangping* F0 contours of /i/, /u/ and /a/ for all both Chinese subjects

From the charts presented above, we can see that: (1) the manifestations of the IF0 in *Yangping* syllables produced by FB and MB are not very clear, since their average F0 contours of /i/ and /u/ are considerably separated from those of /a/. The performance of IF0 appears in MB's average F0 contours from the seventh measuring point. (2) The *Yangping* syllables produced by F3, F4, M3 and M6 have distinctly different F0 levels in terms of /i/, /u/ and /a/, with the order of /u/＞/i/＞/a/ or /u/≥/i/＞/a/. The performance of the IF0 exists in these subjects' average F0 contours mainly from the second half of their F0 contours. (3) The performances of the IF0 in F1, F2, M1, M2, M4 and M6 are unobvious, since their average F0 contours of /i/, /u/ and /a/ are highly concentrated and some of them are even almost overlapping. (4) The average F0 contours of /i/, and /u/ produced by all Hungarian female subjects are almost overlapping and relatively higher that of /a/ produced by them from the second half part of their contours, which indicates the considerably apparent performances of the IF0. On the other hand the average F0 contours of /i/, /u/ and /a/ produced by all Hungarian male subjects are quite concentrated, which prevents the show the obvious appearance of the IF0. (5) Based on those whose average F0 contours reflect the manifestations of IF0 in *Yangping* syllables, we can see that the IF0 in high areas of the individual's own frequency domain is more distinct than that that in low areas, which corresponds to Zhu Xiaonong's statement[185] (See 5.1.3 above).

All in all, the manifestations of IF0 in *Yangping* tone produced by native speaker are fairly significant. However, six of the 10 Hungarian subjects who produce the *Yangping* tone without showing the obvious manifestations of the IF0, in other words, 40% of the Hungarian leaners, are likely to produce *Yangping* tone accompanied by the quite obvious existence of the IF0.

5.2.4. Range

For *Yangping*, the range indicates how steep the contour is. At this point, range refers to the difference between the maximal and minimum value of F0. Therefore, the narrower the range, the less steep the contour of *Yangping*.

185 Zhu 2005.

Table 62: *The average range of 14 Yangping syllables pronounced by 12 subjects*

Subjects	Maximal value of average F0	Minimum value of average F0	Range
F1	221.6	179.6	42
F2	276.5	204.4	72.1
F3	310.6	212.4	98.2
F4	350.9	184.9	166
M1	148.7	104.2	44.5
M2	133.9	106.5	27.4
M3	178.7	113.2	65.5
M4	127.1	124.7	2.7
M5	156.8	96.8	60
M6	114.9	89.6	25.3
FB	233.7	162.4	71.3
MB	153.7	102.6	51.1

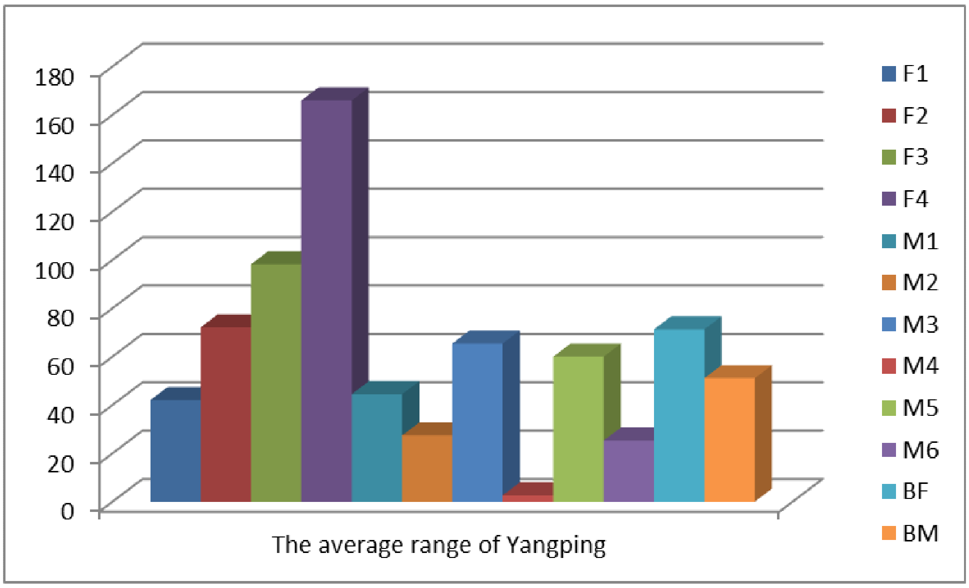


Figure 83. The average range of 14 *Yangping* syllables pronounced by 12 subjects

From the table and the figure presented above, we can see: (1) the average ranges of FB and MB are 71.3Hz and 51.1Hz; (2) the average ranges of F2, M1, M2, M3, M5 are very close to FB's and MB's. In other words, the steepness of the *Yangping* contours pronounced by these 5 subjects are very appropriate; (3) the average ranges of F3 and F4 are much higher than those of FB and MB, especially M4 which is 2-3 times of FB's and MB's; (4) the average ranges of

F1, M2 , M4 and M6 are considerably (and, this case perceptibly and disturbingly) lower than FB's and MB's, which indicates that these four learners do not produce acceptably steep levels in *Yangping*. What is more, the average range of M4's *Yangping* is very narrowand runs parallel with his *Yinping* contour. All in all, 40% of Hungarian university learners do not appear to experience problems in pronouncing *Yangping* in terms of range, but 60% of them do find it difficult.

5.2.5. Summary

In this section, the recordings of 12 subjects are examined in terms of the duration, the fundamental frequency (F0) and the tone range of their *Yangping* syllables. On the basis of measurement, extraction, analysis and comparison, the following conclusions can be drawn.

(1) In terms of the duration, 60% of Hungarian university learners – two female and two male subjects in the sample - differ considerably from the "model" native speakers in terms of the duration of *Yangping* syllables, all in producing much shorter durations. What is more, 20% of Hungarian learners produce deviant orders of durations of /i/, /u/ and /a/, which indicates that the inappropriate production of vowels probably causes problems in duration.

(2) With regard to fundamental frequency (F0), only 20% of the Hungarian subjects pronounce *Yangping* syllables correctly; in other words, their average F0 contours are appropriately steep with a smooth start; in contrasts, 80% of the Hungarian subjects have difficulty producing appropriately flat contours. The deviant contours they produce mainly appear as a contour with a falling start, too steep or too gentle a rising pattern, a fall-rise pattern or a flat pattern, or an inappropriate rising point. As to the pitch value of F0, it emerges that the male Hungarian subjects are closely grouped in terms of pitch around the native male, while all the female Hungarian subjects were higher in pitch than the female native. Concerning the intrinsic vowel fundamental frequency (IFO), it turned out that the effect of IF0 is fairly clearly present in some Hungarian learners when they are producing *Yangping* tone, mainly appearing in the second half part of their average F0

contours. At this point, our findings tend to support to Zhu Xiaoong's claim that the IF0 in high areas of the individual's own frequency domain is more distinct than that in low areas.

(3) Summarizing the findings about pitch range we can say that 40% of these Hungarian university learners have a fairly appropriate range in *Yangping*, that is, their ranges are quite close to those of native speakers. However, 60% of them (two female and four male in the sample) can be considered deviant in the sense that they produce much wider or narrower ranges.

5.3. *Shangsheng*

The third tone, called *Shǎngshēng*, begins at the half-low point 2 and falls to the lowest level 1, then rises to the half-high level 4. The value of the tone in the Chao system is therefore 214. Feng Zhichun in *Xiandai Hanyu* defines *Shangsheng* as a fall-rise tone with a pattern running from quite low to the lowest level of the scale, then rising to half high level[186]. (2008, 100)

Zhu Xiaonong, Zhang Ting and Yi li indicate that *Shangsheng* should be categorized as a "dipping" tone[187].

The main aim of this section is to present the voice samples of the third tone, *Shangsheng*, including the measurement, extraction, analysis and comparison. The recordings are examined in terms of the duration, the fundamental frequency (F0), the tone range and the tone value of the monosyllabic tone produced by the recorded subjects, thus creating a detailed profile of the *Shangsheng* pronunciation of this sample of Hungarian university students (four male, labeled M1, M2 etc., two female, labeled F1, F2 etc.), juxtaposed with the pronunciation of one male and one female native speaker of Chinese (labeled BM and BF respectively). The actual methodology used in the measurement of duration, F0 and tone range was detailed in 4.3 above; note that the initial or "onset" of each syllable was discarded; the 11 measuring points all fall within the nucleus or rime. Thus the word "start" used in the comments about

[186] Feng 2008: 100.

[187] Zhu, Zhang and Yi 2012.

the subjects' performance refers to the beginning of the nucleus, and not to the beginning of the whole syllable.

As we mentioned before, creaky voice often occurs during the production of *Shangsheng*. Because of the thick stiff vocal fold, the F0 range of a sound produced with creaky voice is extremely low, sometimes only 40-50Hz or even lower. Moreover, the vibration of vocal folds is irregular. Due to the extremely low and irregular F0, the F0 of a sound pronounced with creaky voice cannot easily be measured, and as Zhu Xiaonong points out this usually results in the appearance of a broken F0 contour[188]. Of the 12 subjects involved in this study, there are 5 subjects –three female and two male - who frequently pronounce *Shangsheng* with creaky voice, with the result that the F0 of their *Shangsheng* syllables could not be satisfactorily measured. Since determining the duration is still possible, we will use all 12 subjects' data for analyzing durations. However, probably only the 7 subjects who produced *Shangsheng* syllables without creaky voice accompanied are worth analyzing for F0 and range. These include two female subjects: F1, F4; and 5 male subjects: M1, M3, M4, M5, BM.

The 16 *Shangsheng* syllables used in the recordings, as explained in 4.3 above, fit into the following matrix:

Table 63: *Matrix of Shangsheng syllables*

	ǎ	ǐ	ǔ
b	把	笔、比	哺、捕、补
d	沓、打	底、抵	堵、赌
Zero-initials		已、以	五、午

5.3.1. Duration

The method used for determining the duration of the syllable has been presented before. The following table shows the durations of each *Shangsheng* syllable pronounced by 12 subjects, as well as average durations. The unit of duration is the millisecond (ms).

[188] Zhu 2005: 93.

Table 64: *Duration of 16 different Shangsheng syllables pronounced by 12 different subjects.*

Syllables	F1	F2	F3	F4	M1	M2	M3	M4	M5	M6	BF	BM
哺 bǔ	383	613	577	699	380	440	332	483	675	298	480	419
沓 dǎ	419	572	634	690	449	475	304	508	739	358	580	500
笔 bǐ	464	573	629	710	441	448	330	401	786	281	585	584
底 dǐ	476	597	667	739	448	435	304	557	772	265	664	591
捕 bǔ	427	575	613	740	383	441	301	492	766	285	622	546
抵 dǐ	481	573	611	632	433	420	290	464	792	329	632	589
五 wǔ	535	616	662	652	463	394	317	520	662	299	690	675
打 dǎ	435	514	594	719	425	498	275	507	701	347	648	586
已 yǐ	390	538	607	708	398	518	261	477	762	320	674	672
堵 dǔ	445	533	686	731	400	441	308	455	850	282	668	650
比 bǐ	459	564	583	674	406	450	309	391	654	335	681	609
把 bǎ	426	478	617	671	383	460	266	440	648	239	696	623
赌 dǔ	445	542	568	711	409	473	305	510	654	334	649	625
午 wǔ	427	660	589	640	476	543	357	498	739	312	657	658
补 bǔ	459	590	710	696	459	458	309	467	686	331	619	608
以 yǐ	410	646	691	737	442	508	244	528	654	341	712	642
Average	445	574	627	697	426	463	301	481	721	310	641	599

On the basis of the averages displayed in the bottom row of the above table, we made a chart for the further comparison.

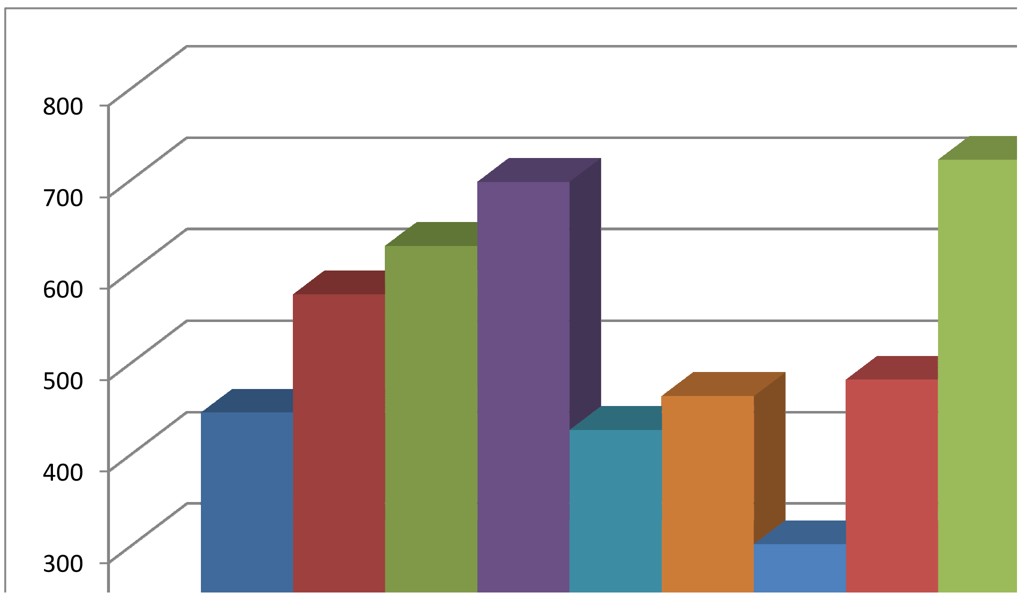


Figure 84. Average duration of *Shangsheng* syllables pronounced by 12 different subjects

The chart clearly shows that the average duration of M5 is the longest, while the average duration of M3 is the shortest, at less than half of M5's. The average durations of BF and BM are 641ms, 599ms, respectively. Only F2's, F3's and F4's average durations are close to BF's and BM's. The average durations of F1, M1, M2, M3, M4 and M6 are considerably shorter than BF's and BM's, especially M3's and M6's, while the average duration of M5 is considerably longer than "model" native speakers'. In other words, 70% Hungarian university learners in this sample – one female and six male subjects - differ considerably from the "model" native speaker in terms of the duration of *Shangsheng* syllables, most in producing much shorter durations. Compared to female subjects, male subjects are much more likely to cause problems in producing appropriate durations of *Shangsheng* syllables.

While, as we have seen, producing a tone too lengthily or too briefly may cause problems in duration, it is probably not the only reason. The inappropriate pronounciation of the vowel which is bearing a tone can also lead to duration problems. Generally speaking, the duration of a given vowel is associated with its height, and each vowel has an intrinsic vowel duration (referred to as IVD)[189]. Catford (1977, 45) believes that the duration of a high vowel is longer than that of a low one. His statement is described as IVD "i>a"[190].

The following table shows the average durations (m) of *Shangsheng* produced by each of the 12 subjects, in terms of different vowels /i/, /u/ and /a/ which are bearing the *Shangsheng* tone, as well as the average durations of all female subjects (F) and male subjects (M). The unit of duration is the millisecond (ms), and "n" refers to the number of sample syllables. The last column in the table shows the subject's "duration order" of /i/, /u/ and /a/, with the longest vowel first and the shortest vowel last. For example, F1's average durations of *Shangsheng* tone which are contained by vowel /i/, /u/ and /a/ are 447ms, 446ms and 427ms, respectively. The "duration order" of /i/, /u/ and /a/ is /a/>/i/>/u/.

[189] Zhu 2005: 72.
[190] Catford 1977: 45.

Table 65: *Average durations (m) of vowels /i/, /u/ and /a/ with Shangsheng tone produced by 12 subjects*

	/i/		/u/		/a/		the order
	m	n	m	n	m	n	
F1	447	7	446	6	427	3	/i/>/u/>/a/
F2	582	7	590	6	521	3	/u/>//i/>a/
F3	631	7	629	6	615	3	/i/>/u/>/a/
F4	700	7	696	6	693	3	/i/>/u/>/a/
M1	428	7	424	6	419	3	/i/>/u/>/a/
M2	463	7	456	6	478	3	/a/>/i/>/u/
M3	290	7	318	6	282	3	/u/>//i/>a/
M4	470	7	489	6	485	3	/u/>/a/>/i/
M5	737	7	719	6	696	3	/i/>/u/>/a/
M6	312	7	306	6	315	3	/a/>/i/>/u/
F	590	7	590	6	564	3	/i/=/u/>/a/
M	496	7	452	6	446	3	/i/>/u/>/a/
FB	658	7	633	6	626	3	/i/>/u/>/a/
MB	615	7	597	6	570	3	/i/>/u/>/a/

From the above table, we can see that: (1) FB's average durations of /i/, /u/ and /a/ are 658ms, 633ms and 626ms, and MB's durations of these three vowels are 615ms, 597ms and 570ms, respectively. In other words, both FB's and MB's duration orders of these three vowels are /i/>/u/>/a/, which corresponds to the "i>a" principle. (2) Many subjects: F1, F3, F4, M3 and M4, produced duration orders of /i/, /u/ and /a/ that are the same as the "model" native speakers', which is /i/>/u/>/a/. (3) The duration order of /i/, /u/ and /a/ produced by M2 is /u/>/i/>/a/. Since both /i/ and /u/ are high vowel, the order in /u/>/i/>/a/ makes sense. (4) 30% of Hungarian subjects produce deviant orders of durations of /i/, /u/ and /a/. M2's and M6's are in the order, which is /a/>/i/>/u/, and M4's in /u/>/a/>/i/. (4) The average durations of /i/, /u/ and /a/ produced by all the Hungarian male subjects are sorted by /i/>/u/>/a/, which is identical to the "model" native speakers', and the average duration order of /i/, /u/ and /a/

produced by all the Hungarian female subjects is /i/=/u/>/a/. Since both /i/ and /u/ are high vowels, therefore F's and M's duration order of vowels are consistent with the "i>a" principle, that is, /i/>/a/. All in all, the deviant orders of durations of /i/, /u/ and /a/ indicate that the inappropriately produced vowels are likely to cause problems in duration.

It is worth mentioning that the average durations of /i/, /u/ and /a/ produced by all the Hungarian female subjects (F) and male subjects (M) are 581ms [(590＋590＋564) ÷ 3], 465ms [(496＋452＋446) ÷ 3], respectively. In other words, although the sample is a small one, there probably is a connection between duration and gender. This conclusion is in contrast with Zhu Xiaonong's view that there is no connection between duration and gender[191].

However, as we have seen duration is a relative concept. It is hard to define a standard or universally "desirable" value of duration, because it depends so strongly on the individual's general rate of speech, not to mention the nature of each specific utterance and the conditions under which it is made. Discussing the duration of specific syllables within the parameters of an individual's own tone system probably makes better sense. This issue will be discussed later in this chapter.

5.3.2. Fundamental frequency

In this section, firstly, each subject's F0 values for every *Shangsheng* monosyllable, together with the means for all the syllables, are presented in a combined table. Secondly, each subject's F0 contours for all the *Shangsheng* monosyllables are plotted in one chart for further analysis and comparison. In each chart the vertical (Y) axis shows the value of fundamental frequency, with Hz as the basic unit of measurement. The horizontal (X) axis presents the 11 measuring points. For better comparison purposes, the vertical axis covers a fixed value range based on the maximum and minimum frequencies recorded by female and male subjects respectively. For female subjects, from 100Hz to 400Hz; for male subjects, from 50Hz to 250Hz. Thirdly, the average contour of all the subject's *Shangsheng* syllables is plotted in another chart. (For details of the methods used for measuring fundamental frequency (F0), see 3.2.1)

[191] Zhu 2005: 71.

5.3.2.1. Individual subjects' F0 data

The F0 data table and charts of every subject will now be provided.

(1) F1

The following table shows the F0 data for subject F1, including 16 *Shangsheng* monosyllables. Each monosyllable has 11 F0 values which correspond to the 11 measuring points. The means of the measurements made at each point are presented in the final row of the table.

Table 66: *F1's F0 values for 16 Shangsheng syllables at 11 measuring points*

Syllables	1	2	3	4	5	6	7	8	9	10	11
哺 bǔ	227	218	205	191	182	179	181	191	206	225	238
沓 dǎ	210	187	175	166	165	164	166	172	191	200	214
笔 bǐ	187	184	173	167	165	162	165	176	192	203	211
底 dǐ	182	186	174	167	165	162	165	176	192	204	212
捕 bǔ	210	217	207	188	173	171	173	184	202	218	222
抵 dǐ	182	185	174	167	165	162	165	182	193	205	212
五 wǔ	198	183	174	170	165	163	166	171	184	204	212
打 dǎ	211	187	174	166	165	164	166	177	192	203	214
已 yǐ	195	186	180	178	178	190	194	198	207	216	225
堵 dǔ	194	179	169	165	162	163	181	173	189	203	221
比 bǐ	217	197	186	177	176	174	176	185	201	210	218
把 bǎ	205	194	177	166	158	152	154	171	186	201	213
赌 dǔ	194	179	169	165	162	163	181	173	189	203	221
午 wǔ	208	217	208	189	173	171	173	184	202	218	222
补 bǔ	220	197	186	178	176	174	176	185	202	211	219
以 yǐ	216	215	203	185	172	171	173	184	201	217	222
Average	202.7	193.1	182.1	173.3	168.7	167.6	172.1	179.9	195.2	208.3	218.3

On the basis of the above table, we made the following two F0 contour charts of F1's *Shangsheng* pronunciation. The first chart shows the contour for every monosyllable, the second shows the mean of all the contours. The two charts are followed by brief comments on the data.

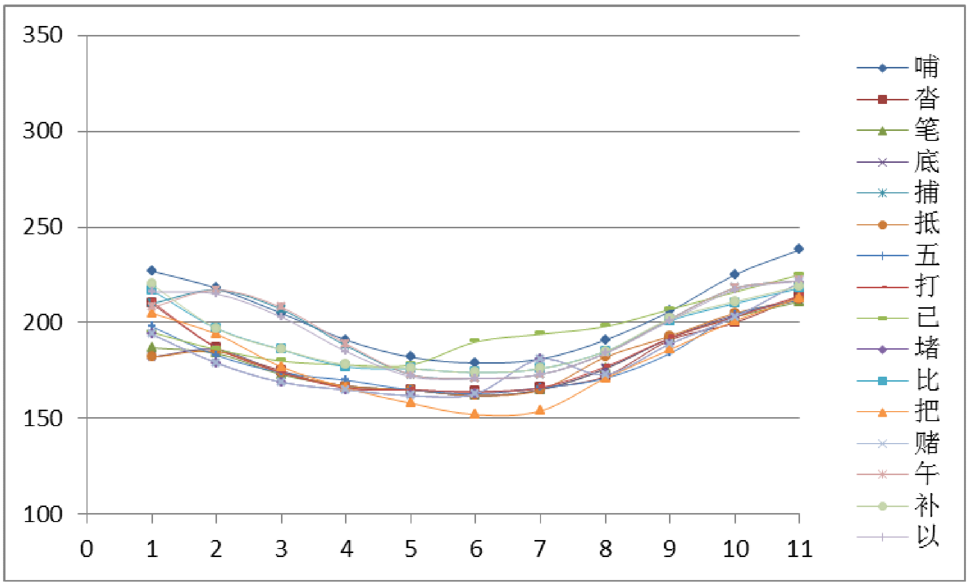


Figure 85. The F0 contour of 16 *Shangsheng* syllables pronounced by F1

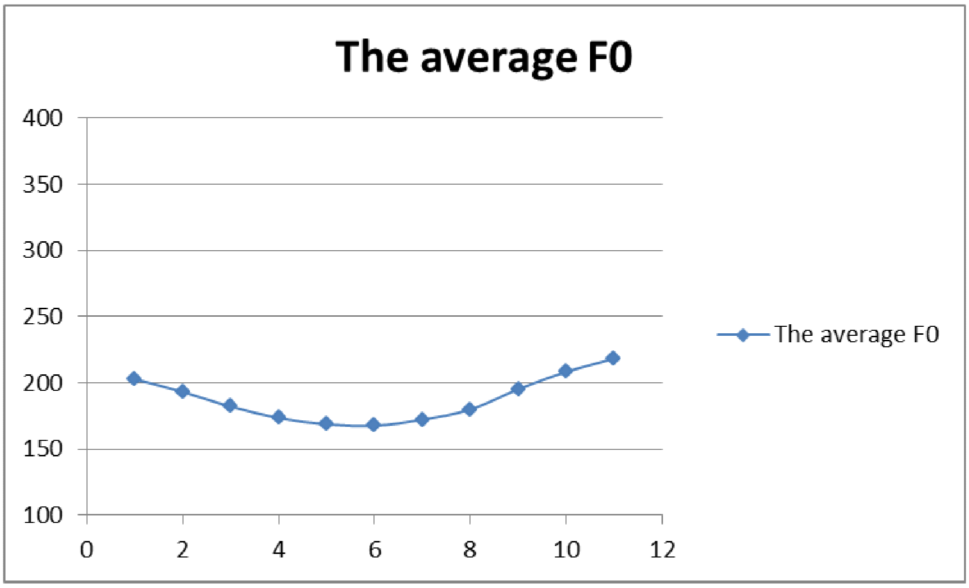


Figure 86. The average F0 contour of 16 *Shangsheng* syllables pronounced by F1

Based on Chart 1, the following conclusions can be made. (1) The F0 range of F1 is 150-240Hz, that is, about 90Hz. (2) The F0 contours of F1 are relatively concentrated. (3) Most F0 contours have a clear falling start, and the rising parts are relatively gentle. (4) The F0 contours of F1 turn out as fall-rise patterns. (5) Most F0 contours begin rising after the first half of the contour.

From Chart 2, we can see that the mean F0 contour of F1 displays a fall-rise pattern with a quite gentle radian and that the rising part does not begin until the seventh measuring point, in other words, at a point about 65% through the whole F0 contour. The value of the mean F0 contour ranges between 160 and 220Hz.

(2) F4

Table 67: *F4's F0 values for 16 Shangsheng syllables at 11 measuring points*

Syllables	1	2	3	4	5	6	7	8	9	10	11
哺 bǔ	301	279	230	204	191	190	191	197	218	273	328
沓 dǎ	295	241	207	189	179	179	177	191	214	254	320
笔 bǐ	301	265	218	198	189	185	184	192	212	248	317
底 dǐ	275	246	206	193	187	183	188	189	205	242	313
捕 bǔ	275	245	211	190	183	181	181	188	208	271	322
抵 dǐ	288	258	218	191	184	178	181	183	203	233	271
五 wǔ	291	248	200	182	175	175	177	186	210	263	329
打 dǎ	283	236	195	183	178	176	176	182	190	212	274
已 yǐ	294	257	211	183	174	173	174	186	214	264	327
堵 dǔ	325	240	195	179	181	182	181	194	224	269	337
比 bǐ	272	234	195	181	176	179	179	186	204	250	303
把 bǎ	255	205	183	171	173	170	178	186	213	253	302
赌 dǔ	288	241	185	175	171	171	172	179	196	262	332
午 wǔ	286	273	230	186	179	175	181	191	214	252	291
补 bǔ	270	235	192	179	177	176	176	179	194	249	304
以 yǐ	291	280	222	188	177	173	176	184	213	251	322
Average	286.9	248.9	206.1	185.8	179.6	177.9	179.5	187.1	208.3	252.9	312

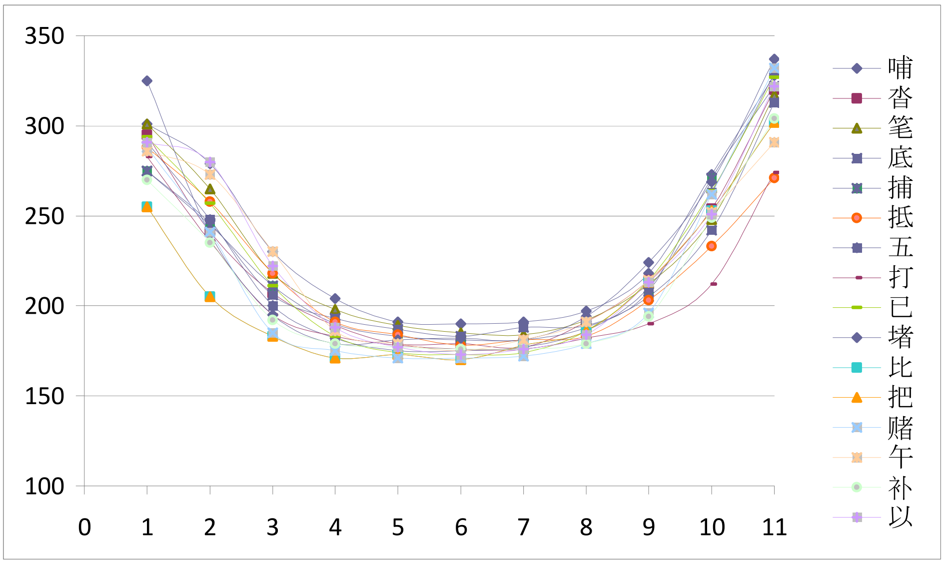


Figure 87. The F0 contour of 16 *Shangsheng* syllables pronounced by F4

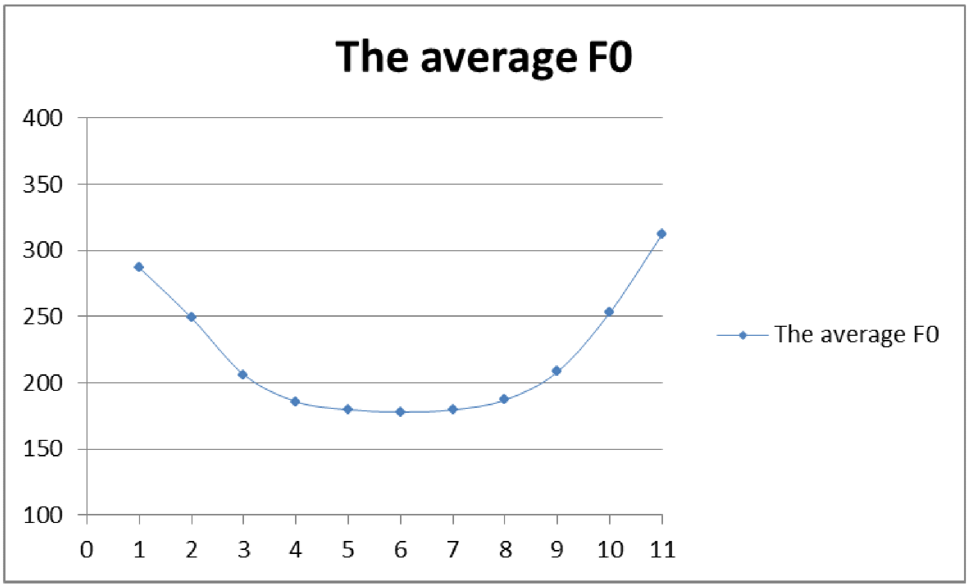


Figure 88. the average F0 contour of 16 *Shangsheng* syllables pronounced by F4

Based on Chart 1, the following conclusions can be made. (1) The F0 range of F4 is 150-350Hz, that is, about 200Hz. (2) The F0 contours of F4 are quite concentrated. (3) Most F0 contours have a very steep falling start, and the rising parts are also very steep. (4) The F0 contours of F4 turn out as fall-rise patterns. (5) Most F0 contours change direction and start rising after the first half of the contour.

From Chart 2, we can see that the mean F0 contour of F4 displays a fall-rise pattern with an disproportionately large radian and that the rising part does not begin until the seventh measuring point, in other words, at a point about 65% through the whole F0 contour. The value of the mean F0 contour ranges between 170 and 320Hz.

(3) M1

Table 68: *M1's F0 values for 16 Shangsheng syllables at 11 measuring points*

Syllables	1	2	3	4	5	6	7	8	9	10	11
哺 bǔ	114	111	105	107	106	105	108	112	122	124	150
沓 dǎ	111	105	105	105	103	103	104	107	112	121	134
笔 bǐ	115	111	100	103	105	106	108	115	116	126	138
底 dǐ	119	109	100	103	105	106	108	115	116	122	138
捕 bǔ	111	107	95	89	96	96	104	112	113	122	133
抵 dǐ	115	111	100	103	105	105	107	115	116	121	138
五 wǔ	111	105	105	105	103	103	105	108	115	123	137
打 dǎ	117	107	101	100	101	99	105	107	110	117	137
已 yǐ	114	105	105	105	103	103	104	107	111	120	132
堵 dǔ	110	109	105	97	96	97	103	112	121	125	139
比 bǐ	119	110	100	103	105	104	106	113	117	118	135
把 bǎ	112	106	94	89	97	96	104	112	114	122	134
赌 dǔ	125	112	108	99	95	99	99	104	115	129	147
午 wǔ	107	107	101	99	96	100	101	110	118	127	137
补 bǔ	111	105	105	105	103	103	105	108	114	123	137
以 yǐ	123	111	107	109	102	106	109	109	116	126	146
Average	114.6	108.2	102.3	101.3	101.3	101.9	105	110.4	115.4	122.9	138.3

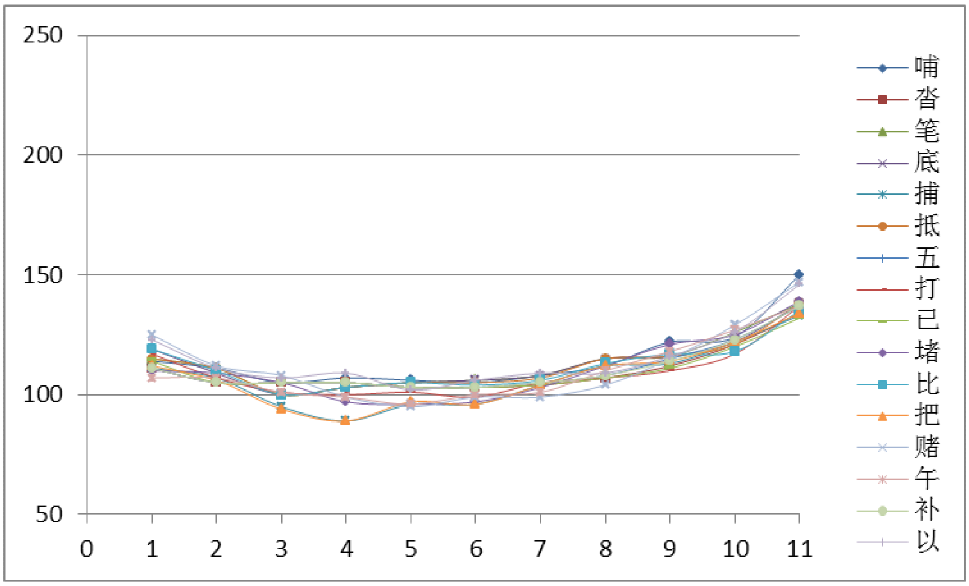


Figure 89. The F0 contour of 16 *Shangsheng* syllables pronounced by M1

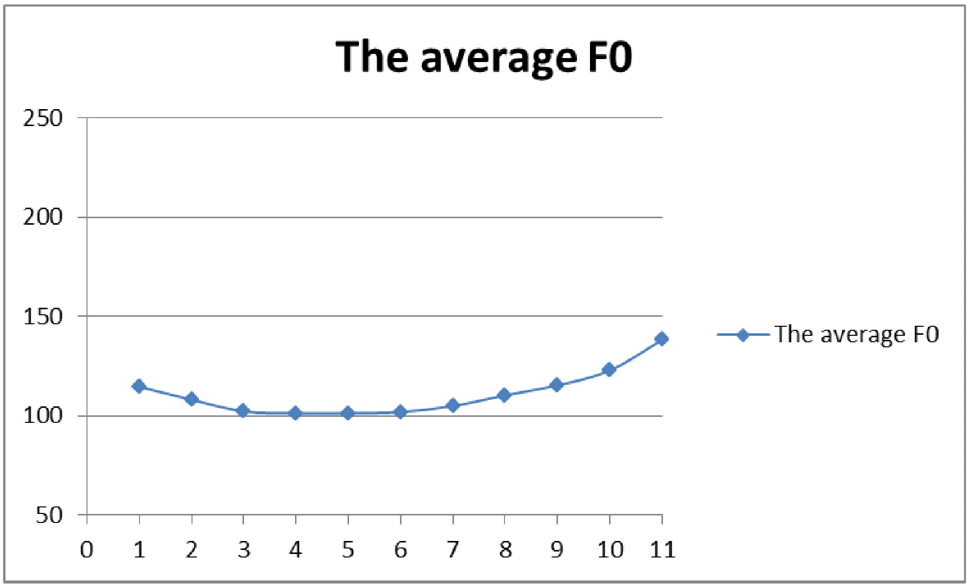


Figure 90. The average F0 contour of 16 *Shangsheng* syllables pronounced by M1

Based on Chart 1, the following conclusions can be made. (1) The F0 range of M1 is 80-150Hz, that is, about 70Hz. (2) The F0 contours of M1 are highly concentrated. (3) Most F0 contours have a quite gentle falling start, and the rising parts are also relatively gentle. (4) The F0 contours of F4 turn out as fall-rise patterns. (5) Most F0 contours change direction and start rising at the middle of the contour.

From Chart 2, we can see that the mean F0 contour of F1 displays a fall-rise pattern with a quite gentle radian and that the rising part does not begin until the sixth measuring point, in other words, at a point about 55% through the whole F0 contour. The value of the mean F0 contour ranges between 100 and 140Hz.

(4) M3

Table 69: *M3's F0 values for 16 Shangsheng syllables at 11 measuring points*

Syllables	1	2	3	4	5	6	7	8	9	10	11
哺 bǔ	129	118	115	114	118	119	124	148	185	221	237
沓 dǎ	117	114	109	106	104	105	108	119	139	168	185
笔 bǐ	112	109	111	112	113	115	122	140	164	190	201
底 dǐ	114	109	105	107	111	113	112	115	128	151	178
捕 bǔ	113	111	111	112	111	111	115	120	137	165	187
抵 dǐ	114	112	110	112	113	115	124	139	159	183	197
五 wǔ	108	107	107	110	111	113	121	138	168	198	217
打 dǎ	112	110	110	108	107	108	113	124	142	165	179
已 yǐ	116	118	118	119	118	121	131	149	171	191	200
堵 dǔ	119	114	114	113	114	115	121	137	163	191	210
比 bǐ	117	115	112	111	112	112	113	120	139	157	182
把 bǎ	111	110	109	107	108	110	111	116	129	147	166
赌 dǔ	115	111	111	110	108	108	113	129	155	182	207
午 wǔ	117	110	111	114	114	116	118	133	159	189	208
补 bǔ	117	115	115	115	113	112	114	117	133	160	189
以 yǐ	109	110	110	111	116	122	136	151	169	184	190
Average	115	112.1	111.1	111.3	111.9	113.4	118.5	130.9	152.5	177.6	195.8

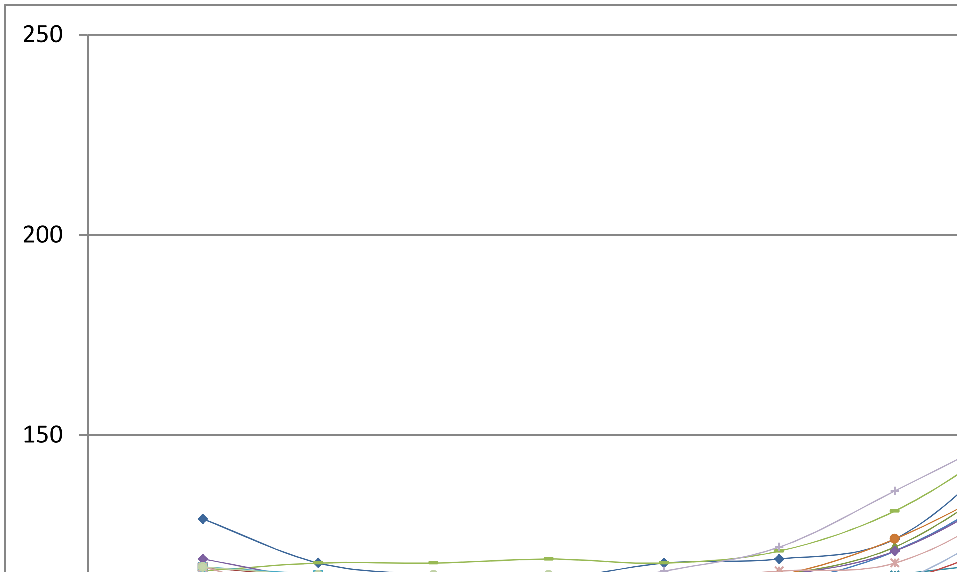


Figure 91. The F0 contour of 16 *Shangsheng* syllables pronounced by M3

Figure 92. The average F0 contour of 16 *Shangsheng* syllables pronounced by M3

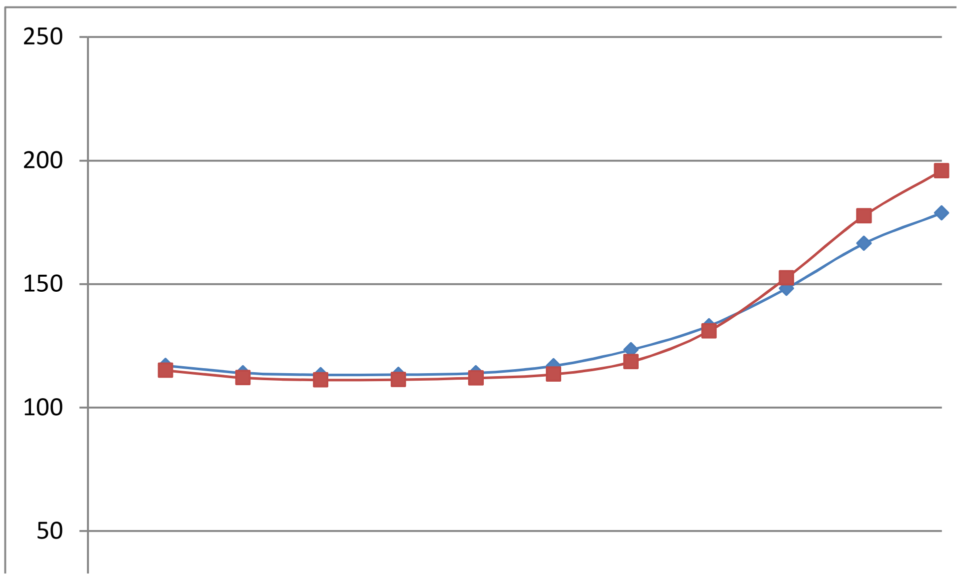


Figure 93. The average F0 contours of *Yangping* and *Shangsheng* syllables pronounced by M4

Based on Chart 1, the following conclusions can be made. (1) The F0 range of M3 is 100-250Hz, that is, about 150Hz. (2) The F0 contours of M3 are highly concentrated. (3) Most F0 contours have a quite flat start and the rising parts are noticeably steep. (4) The F0 contours of M3 turn out as rising patterns. (5) Most F0 contours change direction and start rising after the first half of the contour.

From Chart 2, we can see that the mean F0 contour of M3 displays a rising pattern with a flat start and that the rising part does not begin until the fifth measuring point, in other words, at a point about 45% through the whole F0 contour. The value of the mean F0 contour ranges between 110 and 200Hz. For the purpose of comparison, we plotted M3's average F0 contours for *Yangping* and *Shangsheng* together on the same Chart (above). The result clearly shows that two F0 contours are very similar, especially the first halves, which are almost identical. In other words, the *Shangsheng* syllables pronounced by M3 appear as *Yangping*.

(5) M4

Table 70: *M4's F0 values for 16 Shangsheng syllables at 11 measuring points*

Syllables	1	2	3	4	5	6	7	8	9	10	11
哺 bǔ	133	128	122	118	112	108	107	111	118	128	133
沓 dǎ	127	118	113	109	106	100	98	104	117	127	133

笔 bǐ	123	123	120	118	114	111	109	110	116	125	135
底 dǐ	127	120	118	114	109	105	103	106	111	117	130
捕 bǔ	118	118	114	111	106	106	108	110	112	117	137
抵 dǐ	123	119	116	111	106	103	104	108	117	131	141
五 wǔ	122	117	118	116	111	102	99	103	114	123	130
打 dǎ	120	117	111	107	101	100	104	108	114	123	130
已 yǐ	129	121	119	114	107	102	101	103	115	133	145
堵 dǔ	126	125	122	117	111	109	109	114	125	136	154
比 bǐ	117	115	113	111	108	105	104	107	109	117	126
把 bǎ	127	122	117	110	104	102	98	98	104	115	127
赌 dǔ	124	123	120	114	107	104	105	112	118	124	128
午 wǔ	125	124	124	122	114	108	103	104	106	113	127
补 bǔ	124	122	119	114	107	104	104	108	118	129	138
以 yǐ	121	121	120	117	110	106	105	107	117	136	146
Average	124.1	120.8	117.9	113.9	108.3	104.7	103.8	107.1	114.4	124.6	135

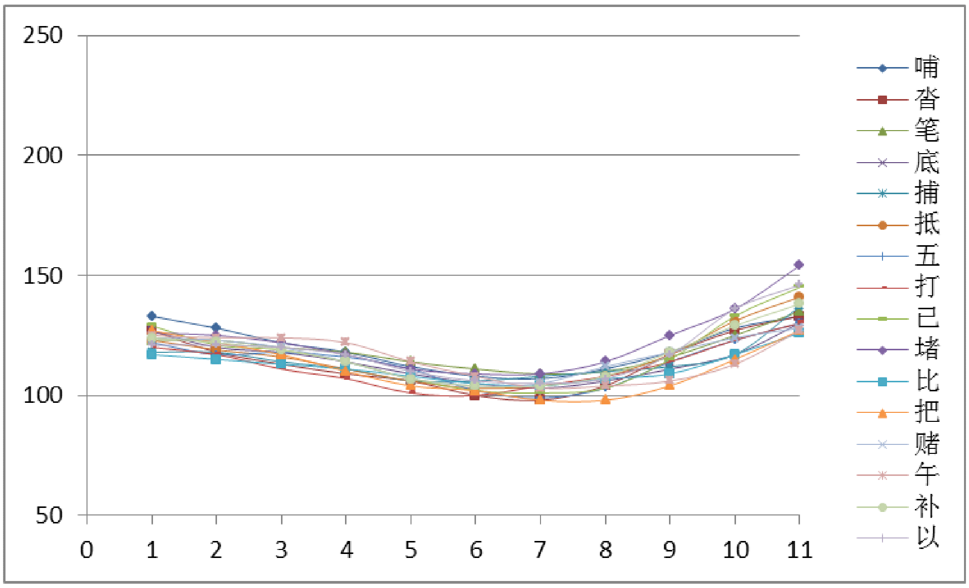


Figure 94. The F0 contour of 16 *Shangsheng* syllables pronounced by M4

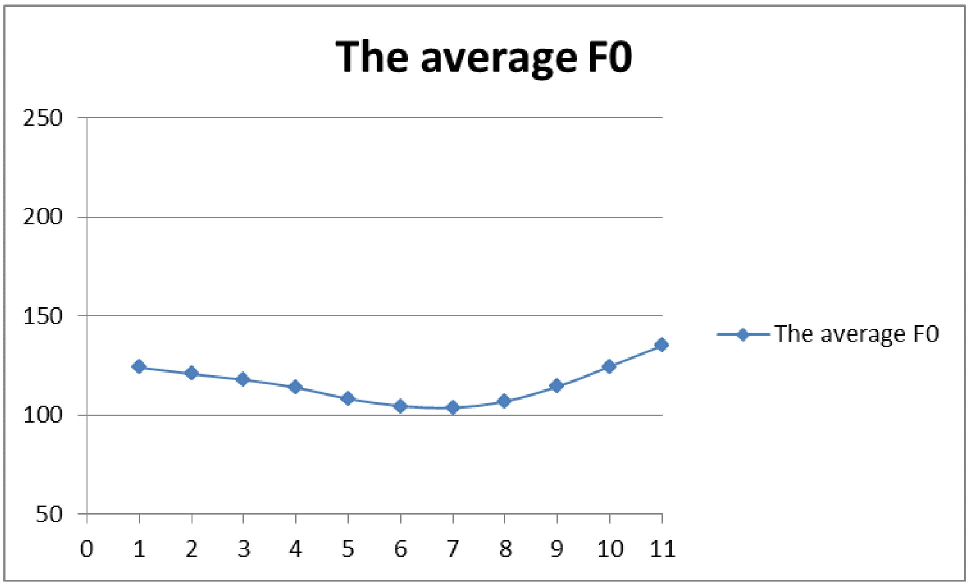


Figure 95. The average F0 contour of 16 *Shangsheng* syllables pronounced by M4

Based on Chart 1, the following conclusions can be made. (1) The F0 range of M4 is 90-160Hz, that is, about 70Hz. (2) The F0 contours of M4 are highly concentrated. (3) Most F0 contours have a clear falling start and the rising parts are relatively gentle. (4) The F0 contours of M4 turn out as fall-rise patterns. (5) Most F0 contours change direction and start rising after the first half of the contour.

From Chart 2, we can see that the mean F0 contour of M4 displays a fall-rise pattern with a relatively gentle radian and that the rising part does not begin until the eighth measuring point, in other words, at a point about 75% through the whole F0 contour. The value of the mean F0 contour ranges between 100 and 140Hz.

(6) M5

Table 71: *M5's F0 values for 16 Shangsheng syllables at 11 measuring points*

Syllables	1	2	3	4	5	6	7	8	9	10	11
哺 bǔ	108	101	97	96	94	94	94	95	102	123	147
沓 dǎ	108	100	97	96	96	95	97	101	109	128	146
笔 bǐ	113	102	100	98	97	97	98	100	111	130	152
底 dǐ	105	101	98	97	97	98	98	103	116	141	159
捕 bǔ	116	104	100	100	100	99	100	101	106	120	143
抵 dǐ	112	103	101	99	99	98	98	100	105	116	140
五 wǔ	110	109	107	104	101	100	99	99	102	120	141

打 dǎ	113	103	100	97	96	95	95	97	100	113	135
已 yǐ	122	108	103	102	101	101	102	108	123	146	170
堵 dǔ	100	101	99	97	98	99	103	116	135	159	170
比 bǐ	120	106	102	100	98	98	98	99	104	117	139
把 bǎ	108	101	99	97	96	96	96	98	102	115	122
赌 dǔ	109	103	100	99	99	98	100	103	117	133	155
午 wǔ	113	107	103	102	99	98	98	97	100	111	134
补 bǔ	110	103	100	98	98	98	98	101	111	128	152
以 yǐ	126	106	102	99	99	97	98	98	104	118	141
Average	112.1	103.6	100.5	98.8	98	97.6	98.3	101	109.2	126.1	146.6

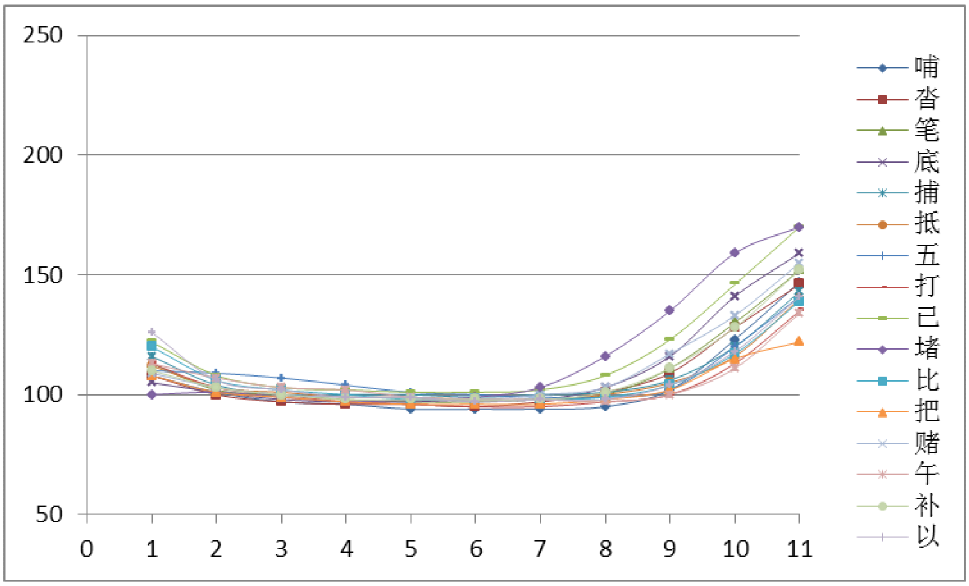


Figure 96. The F0 contour of 16 *Shangsheng* syllables pronounced by M5

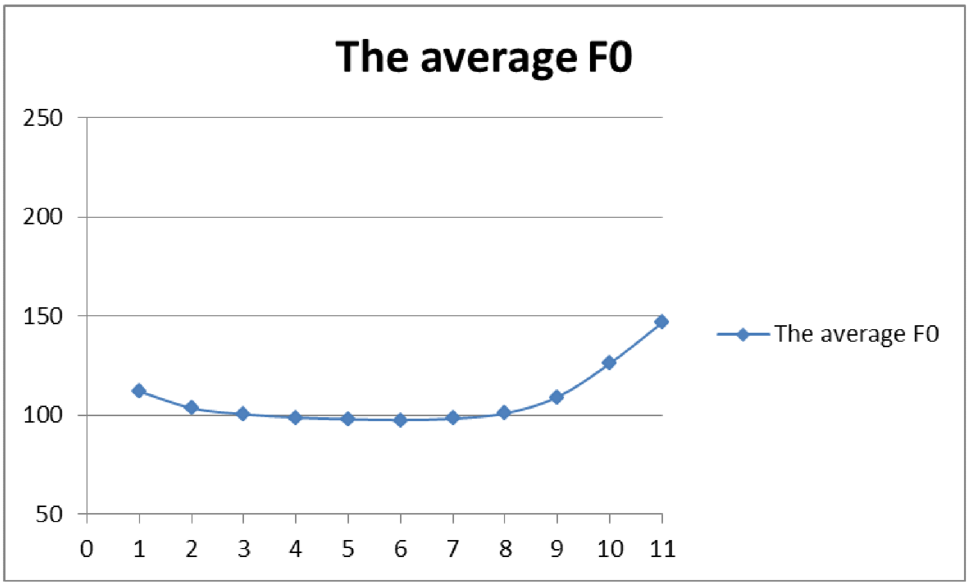


Figure 97. The average F0 contour of 16 *Shangsheng* syllables pronounced by M5

Based on Chart 1, the following conclusions can be made. (1) The F0 range of M5 is 90-170Hz, that is, about 80Hz. (2) The F0 contours of M5 are highly concentrated. (3) Most F0 contours have a clear falling start and the rising parts are quite steep. (4) The F0 contours of M3 turn out as fall-rise patterns. (5) Most F0 contours change direction and start rising after the first half of the contour.

From Chart 2, we can see that the mean F0 contour of M5 displays a rising pattern with a very gentle radian and the rising part does not begin until the seventh measuring point, in other words, at a point about 65% through the whole F0 contour. The F0 contour contains a quite flat section in the middle. The value of the mean F0 contour ranges between 90 and 150Hz.

(7) BM

Table 72: *BM's F0 values for 16 Shangsheng syllables at 11 measuring points*

Syllables	1	2	3	4	5	6	7	8	9	10	11
哺 bǔ	97	91	86	83	83	83	84	85	88	93	95
沓 dǎ	94	89	83	82	82	84	85	86	89	90	98
笔 bǐ	102	92	91	90	90	91	93	95	97	105	113
底 dǐ	104	97	91	93	93	94	96	99	103	110	117
捕 bǔ	103	97	91	90	93	94	95	96	101	108	109
抵 dǐ	100	88	86	90	91	91	91	92	97	104	107
五 wǔ	99	98	92	90	88	88	89	94	95	99	108

打 dǎ	97	90	86	84	85	87	88	91	98	101	100
已 yǐ	104	101	94	89	86	84	86	89	95	100	102
堵 dǔ	108	93	86	87	85	86	86	88	93	97	98
比 bǐ	101	93	87	88	88	88	88	91	95	96	99
把 bǎ	105	93	86	86	86	87	87	89	94	97	99
赌 dǔ	103	92	86	86	86	87	87	89	93	97	99
午 wǔ	98	93	90	88	88	88	92	96	99	106	111
补 bǔ	95	90	86	87	86	88	90	94	103	108	108
以 yǐ	110	94	85	83	84	85	86	89	100	109	113
Average	101.3	93.2	87.9	87.3	87.1	87.8	88.9	91.4	96.3	101.3	104.8

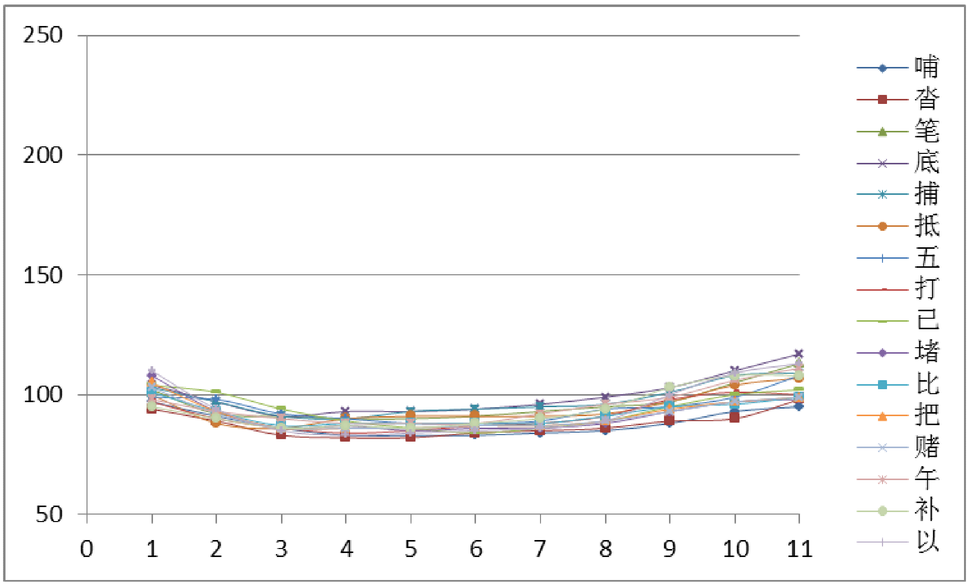


Figure 98. The F0 contour of 16 *Shangsheng* syllables pronounced by BM

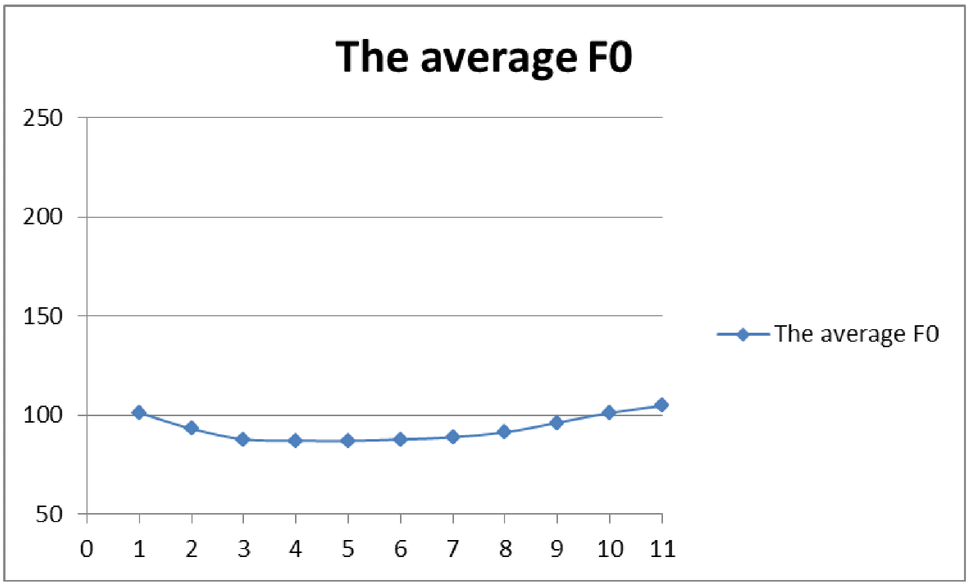


Figure 99. The average F0 contour of 16 *Shangsheng* syllables pronounced by BM

Based on Chart 1, the following conclusions can be made. (1) The F0 range of BM is 80-120Hz, that is, about 40Hz. (2) The F0 contours of M5 are highly concentrated. (3) Most F0 contours have a clear falling start and the rising parts are relatively gentle. (4) The F0 contours of BM turn out as fall-rise patterns. (5) Most F0 contours change direction and start rising after the first half of the contour.

From Chart 2, we can see that the mean F0 contour of BM displays a rising pattern with a very gentle radian and the rising part does not begin until the sixth measuring point, in other words, at a point about 55% through the whole F0 contour. The value of the mean F0 contour ranges between 80 and 110Hz.

5.3.2.2. General comments

From the tables and figures presented above, we can see that only about 33% of the Hungarian subjects (F1 and M1) pronounce *Shangsheng* syllables with an appropriate fall-rise contour. Around 67% of the subjects produce deviant contours, mainly due to an extremely sharp fall-rise pattern, an excessively steep, a rising pattern without a clear falling start, or because the rise begins at an inappropriate point .

In order to better compare the 6 Hungarian subjects with the 1 native speaker in general terms, we made a chart for the average F0 contours of *Shangsheng* syllables pronounced by all 7 subjects. The vertical axis covers a fixed value range which is from 0Hz to 350Hz.

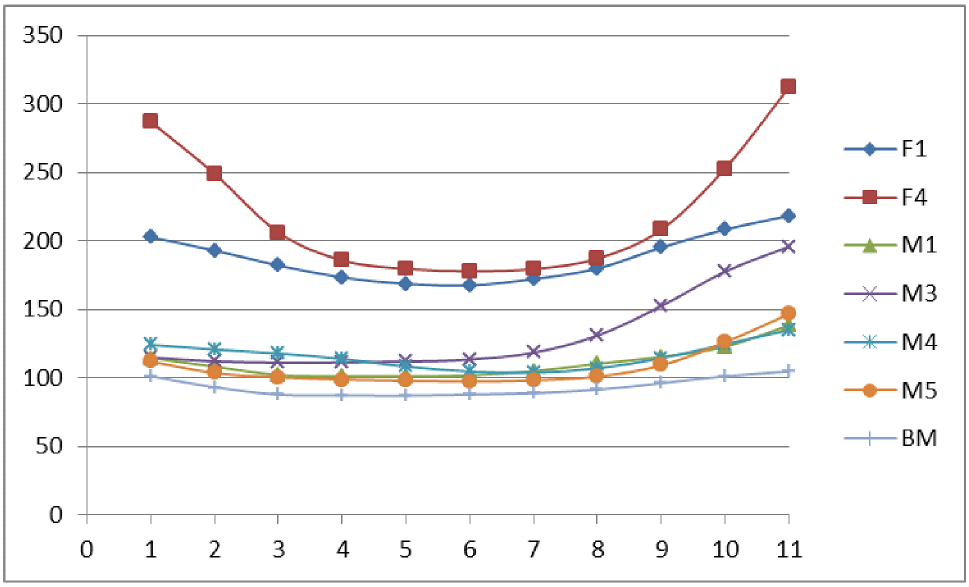


Figure 100. The average F0 contours of *Shangsheng* syllables pronounced by 7 subjects

On the basis of the average F0 contours, the following conclusions can be made:

(1) About 83% of the Hungarian university students have a fall-rise average F0 contour when pronouncing *Shangsheng* syllables;

(2) About 33% of the Hungarian subjects pronounce *Shangsheng* syllables correctly; in other words, their average F0 contours are appropriate fall-rise patterns;

(3) About 67% of the Hungarian subjects do not pronounce *Shangsheng* syllables appropriately, even though most of them do manage to pronounce *Shangsheng* syllables with a fall-rise pattern. The deviant contours are as follows: F4's F0 contour displays a fall-rise pattern, but has an inappropriately sharp falling start, and its rising part is very steep; M3 pronounces *Shangsheng* syllables with an excessively smooth start, therefore, the F0 contour appears as an rising pattern which is close to the same subject's *Yangping*; M4's F0 contour has a quite appropriate falling start and a quite appropriate rising part, but it starts to rise relatively late; M5's F0 contour has an appropriate start but follows a too steep rising part.

(4) The F0 values produced by females are considerably higher than males': about twice as high;

(5) Of the 12 Hungarian subjects, only 6 produced entire and regular F0 contours, two female

and 4 male, which indicates that creaky voice happens considerably often in *Shangsheng* in Hungarian university learners. This does not necessarily indicate a specifically non-native or Hungarian phenomenon – we note that one of our native-speaker subjects, FB, also used creaky voice. Clearly, the fact that BF's data were unusable at this point, owing to the creaky voice, reduces the significance of the learner/native F0 and range comparisons. Attempts to identify a suitable alternative native-speaker (one who did not have the creaky voice phenomenon) met with greater difficulties than were expected, as all the available candidates produced creaky voice. In the end, we decided to focus on MB, but we must acknowledge that more research into *shangsheng* fundamental frequency and range is required.

5.3.3. Intrinsic vowel fundamental frequency

The intrinsic vowel fundamental frequency, shortened to IF0, refers to the fact that, according to Lehiste (1970) the F0 of a high vowel is higher than that of a low one when the two vowels are produced under the same conditions[192]. Zhu Xiaonong believes that the IF0 in high areas of the individual's own frequency domain is more distinct than that in low areas, and the IF0 appears more obviously in female than in male speakers[193] .

For the purposes of defining the relationship between the height of the F0 contours and vowels, we classified all *Shangsheng* syllables produced by each subject in terms of /i/, /u/ and /a/, and averaged the F0 values of each category. The following table shows the relevant details for each of the subjects whose recorded data were used in this part of the study and the average F0 of the Hungarian female subjects and the Hungarian male subjects (labeled F, M, respectively).

Table 73: *Shangsheng F0 of each Hungarian female subject and Hungarian male subject, with averages for all female and all male subjects.*

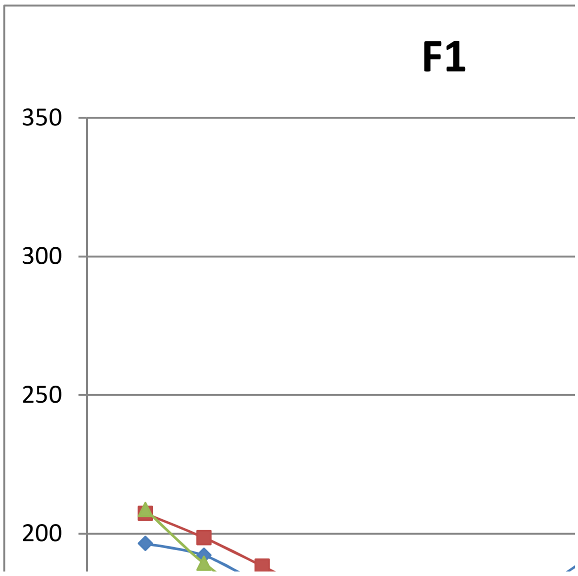
		1	2	3	4	5	6	7	8	9	10	11	ms
F1	/i/	196.5	192.1	181.7	173.5	170.2	170.2	173	183.5	197.7	209.2	216.7	447
	/u/	207.3	198.6	188.3	178	170.4	169.1	175.9	180.1	196.3	211.7	222.1	446
	/a/	208.7	189.3	175.3	166	162.7	160	162	173.3	189.7	201.3	213.7	427

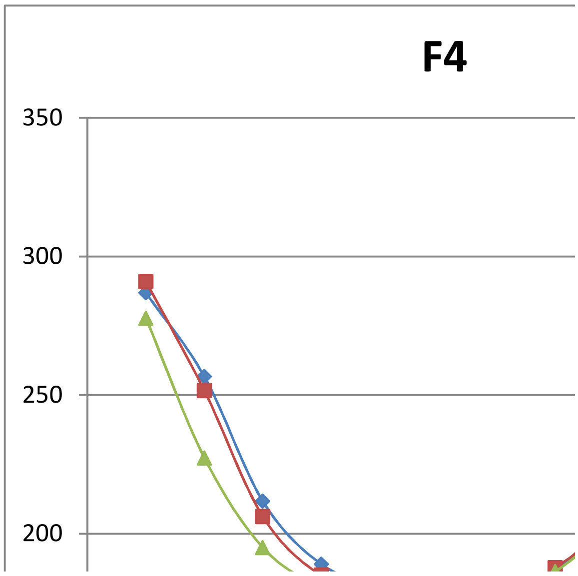
[192] Lehiste, Ilse 1970.

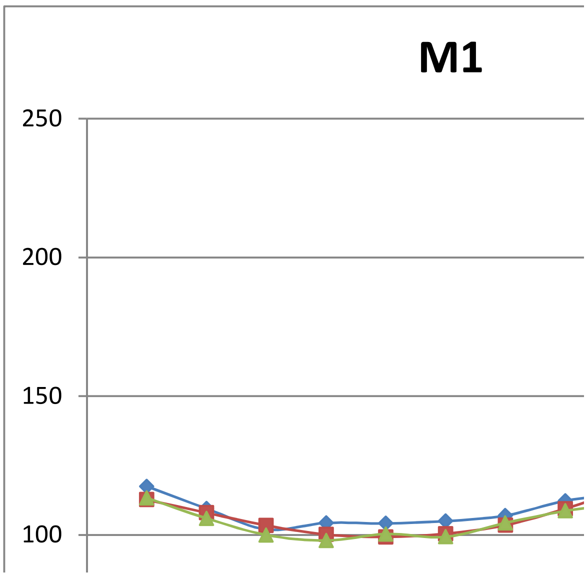
[193] Zhu 2005: 79.

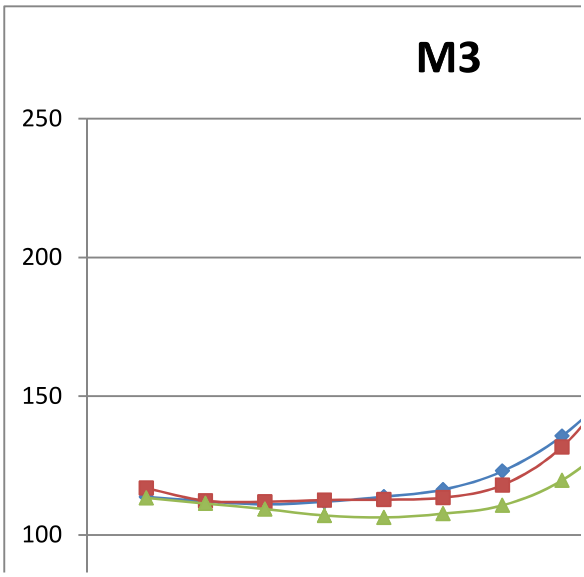
F4	/i/	286.8	256.7	211.7	189	181.2	178.5	180.3	186.7	208.5	248	308.8	700
	/u/	290.9	251.6	206.1	185	179.6	178.6	179.9	187.7	209.1	262.7	320.4	696
	/a/	277.7	227.3	195	181	176.7	175	177	186.3	205.7	239.7	298.7	693
M1	/i/	117.5	109.5	102	104.	104.2	105	107	112.3	115.3	122.2	137.8	428
	/u/	112.7	108	103.4	100.1	99.3	100.4	103.6	109.4	116.9	124.7	140	424
	/a/	113.3	106	100	98	100.3	99.33	104.3	108.7	112	120	135	419
M3	/i/	113.7	112.2	111	112	113.8	116.3	123	135.7	155	176	191.3	290
	/u/	116.9	112.3	112	112.6	112.7	113.4	118	131.7	157.1	186.6	207.9	318
	/a/	113.3	111.3	109.3	107	106.3	107.7	110.7	119.7	136.7	160	176.7	282
M4	/i/	123.3	119.8	117.7	114.2	109	105.3	104.3	106.8	114.2	126.5	137.2	470
	/u/	124.6	122.4	119.9	116	109.7	105.9	105	108.9	115.9	124.3	135.3	489
	/a/	124.7	119	113.7	108.7	103.7	100.7	100	103.3	111.7	121.7	130	485
M5	/i/	116.3	104.3	101	99.2	98.5	98.2	98.7	101.3	110.5	128	150.2	737
	/u/	109.4	104	100.9	99.4	98.4	98	98.9	101.7	110.4	127.7	148.9	719
	/a/	109.7	101.3	98.7	96.7	96	95.3	96	98.7	103.7	118.7	134.3	696
F	/i/	241.7	224.4	196.7	181.3	175.7	174.3	176.7	185.1	203.1	228.6	262.8	574
	/u/	249.1	225.1	197.2	181.5	175	173.9	177.9	183.9	202.7	237.2	271.3	571
	/a/	243.2	208.3	185.2	173.5	169.7	167.5	169.5	179.8	197.7	220.5	256.2	560
M	/i/	117.7	111.5	107.9	107.4	106.4	106.2	108.3	114	123.8	138.2	154.1	481
	/u/	115.9	111.7	109	107	105	104.4	106.4	112.9	125.1	140.8	158	488
	/a/	115.3	109.5	105.5	102.6	101.5	100.8	102.8	107.6	116	130.1	144	471
MB	/i/	103.5	94.2	89	88.8	88.7	88.8	90	92.5	97.8	104	108.5	615
	/u/	100.4	93.4	88.1	87.3	87	87.7	89	91.7	96	101.1	104	597
	/a/	98.7	90.7	85	84	84.3	86	86.7	88. 7	93. 7	96	99	570

Based on the above table, we made charts of the F0 contours of /i/, /u/ and /a/ for each subject. In each chart the vertical (Y) axis shows the value of fundamental frequency, with Hz as the basic unit of measurement. The horizontal (X) axis presents the 11 measuring points. For the pupose of better comparison, the vertical axis covers a fixed value range based on the maximum and minimum frequencies recorded by female and male subjects respectively. For female subjects, the range is from 100Hz to 350Hz; for male subjects, from 0Hz to 250Hz.


F1
350
300
250
200
F1


F4
350
300
250
200
F4


M1
250
200
150
100
M1


M3
250
200
150
100
M3

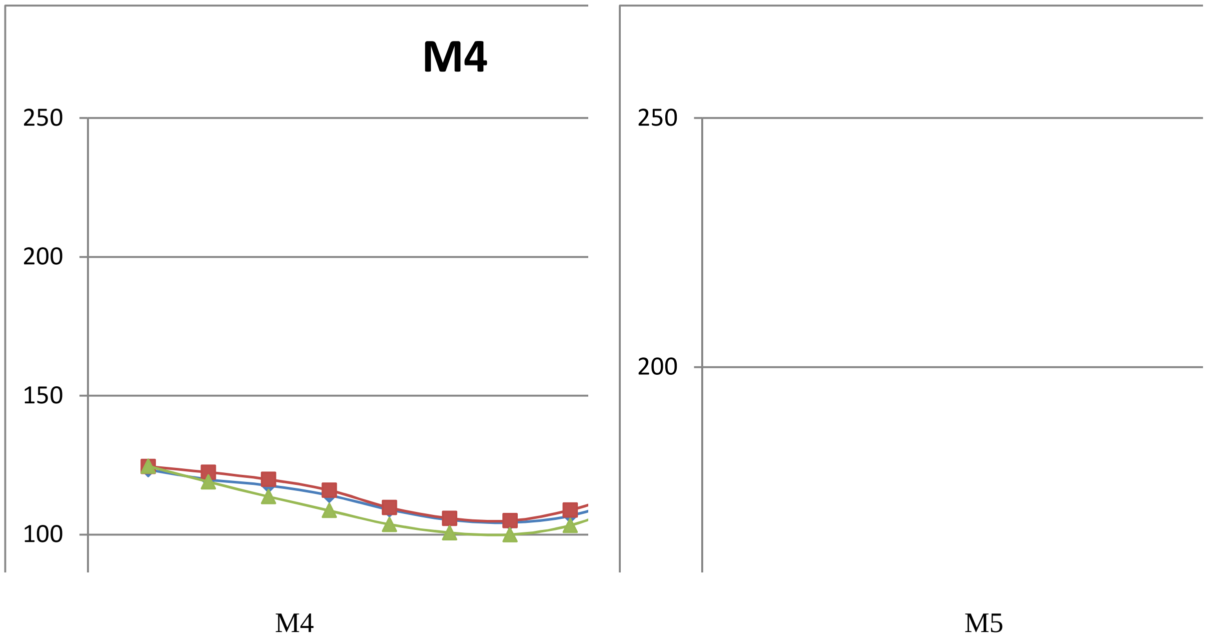


Figure 101. Average *Shangsheng* F0 contours of /i/, /u/ and /a/ for each Hungarian subject

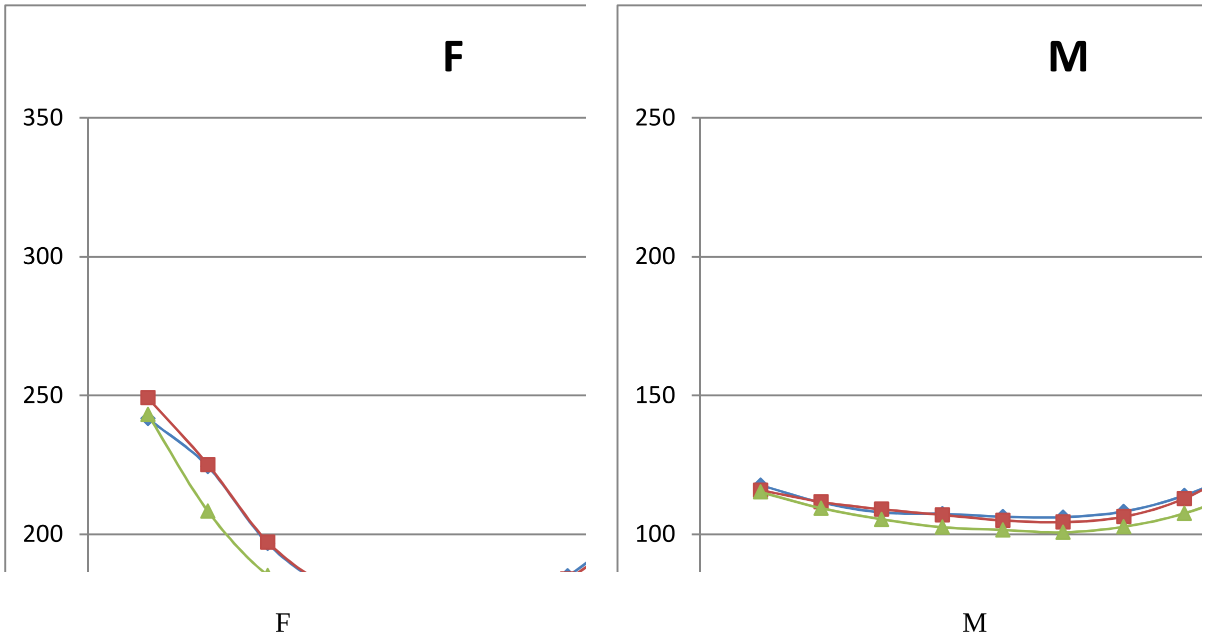


Figure 102. Average *Shangsheng* F0 contours of /i/, /u/ and /a/ for Hungarian subjects

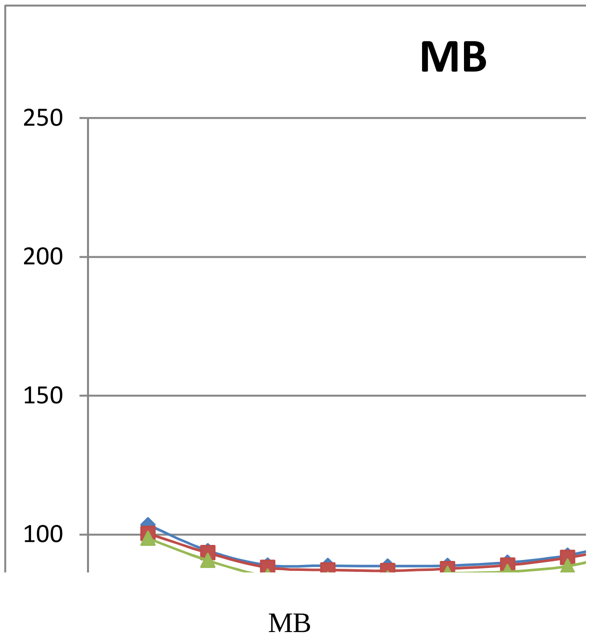


Figure 103. Average *Shangsheng* F0 contours of /i/, /u/ and /a/ for male Chinese subject

From the charts presented above, we can see that: (1) The manifestations of the IF0 in *Shangsheng* syllables produced by MB are not very obvious, since their average F0 contours of /i/, /u/ and /a/ are very concentrated and almost overlapping. (2) The *Shangsheng* syllables produced by the female subjects and M3 have distinctly different F0 levels in terms of /i/, /u/ and /a/, with the order of /u/＞/a/ or /i/＞/a/. Note that the IF0 mainly exist in F4's and M3's high frequency domain when these two subjects produce *Shangsheng* syllables. (3) The performances of the IF0 in M1 and M4 are not at all obvious, since their average F0 contours of /i/, /u/ and /a/ are highly concentrated and almost overlapping. (4) The average F0 contours of /i/, and /u/ produced by all the Hungarian female subjects are almost overlapping and relatively higher than of /a/ produced by them, which indicates the fairly obvious performances of the IF0. And the average F0 contours of /i/, /u/ and /a/ produced by all the Hungarian male subjects are virtually the same as those produced by all the Hungarian female subjects, but with the IF0 appearing in their high frequency domains. In other words, the performances of the IF0 are relatively obvious in F's and M's average F0 contours. (5) Based on those whose average F0 contours reflect the manifestations of the IF0 in *Shangsheng* syllables, we can see that the IF0 in high areas of the individual's own frequency domain is more distinct than that in low areas, which corresponds to Zhu Xiaonong's statement[194] (See 5.1.3).

All in all, the manifestations of the IF0 in *Shangsheng* tone produced by native speaker are not very significant. However, three of the six Hungarian subjects who produce the *Shangsheng* tone without showing the obvious manifestations of the IF0, in other words, 50% of the Hungarian leaners in our sample, are likely to produce the *Shangsheng* tone accompanied by the quite obvious existence of the IF0. What is more, the IF0 is reflected quite obviously in all the Hungarian female subjects and 25% of the male subjects, which is in consistent with Zhu Xiaonong's view[195]: that is, that the IF0 appears more obviously in female than in male speakers. (See 5.1.3 above)

5.3.4. Range

For *Shangsheng*, the ranges indicate how decline and steep the contour is. At this point, range refers to the difference between the maximal and minimum value of F0. There are two ranges should be counted, one on the first half part of F0 contour, another on the second half part (labelled range 1 and range 2). Therefore, the narrower the range 1, the less steep the falling part; the narrower the range 2, the less steep the rising part.

Table 74: *The average ranges of 16 Shangsheng syllables pronounced by 7 subjects*

Subjects	The provious maximal value of average F0	Minimum value of average F0	The posterior maximal value of average F0	Rang 1	Rang 2
F1	202.7	167.6	218.3	35.1	50.7
F4	286.9	177.9	312	109	134.1
M1	114.6	101.3	138.3	13.3	37
M3	115	111.1	195.8	3.9	84.7
M4	124.1	103.8	135	20.3	31.6
M5	112.1	97.6	146.6	14.5	49
BM	101.3	87.1	104.8	14.2	17.7

[194] Zhu 2005.

[195] Zhu 2005.

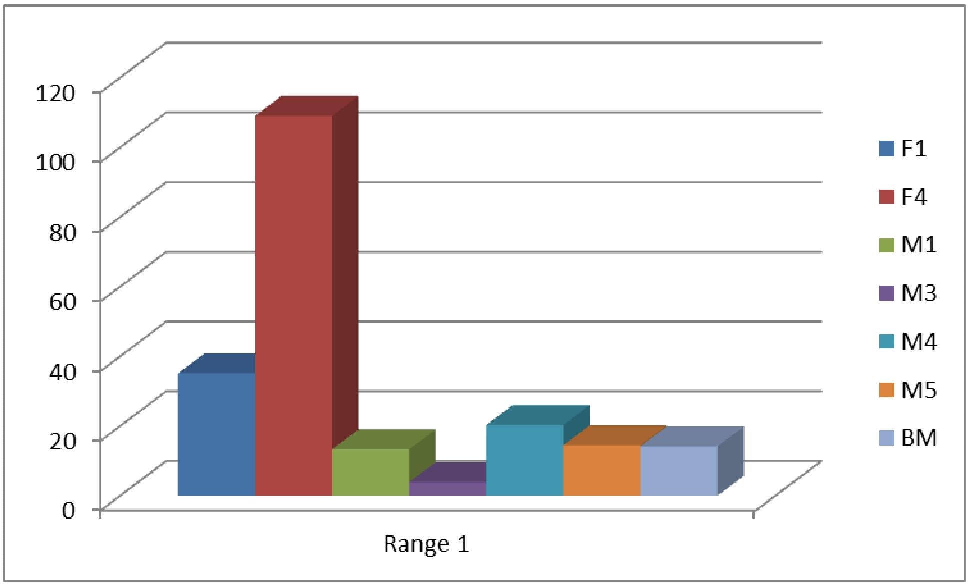


Figure 104. The average range 1 of 16 *Shangsheng* syllables pronounced by 7 subjects

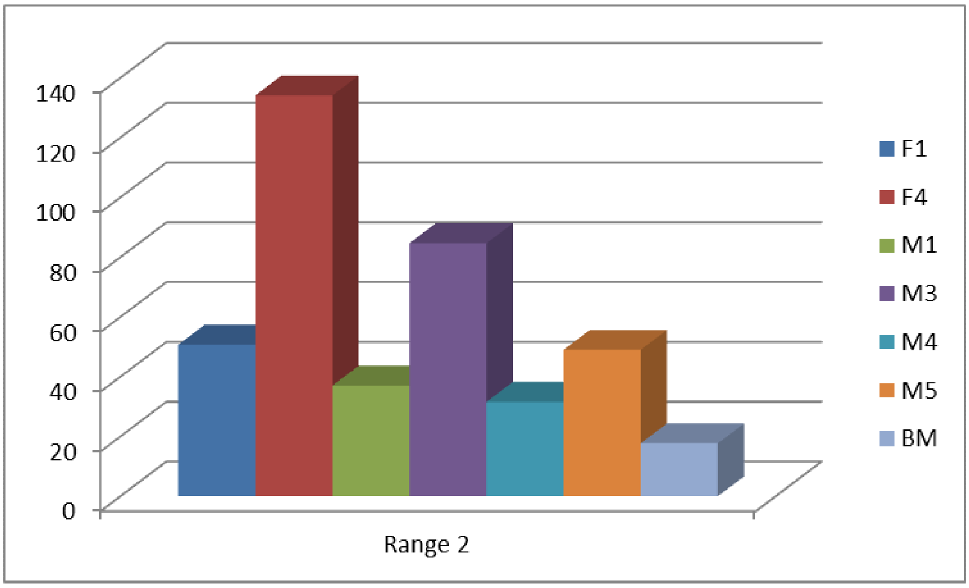


Figure 105. The average range 2 of 16 *Shangsheng* syllables pronounced by 7 subjects

From the table and the figure presented above, we can see: (1) the average range 1 and range 2 of BM are 14.2Hz and 17.7Hz; (2) the average ranges of F1, M1 and M4 are relatively close to BM's. In other words, the decline and steepness of the *Shangsheng* contours pronounced by these 3 subjects are relatively appropriate; (3) the average ranges of F4 are much wider than those of BM, almost 8 times; (4) the average range 1 of M3 is quite narrower than BM's, while the range 2 is considerably wider, which indicates that M3 does not produce acceptably steep falling and rising contours in *Shangsheng*. What is more, the average range 1 of M3's

Shangsheng is very narrow and is almost identical with his *Yangping* contour; (5) the average ranges of M5's is extremely close to BM's, while the range 2 is relatively wider. All in all, 50% of the Hungarian university learners whose recordings were usable in this phase of the investigation do not appear to experience problems in pronouncing *Shangsheng* in terms of range, but 50% of them do find it difficult.

5.3.5. Summary

In this section, the recordings of 7 subjects are examined in terms of the duration, the fundamental frequency (F0) and the tone range of their *Shangsheng* syllables. On the basis of measurement, extraction, analysis and comparison, the following conclusions can be drawn.

(1) In terms of the duration, 70% Hungarian university learners – one female and six male subjects in the sample - differ considerably from the "model" native speaker in terms of the duration of *Shangsheng* syllables, most in producing much shorter durations. And, male subjects are more likely to have problems in producing appropriate durations of *Shangsheng* syllables. What is more, 30% of Hungarian learners produce deviant orders of durations of /i/, /u/ and /a/, which indicates that the inappropriate production of vowels probably causes problems in duration.

(2) With regard to fundamental frequency (F0), only about 33% of the Hungarian subjects pronounce *Shangsheng* syllables correctly; in other words, their average F0 contours are appropriate fall-rise patterns; in contrast, about 67% of the Hungarian subjects have difficulty producing appropriately fall-rise contours. The deviant contours they produce mainly appear as a contour with an extremely sharp fall-rise pattern, an excessively steep, a rising pattern without a clear falling start, or an inappropriate point. As to the pitch value of F0, it emerges that the male Hungarian subjects are closely grouped in terms of pitch around the native male, while all the female Hungarian subjects were, not surprisingly, higher in pitch than the male native. Concerning the intrinsic vowel fundamental frequency (IF0), it turned out that the effect of IF0 is quite obvious in Hungarian learners when they are producing the *Shangsheng* tone, especially among female learners, with 100% in the sample showing

IF0 effects.

(3) Summarizing the findings about pitch range we can say that 50% of these Hungarian university learners (one female and two male in the sample) have a fairly appropriate range in *Shangsheng*, that is, their ranges are relatively close to those of native speakers. However, 50% of them (one female and two male in the sample) can be considered deviant in the sense that they produce much wider or narrower ranges.

5.4. *Qusheng*

The fourth tone, also called *QùShēng*, starts at the highest pitch level (5) and falls to the lowest (1). The value of tone is therefore 51. Feng Zhichun in *Xiandai Hanyu* defines Qusheng as a full falling tone or high falling tone with its pitch starting from the highest then falling to the lowest[196].

In general, the start is not the tone goal in terms of tone components, but the falling tone constitutes an exception to this rule due to the fact that its peak lies in the start. The linguistic goal of the falling tone may occur at the very beginning of the contour or (in many cases) the adjacent peak point comes 5%-20% into the start and the body. The linguistic goal of many tones in certain Chinese varieties lies in the body, but Zhu Xiaonong points out that for the high-rising tone or the high-falling tone the body only constitues a transitional section[197]. The main aim of this section is to present the voice samples of the fourth tone, *Qusheng*, which were recorded by 12 subjects, including the measurement, extraction, analysis and comparison. The recordings are examined in terms of the duration, the fundamental frequency (F0), the tone range and the tone value of the monosyllabic tone produced by the recorded subjects, thus creating a detailed profile of the *Qusheng* pronunciation of this sample of Hungarian university students (six male, labeled M1, M2 etc., four female, labeled F1, F2 etc.), juxtaposed with the pronunciation of one male and one female native speaker of Chinese (labeled MB and FB respectively). The actual methodology used in the measurement of

---

[196] Feng 2008: 100.
[197] Zhu 2010: 278.

duration, F0 and tone range was detailed in 4.3 above; note that as with the three other tones the initial or “onset” of each syllable was discarded; the 11 measuring points all fall within the nucleus or rime. Thus the word “start” used in the comments about the subjects’ performance refers to the beginning of the nucleus, and not to the beginning of the whole syllable.

Because of the existence of creaky voice, the F0 of F2’s, M2’s and M6’s *Qusheng* syllables could not easily be measured. Since determining the duration is still possible, we will use 12 subjects’ data for analyzing durations. However, only the 9 subjects who produced *Qusheng* syllables without creaky voice accompanied will be analyzed for F0 and range. They are 4 female subjects: F1, F3, F4, FB; and 5 male subjects: M1, M3, M4, M5, MB. Note that creaky voice also occurs in FB’s production of *Qusheng*, but since the F0 of FB’s *Qusheng* syllables could still be measured, although not very satisfactorily, this subject’s *Qusheng* data were retained. However, it should be noted that as a result of creaky voice the ends of the F0 of her *Qusheng* syllables are relatively lower, giving her a very wide pitch range.

The 15 *Qusheng* syllables used in the recordings, as explained in 4.3 above, fit into the following matrix:

Table 75: *Matrix of Qusheng syllables*

	à	ì	ù
b	罢、爸	币、必	部、布
d	大	第、弟	度、杜
Zero-initials		易、意	误、物

5.4.1. Duration

The method used for determining the duration of the syllable has been presented before. The following table shows the durations of each *Qusheng* syllable pronounced by 12 subjects, as well as average durations. The unit of duration is the millisecond (ms).

Table 76: *Duration of 15 different Qusheng syllables pronounced by 12 different subjects.*

Syllables	F1	F2	F3	F4	M1	M2	M3	M4	M5	M6	FB	MB
第 dì	279	310	183	215	215	174	181	284	241	277	306	245
大 dà	224	292	170	232	199	200	140	273	219	317	301	220
易 yì	316	310	189	244	214	185	196	269	183	377	284	329
罢 bà	254	261	182	239	181	148	179	335	194	350	268	237
误 wù	351	241	197	264	210	197	151	291	300	381	339	300
度 dù	354	414	176	233	252	173	196	291	204	388	230	272
爸 bà	279	295	189	250	181	205	147	224	131	360	195	187
部 bù	284	307	167	205	207	161	169	269	195	368	284	245
币 bì	237	323	171	192	198	183	147	238	168	391	298	309
布 bù	272	347	207	226	201	189	130	258	194	414	300	267
意 yì	359	301	201	230	174	188	201	283	219	433	323	317
物 wù	348	416	146	285	241	235	193	283	270	421	300	413
弟 dì	278	308	176	261	201	228	166	250	227	444	281	319
必 bì	273	243	127	265	231	191	158	267	222	419	378	297
杜 dù	265	304	178	227	203	214	167	182	176	473	228	268
Average	290	312	177	238	207	191	168	267	210	388	288	282

On the basis of the averages displayed in the bottom row of the above table, we made a chart for further comparison.

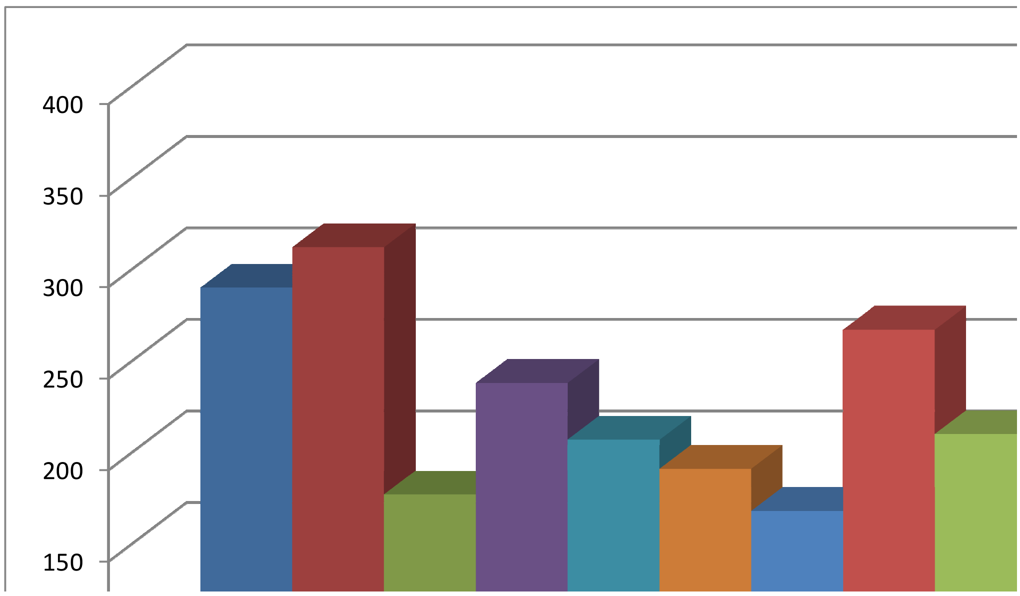


Figure 106. Average duration of *Qusheng* syllables pronounced by 12 different subjects

The chart clearly shows that the average duration of M6 is the longest, while the average duration of M3 is the shortest. The average durations of FB and MB are 288ms, 282ms, respectively. The average durations of F1, F2 and M4 are close to FB's and MB's, the average durations of F4, M1 and M5 are relatively shorter than FB's and MB's. The average duration of F3, M2 and M3 are much shorter than FB's and MB's, while the average duration of M6 is much longer than those of "model" native speakers. In other words, 40% of Hungarian university learners – mainly male subjects in the sample - differ considerably from the "model" native speakers in terms of the duration of *Qusheng* syllables, most often in producing much shorter durations.

As we have said, there can be no doubt producing a tone in too long or too short a way often causes problems in duration, but it is probably not the only reason. The inappropriate pronounciation of the vowel which is bearing a tone also can lead to duration problems. Generally speaking, the duration of a vowel is associated with its height, and each vowel has an intrinsic vowel duration (referred to as IVD)[198] (Zhu Xiaonong 2005). Catford (1977)[199], believes that the duration of a high vowel is longer than that of a low one. His statement is described as IVD "i>a".

The following table shows the average durations (m) of *Qusheng* produced by 12 subjects, in terms of different vowels /i/, /u/ and /a/ which are bearing the *Qusheng* tone, as well as the average durations of all the female subjects (F) and the male subjects (M). The unit of duration is the millisecond (ms), and "n" refers to the number of sample syllables. The last column in the table shows the subject's "duration order" of /i/, /u/ and /a/, with the longest vowel first and the shortest vowel last. For example, F1's average durations of *Qusheng* tone which are contained by the vowels /i/, /u/ and /a/ are 290ms, 312ms and 252ms, respectively. The "duration order" of /i/, /u/ and /a/ is /u/>/i/>/a/.

[198] Zhu 2005: 72.
[199] Catford 1977: 45.

Table 77: *Average durations (m) of vowels /i/, /u/ and /a/ with Qusheng tone produced by 12 subjects*

	/i/		/u/		/a/		the order
	m	n	m	n	m	n	
F1	290	6	312	6	252	3	/u/>/i/>/a/
F2	299	6	338	6	283	3	/u/>/i/>/a/
F3	175	6	179	6	180	3	/a/>/u/>/i/
F4	235	6	240	6	240	3	/u/=/a/>/i/
M1	206	6	219	6	187	3	/u/>/i/>/a/
M2	192	6	195	6	184	3	/u/>/i/>/a/
M3	175	6	168	6	155	3	/i/>/u/>/a/
M4	265	6	262	6	277	3	/a/>/i/>/u/
M5	210	6	223	6	181	3	/u/>/i/>/a/
M6	390	6	408	6	342	3	/u/>/i/>/a/
F	250	6	267	6	239	3	/u/>/i>/a/
M	240	6	246	6	221	3	/u/>/i>/a/
FB	312	6	280	6	255	3	/i/>/u/>/a/
MB	303	6	294	6	215	3	/i/>/u/>/a/

From the above table, we can see that: (1) FB's average durations of /i/, /u/ and /a/ are 312ms, 280ms and 255ms, and MB's durations of these three vowels are 303ms, 294ms and 215ms, respectively. In other words, the orders of both FB's and MB's durations of these three vowels are /i/>/u/>/a/, which corresponds to the "i>a" principle. (2) Only M3's duration order of /i/, /u/ and /a/ is the same as the "model" native speakers', which is /i/>/u/>/a/. (3) Many subjects: F1, F2, M1, M2, M5 and M6 produce the same duration order of /i/, /u/ and /a/, which is /u/>/i/>/a/. Since both /i/ and /u/ are high vowels, the order in /u/>/i/>/a/ makes sense. (4) 30% of the Hungarian subjects produce deviant orders of durations of /i/, /u/ and /a/. F4's is in the order of /u/=/a/>/i/, M4's in /a/>/i/>/u/ and F3's in /a/>/u/>/i/. In other words, they all produce the duration order of /i/, /u/, /a/ as "i＜a", which is in contrast with the "i>a" principle. (4) The average durations of /i/, /u/ and /a/ produced by all the Hungarian female and male subjects are sorted by /u/>/i/>/a/, which can be considered as identical to the “model” native speakers’, since both /i/ and /u/ are high vowels. All in all, the deviant orders of durations of /i/, /u/ and /a/ indicate that the inappropriately produced vowels are likely to cause problems in duration.

It is worth mentioning that the average durations of /i/, /u/ and /a/ produced by all the Hungarian female subjects (F) and male subjects (M) are 252ms [(250＋267＋239) ÷ 3], 236ms [(240＋246＋221) ÷ 3], respectively. In other words, although the sample is admittedly small, there are some grounds for believing in the possibility of a connection between duration and gender. This conclusion is, however, in contrast with the view expressed by Zhu Xiaonong, who believes there is no connection between duration and gender[200].

However, duration is a relative concept. It is hard to define a standard or universally “desirable” value of duration, because it depends so strongly on the individual’s general rate of speech, not to mention the nature of each specific utterance and the conditions under which it is made. Discussing the duration of specific syllables within the parameters of an individual’s own tone system probably makes better sense. This issue will be discussed later in this chapter.

5.4.2. Fundamental frequency

The fundamental frequency (“F0”or “F_0” in PRAAT-processed data), namely tone pitch, is a universal, indispensable and very distinctive feature of pronunciation in Chinese. In this section, firstly, each subject’s F0 values for every *Qusheng* monosyllable, together with the means for all the syllables, are presented in a coMBined table. Secondly, each subject’s F0 contours for all the *Qusheng* monosyllables are plotted in one chart for further analysis and comparison. In each chart the vertical (Y) axis shows the value of fundamental frequency, with Hz as the basic unit of measurement. The horizontal (X) axis presents the 11 measuring points. For better comparison purposes, the vertical axis covers a fixed value range based on the maximum and minimum frequencies recorded by female and male subjects respectively. For female subjects, the range is from 100Hz to 400Hz; for male subjects, from 0Hz to 200Hz.

[200] Zhu 2005: 77.

Thirdly, the average contour of all the subject's *Qusheng* syllables is plotted in another chart. (For details of the methods used for measuring fundamental frequency (F0), see 3.2.1)

5.4.2.1. Individual subjects' F0 data

The F0 data table and charts of every subject will now be provided.

(1) F1

The following table shows the F0 data for subject F1, including 15 *Qusheng* monosyllables. Each monosyllable has 11 F0 values which correspond to the 11 measuring points. The means of the measurements made at each point each point are presented in the final row of the table.

Table 78: *F1's F0 values for 15 Qusheng syllables at 11 measuring points*

Syllables	1	2	3	4	5	6	7	8	9	10	11
第 dì	258	255	251	241	228	212	201	193	187	182	178
大 dà	253	251	247	241	232	219	207	196	187	180	180
易 yì	253	244	247	243	232	217	200	186	177	173	166
罢 bà	250	247	243	238	228	218	208	195	186	179	174
误 wù	265	248	249	247	237	223	208	197	183	172	170
度 dù	255	255	251	241	225	209	193	183	175	167	164
爸 bà	239	238	233	224	214	204	196	185	174	167	164
部 bù	262	258	251	240	228	216	204	194	185	179	174
币 bì	257	251	248	242	233	219	206	195	186	177	171
布 bù	262	256	248	237	225	213	201	192	183	178	173
意 yì	255	255	248	240	228	213	201	188	179	173	166
物 wù	263	265	260	249	231	215	202	191	181	169	165
弟 dì	250	251	249	241	227	213	199	185	178	173	168
必 bì	247	244	238	228	215	202	190	180	175	171	169
杜 dù	265	261	257	247	233	216	203	195	189	182	179
Average	255.6	251.9	248	239.9	227.7	213.9	201.3	190.3	181.7	174.8	170.7

On the basis of the above table, we made the following two F0 contour charts of F1's *Qusheng* pronunciation. The first chart shows the contour for every monosyllable, the second shows the mean of all the contours. The two charts are followed by brief comments on the data.

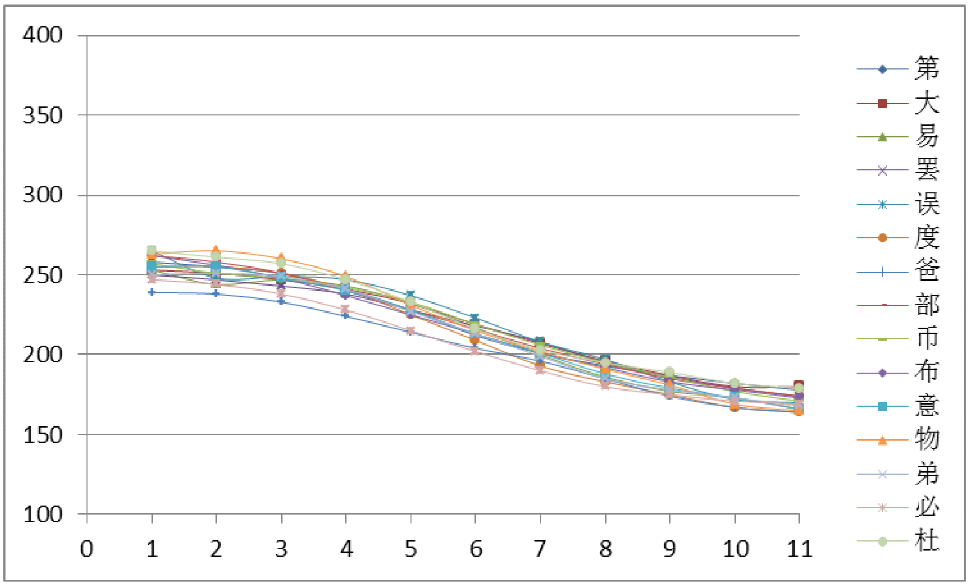


Figure 107. The F0 contour of 15 *Qusheng* syllables pronounced by F1

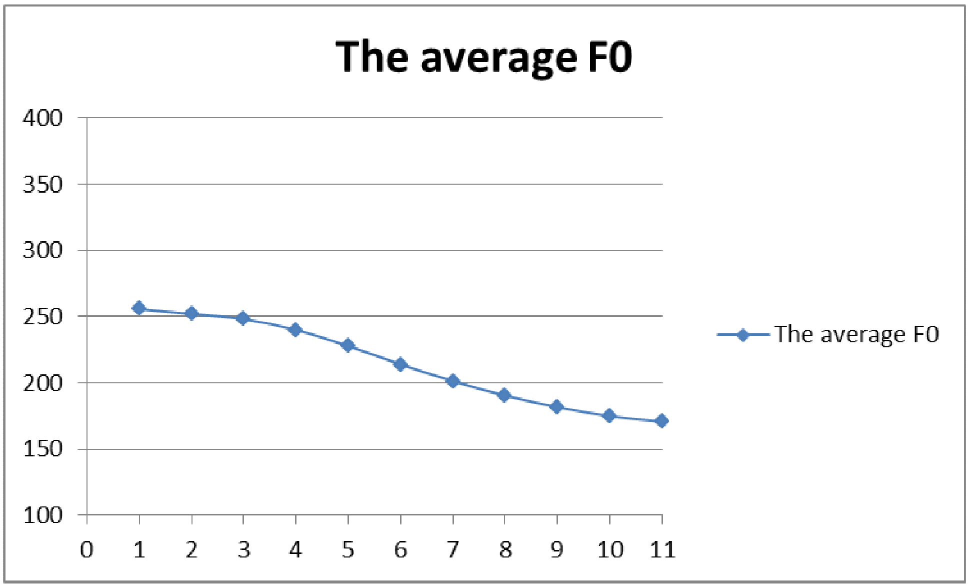


Figure 108. The average F0 contour of 15 *Qusheng* syllables pronounced by F1

Based on Chart 1, the following conclusions can be made. (1) The F0 range of F1 is 160-270Hz, that is, about 110Hz. (2) The F0 contours of F1 are highly concentrated. (3) Most F0 contours of F1 have a considerably smooth start. (4) Overall, the falling parts of most F0 contours are noticeably steep, and all F0 contours of F1 turn out as falling patterns.

From Chart 2, we can see that the mean F0 contour of F1 presents a falling pattern with a relatively delayed falling part. The value of the mean F0 contour ranges between 170 and 260Hz.

(2) F3

Table 79: *F3's F0 values for 15 Qusheng syllables at 11 measuring points*

Syllables	1	2	3	4	5	6	7	8	9	10	11
第 dì	251	242	234	222	208	200	194	183	178	173	173
大 dà	255	236	226	216	205	194	181	170	169	167	170
易 yì	238	236	238	237	229	219	206	193	186	177	168
罢 bà	255	236	226	216	205	194	181	170	169	167	170
误 wù	245	240	237	226	212	197	187	180	178	184	184
度 dù	251	240	229	217	205	193	184	177	173	176	188
爸 bà	202	196	190	185	185	186	187	186	188	190	198
部 bù	250	242	232	220	209	201	196	192	189	186	187
币 bì	225	223	214	207	203	199	196	191	190	191	193
布 bù	238	228	221	215	213	209	205	201	201	203	205
意 yì	220	222	220	214	207	202	197	194	188	179	180
物 wù	229	225	223	219	213	208	201	195	188	184	183
弟 dì	215	212	207	204	204	204	202	197	195	197	204
必 bì	235	229	226	223	216	207	205	199	197	198	195
杜 dù	256	243	231	217	204	194	188	182	178	184	192
Average	237.7	230	223.6	215.9	207.9	200.5	194	187.3	184.5	183.7	186

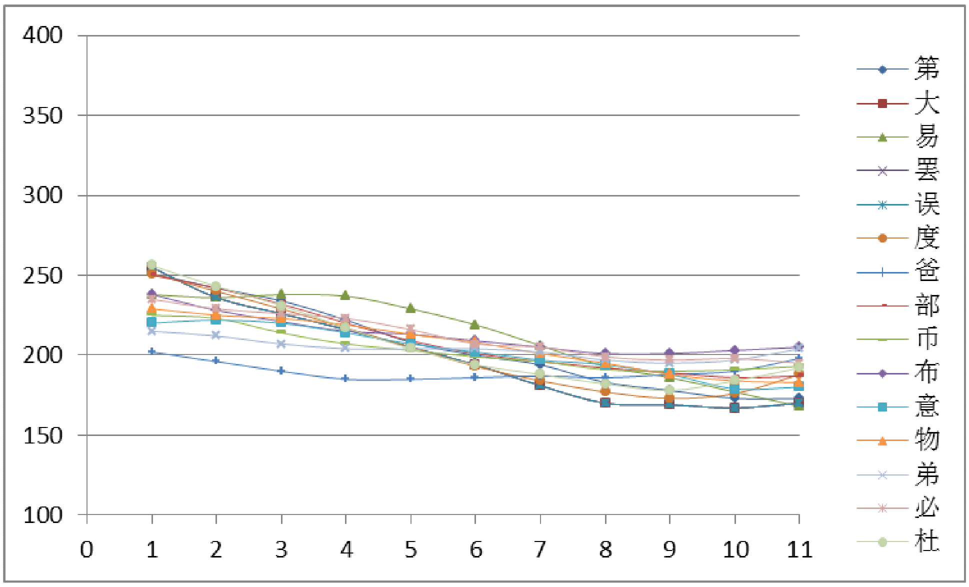


Figure 109. The F0 contour of 15 *Qusheng* syllables pronounced by F3

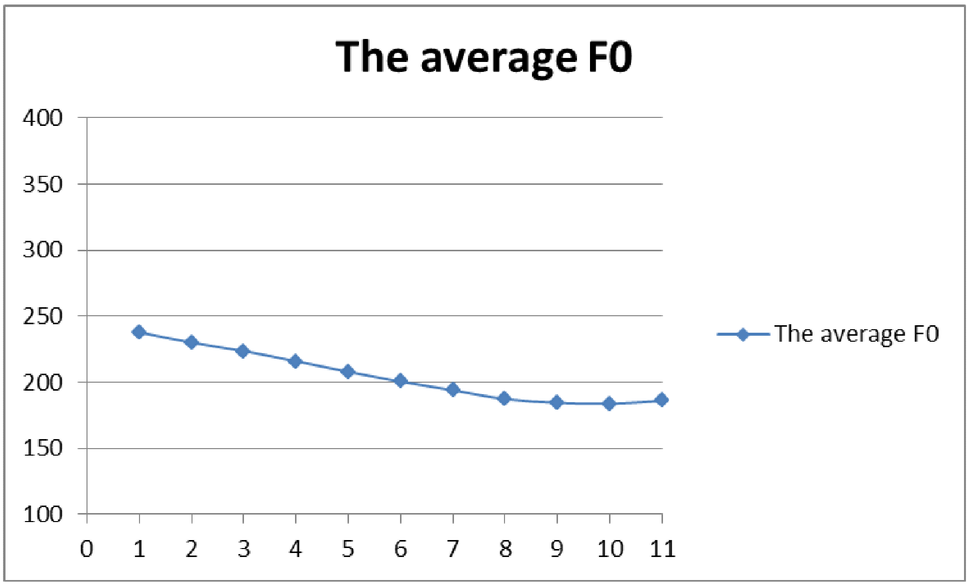


Figure 110. The average F0 contour of 15 *Qusheng* syllables pronounced by F3

Based on Chart 1, the following conclusions can be made. (1) The F0 range of F3 is 160-260Hz, that is, about 100Hz. (2) The F0 contours of F3 are relatively concentrated. (3) Most F0 contours of F3 start falling at the beginning quite obviously. (4) Overall, the falling parts of most F0 contours are considerably gentle. (5) Most F0 contours of F3 turn out as falling patterns, but include a smooth or even rising ending.

From Chart 2, we can see that the mean F0 contour of F3 presents a falling pattern with a straight falling start, but followed by a quite smooth ending. The value of the mean F0 contour ranges between 180 and 240Hz.

(3) F4

Table 80: *F4's F0 values for 15 Qusheng syllables at 11 measuring points*

Syllables	1	2	3	4	5	6	7	8	9	10	11
第 dì	348	338	320	293	247	226	202	190	182	172	172
大 dà	309	284	270	245	227	212	198	189	183	175	167
易 yì	315	317	298	271	240	220	202	187	183	177	174
罢 bà	298	292	273	247	228	216	206	196	189	180	171
误 wù	293	296	286	262	233	211	196	186	179	175	170
度 dù	310	306	292	266	248	229	212	193	182	176	173
爸 bà	293	292	276	249	219	202	191	182	176	168	162
部 bù	342	331	319	302	278	246	220	194	180	175	173

币 bì	341	330	317	303	283	261	242	219	197	179	168
布 bù	311	317	300	273	240	213	190	174	169	169	171
意 yì	339	328	307	282	252	230	216	203	187	167	163
物 wù	277	293	296	282	253	223	205	192	179	176	174
弟 dì	300	309	307	290	264	241	226	207	188	171	171
必 bì	366	350	339	320	290	252	231	206	190	176	174
杜 dù	311	306	294	268	249	231	214	196	183	177	173
Average	316.9	312.6	299.6	276.9	250.1	227.5	210.1	194.3	183.1	174.2	170.4

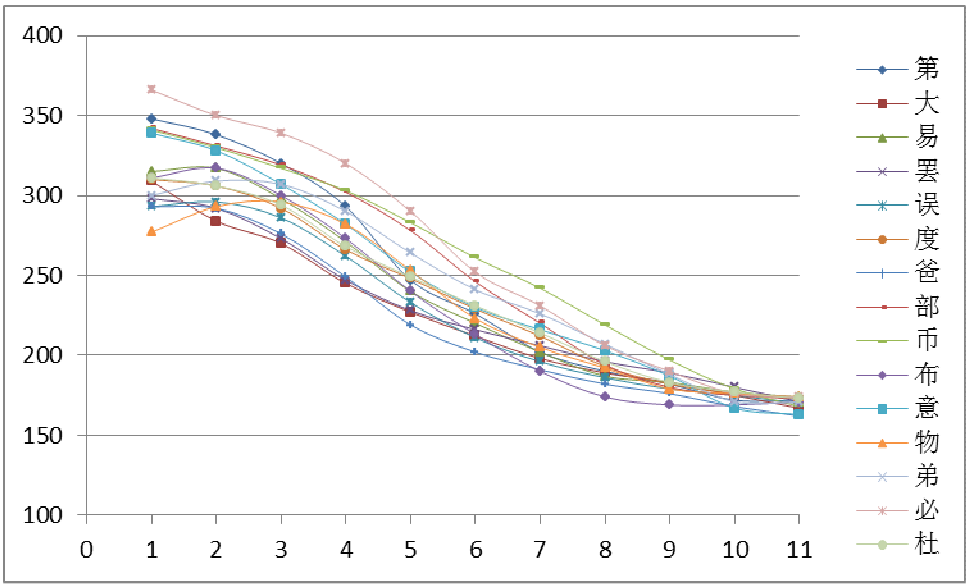


Figure 111. The F0 contour of 15 *Qusheng* syllables pronounced by F4

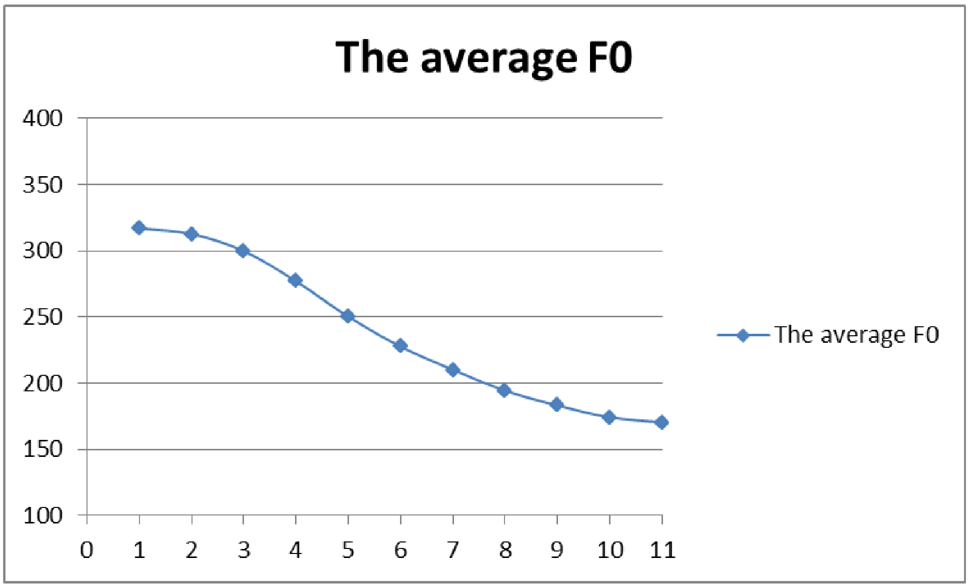


Figure 112. The average F0 contour of 15 *Qusheng* syllables pronounced by F4

Based on Chart 1, the following conclusions can be made. (1) The F0 range of F4 is 160-370Hz, that is, about 210Hz. (2) The F0 contours of F4 are relatively spread out. (3) Many F0 contours of F4 have a quite clear falling start. (4) Overall, the falling parts of most F0 contours are very steep, and most F0 contours of F4 turn out as falling patterns.

From Chart 2, we can see that the mean F0 contour of F4 presents a falling pattern with a straight falling start and a very steep falling part. The value of the mean F0 contour ranges between 170 and 320Hz.

(4) M1

Table 81: *M1's F0 values for 15 Qusheng syllables at 11 measuring points*

Syllables	1	2	3	4	5	6	7	8	9	10	11
第 dì	169	173	164	156	146	135	124	116	110	102	101
大 dà	140	148	142	137	130	120	110	104	101	98	89
易 yì	197	199	194	186	178	166	151	134	120	104	101
罢 bà	163	161	160	158	155	150	141	131	119	103	96
误 wù	198	199	198	190	177	162	142	125	114	103	83
度 dù	173	176	169	160	148	132	121	107	98	85	80
爸 bà	156	153	148	144	140	134	128	118	111	96	98
部 bù	186	193	185	177	166	153	143	134	122	114	102
币 bì	178	180	176	169	160	151	138	124	115	112	107
布 bù	190	191	187	179	168	152	135	128	121	115	115
意 yì	191	191	191	188	179	167	153	139	123	114	109
物 wù	196	196	192	187	179	167	150	131	116	111	110
弟 dì	194	196	191	183	176	168	151	133	121	111	111
必 bì	180	180	174	165	158	148	136	123	124	118	108
杜 dù	162	167	162	156	149	137	126	115	107	104	107
Average	178.2	180.2	175.5	169	160.6	149.5	136.6	124.1	114.8	106	101.1

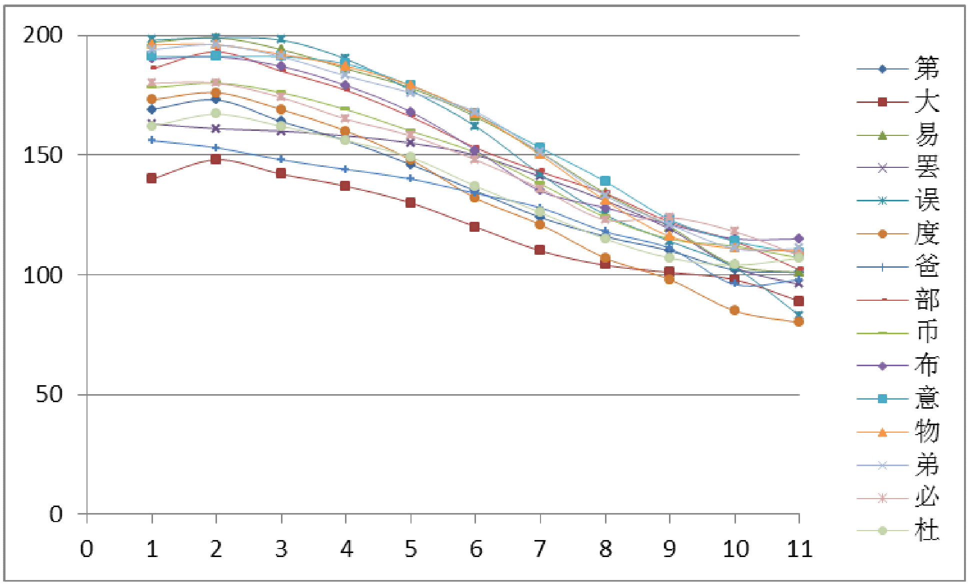


Figure 113. The F0 contour of 15 *Qusheng* syllables pronounced by M1

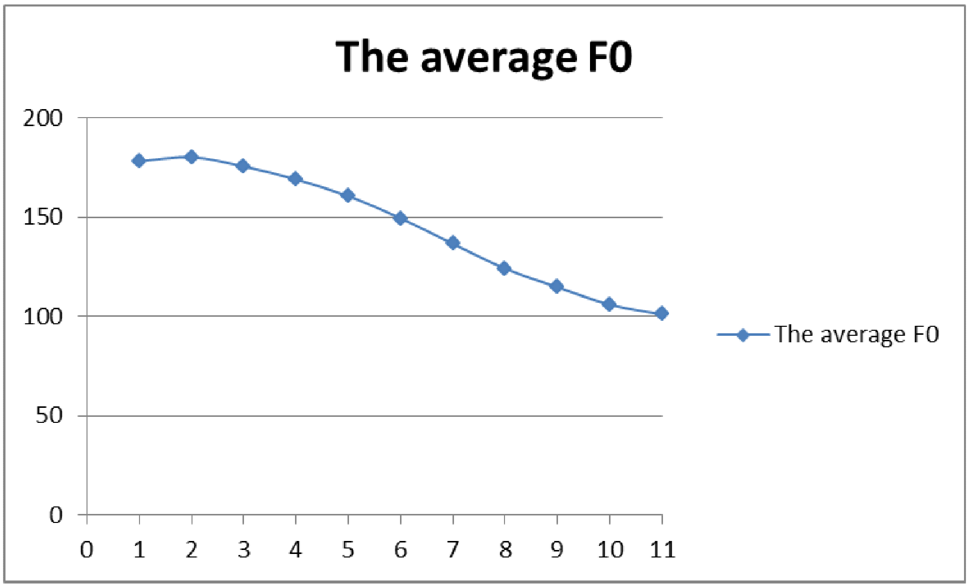


Figure 114. The average F0 contour of 15 *Qusheng* syllables pronounced by M1

Based on Chart 1, the following conclusions can be made. (1) The F0 range of M1 is 80-200Hz, that is, about 120Hz. (2) The F0 contours of M1 are relatively spread out. (3) Most F0 contours of M1 have a fairly smooth start. (4) Overall, the falling parts of most F0 contours are noticeably steep, and almost all F0 contours of M1 turn out as falling patterns.

From Chart 2, we can see that the mean F0 contour of M1 presents a falling pattern with a delayed falling start. The value of the mean F0 contour ranges between 100 and 190Hz.

(5) M3

Table 82: *M3's F0 values for 15 Qusheng syllables at 11 measuring points*

Syllables	1	2	3	4	5	6	7	8	9	10	11
第 dì	154	154	152	149	142	134	127	120	113	105	95
大 dà	137	134	131	130	125	120	119	114	106	100	117
易 yì	145	155	164	169	166	158	143	126	113	103	93
罢 bà	138	138	137	136	131	128	122	119	113	107	95
误 wù	145	160	172	178	179	174	166	155	141	126	105
度 dù	129	129	128	127	127	129	131	133	137	136	131
爸 bà	139	137	134	131	127	123	118	112	108	107	105
部 bù	152	157	156	152	145	137	129	123	115	105	92
币 bì	130	131	131	130	128	128	127	125	126	126	115
布 bù	145	152	154	154	151	146	139	131	124	119	114
意 yì	157	163	169	173	169	160	143	125	116	110	104
物 wù	143	152	159	162	160	151	137	124	113	106	105
弟 dì	170	172	168	163	156	146	135	122	113	107	103
必 bì	171	178	181	182	181	175	165	152	133	117	99
杜 dù	172	171	166	158	149	139	130	122	116	107	104
Average	148.5	152.2	153.5	152.9	149.1	143.2	135.4	126.9	119.1	112.1	105.1

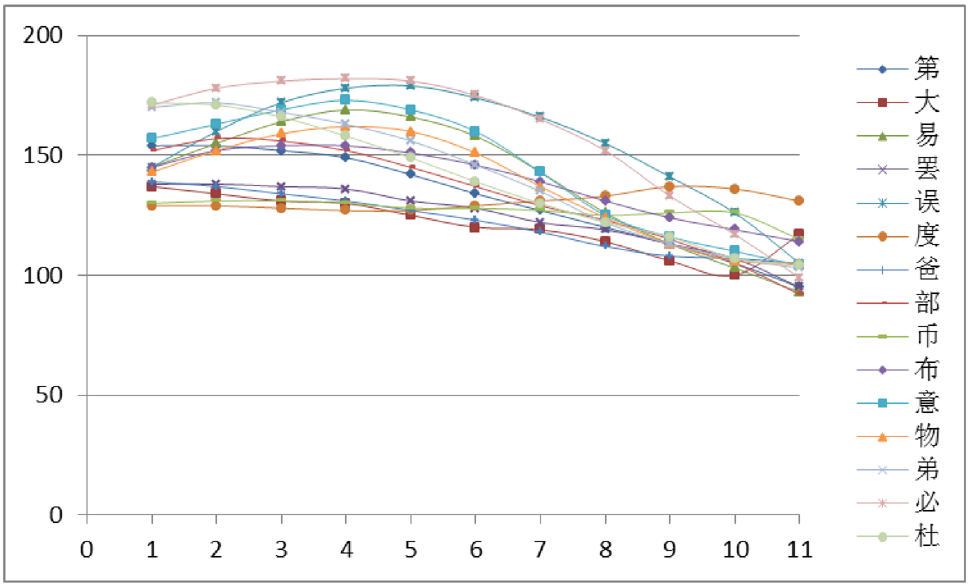


Figure 115. The F0 contour of 15 *Qusheng* syllables pronounced by M3

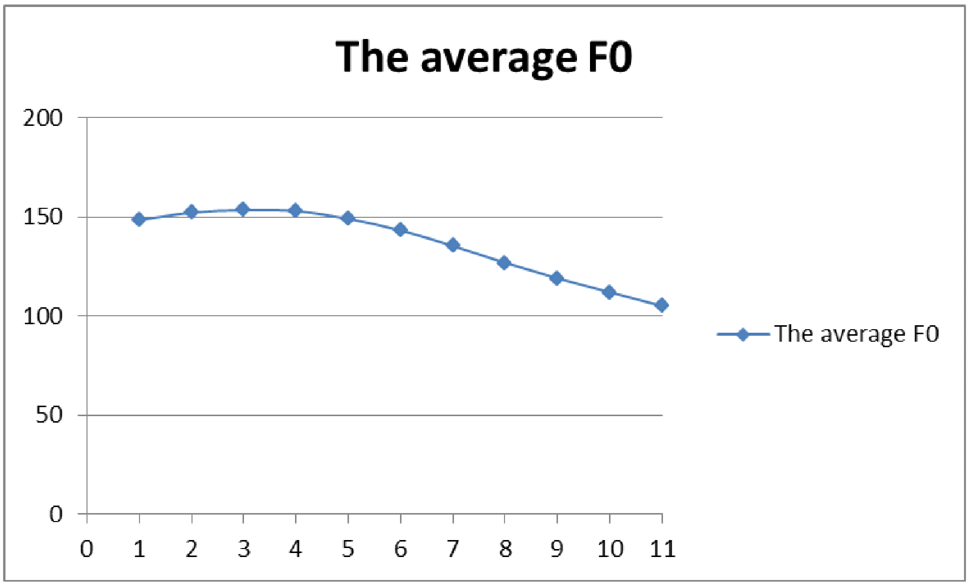


Figure 116. The average F0 contour of 15 *Qusheng* syllables pronounced by M3

Based on Chart 1, the following conclusions can be made. (1) The F0 range of M3 is 90-190Hz, that is, about 120Hz. (2) The F0 contours of M3 are considerably spread out. (3) Most F0 contours of M3 have a clear rising or smooth start, and the inflection point of falling comes relatively late. (4) Overall, the falling parts of most F0 contours are relatively gentle, and almost F0 contours of M3 turn out as falling patterns.

From Chart 2, we can see that the mean F0 contour of M3 presents a falling pattern with a considerably delayed falling start. The value of the mean F0 contour ranges between 100 and 160Hz.

(6) M4

Table 83: *M4's F0 values for 15 Qusheng syllables at 11 measuring points*

Syllables	1	2	3	4	5	6	7	8	9	10	11
第 dì	127	122	117	113	109	105	102	101	98	96	94
大 dà	126	121	115	109	106	103	103	101	99	100	99
易 yì	115	116	114	113	111	109	107	106	103	102	103
罢 bà	115	113	113	113	113	114	114	115	116	116	118
误 wù	115	114	114	115	115	114	112	110	107	105	102
度 dù	123	120	117	115	113	112	110	107	105	106	100
爸 bà	116	121	113	109	106	103	100	98	98	98	94
部 bù	118	115	113	111	109	107	105	104	103	99	98

币 bì	119	117	113	109	104	101	99	97	95	97	96
布 bù	125	120	117	114	112	110	108	107	107	106	107
意 yì	113	113	114	113	111	111	109	106	103	102	106
物 wù	120	115	115	114	112	110	107	105	102	100	99
弟 dì	117	116	113	110	107	106	105	105	103	103	110
必 bì	117	115	115	115	115	115	115	114	114	116	114
杜 dù	127	124	122	119	115	112	110	108	107	105	105
Average	119.5	117.5	115	112.8	110.5	108.8	107.1	105.6	104	103.4	103

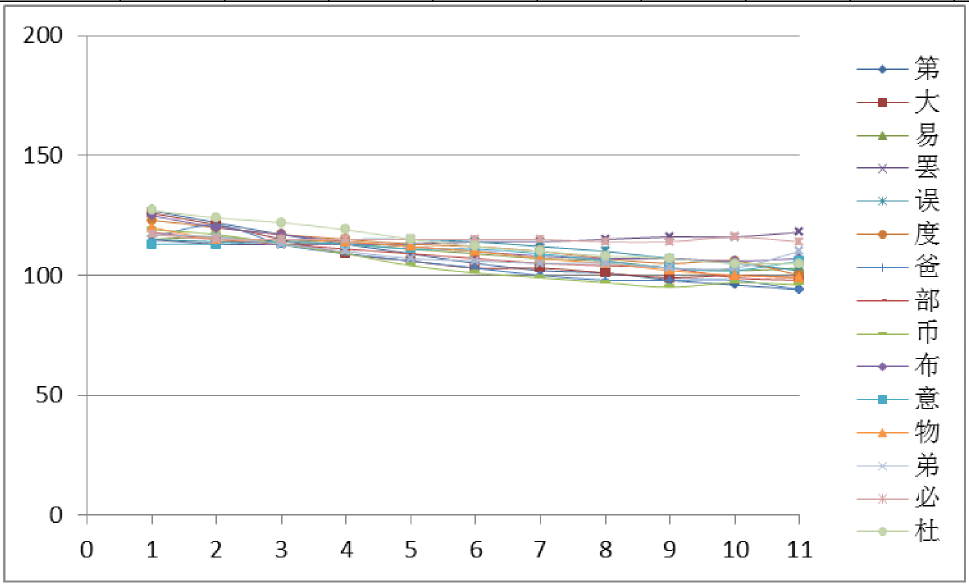


Figure 117. The F0 contour of 15 *Qusheng* syllables pronounced by M4

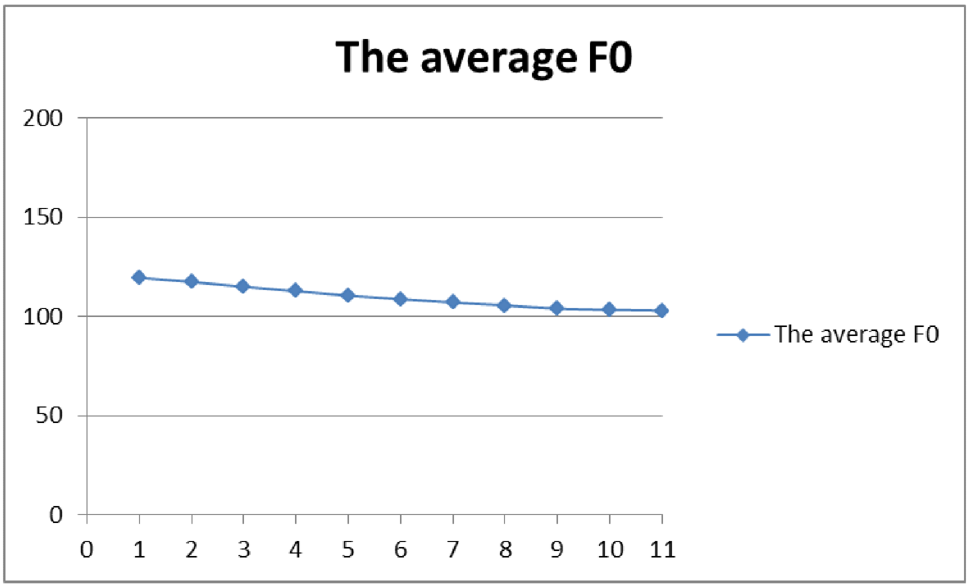


Figure 118. The average F0 contour of 15 *Qusheng* syllables pronounced by M4

Based on Chart 1, the following conclusions can be made. (1) The F0 range of M4 is 90-130Hz, that is, about 40Hz. (2) The F0 contours of M4 are highly concentrated. (3) Overall, the falling parts of most F0 contours are very gentle, although almost F0 contours of M3 turn out as falling patterns.

From Chart 2, we can see that the mean F0 contour of M4 presents a falling pattern but with a relatively straight falling part. The value of the mean F0 contour ranges between 100 and 120Hz.

(7) M5

Table 84: *M5's F0 values for 15 Qusheng syllables at 11 measuring points*

Syllables	1	2	3	4	5	6	7	8	9	10	11
第 dì	190	184	178	161	126	106	95	89	87	89	93
大 dà	181	177	168	147	122	104	97	94	91	90	91
易 yì	164	166	168	165	159	151	126	113	101	94	89
罢 bà	177	171	166	159	144	125	107	97	93	91	90
误 wù	156	161	169	177	179	169	140	111	96	92	89
度 dù	192	189	183	168	138	113	100	94	91	87	92
爸 bà	156	126	120	115	111	109	107	103	99	97	97
部 bù	190	189	187	182	169	148	122	104	97	94	91
币 bì	189	184	181	176	169	156	133	115	105	95	92
布 bù	189	183	180	176	170	153	132	113	100	94	91
意 yì	171	171	169	164	151	129	112	100	96	92	88
物 wù	133	138	145	155	164	171	169	153	123	100	92
弟 dì	189	183	180	173	151	122	104	95	92	88	88
必 bì	188	188	185	177	161	128	106	95	91	88	88
杜 dù	177	175	173	168	157	139	121	104	96	93	91
Average	176.1	172.3	170.1	164.2	151.4	134.9	118.1	105.3	97.2	92.3	90.8

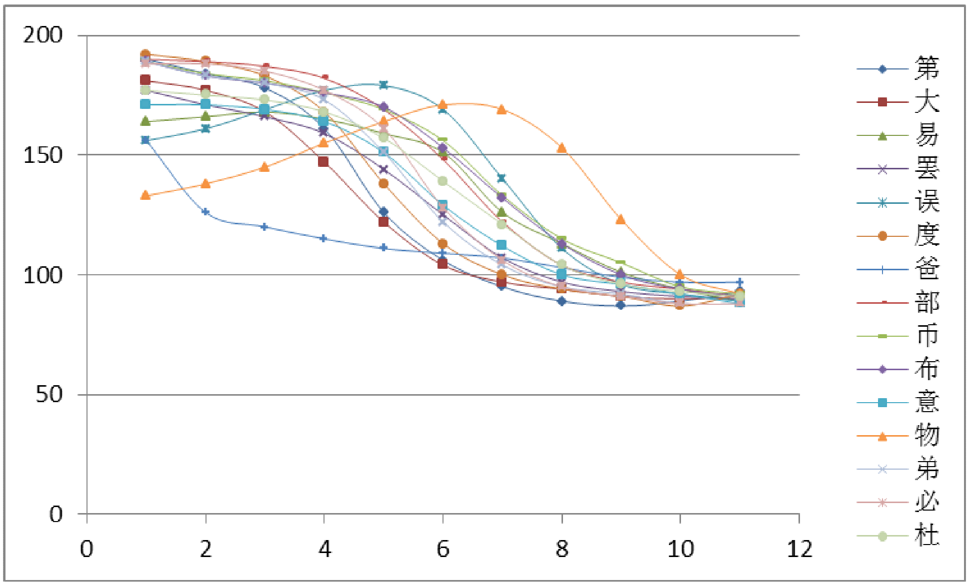


Figure 119. The F0 contour of 15 *Qusheng* syllables pronounced by M5

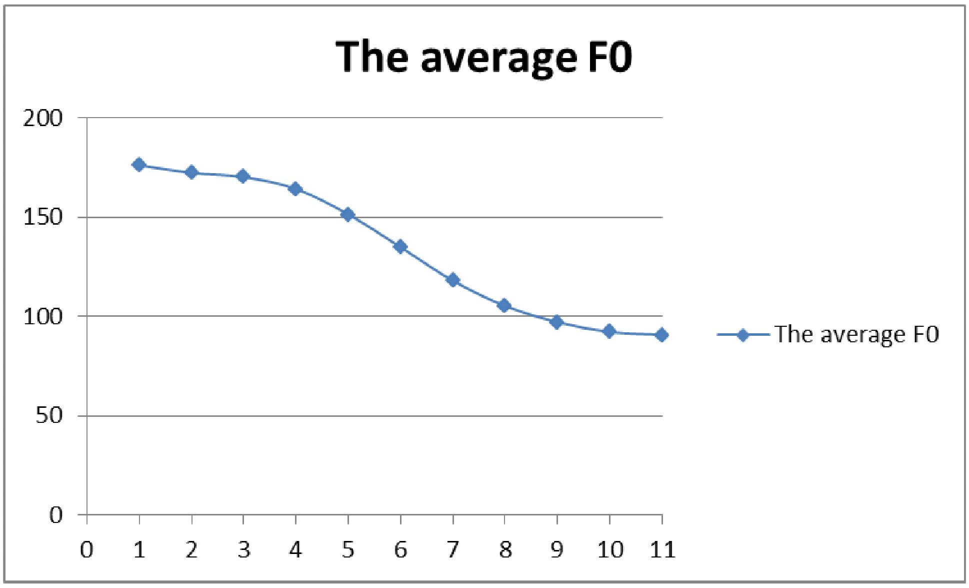


Figure 120. The average F0 contour of 15 *Qusheng* syllables pronounced by M5

Based on Chart 1, the following conclusions can be made. (1) The F0 range of M5 is 80-200Hz, that is, about 120Hz. (2) The F0 contours of M1 are relatively concentrated, except for the contours of 误 wù, 爸 bà and 物 wù. (3) Many F0 contours of M5 have a relatively smooth start, and a quite flat end as well, Again, 误 wù, 爸 bà and 物 wù are the exceptions. (4) Overall, most F0 contours of M5 turn out as declining "S" patterns.

From Chart 2, we can see that the mean F0 contour of M5 presents a declining "S" pattern with a delayed falling start and a relative smooth end, but a noticeably steep part in the middle.

The value of the mean F0 contour ranges between 90 and 180Hz.

(8) FB

Table 85: *FB's F0 values for 15 Qusheng syllables at 11 measuring points*

Syllables	1	2	3	4	5	6	7	8	9	10	11
第 dì	263	251	245	234	221	205	188	176	164	151	132
大 dà	237	236	230	218	204	187	173	159	146	131	124
易 yì	241	237	242	235	223	204	183	170	157	142	129
罢 bà	257	250	242	233	224	211	196	175	153	141	138
误 wù	257	250	242	234	219	201	182	161	145	129	113
度 dù	263	262	254	241	223	202	181	162	147	134	114
爸 bà	252	259	254	243	226	212	198	184	162	150	129
部 bù	250	254	246	231	210	193	173	155	145	132	122
币 bì	256	256	250	236	217	199	185	173	158	139	120
布 bù	261	260	257	244	217	199	184	169	155	149	145
意 yì	226	237	241	234	217	198	181	165	151	138	130
物 wù	236	252	250	240	225	202	175	152	137	127	130
弟 dì	252	254	244	227	209	192	178	166	150	134	119
必 bì	270	248	240	229	213	195	179	167	149	128	108
杜 dù	253	252	242	222	203	192	179	163	144	126	120
Average	251.6	250.5	245.3	233.4	216.7	199.5	182.3	166.5	150.9	136.7	124.9

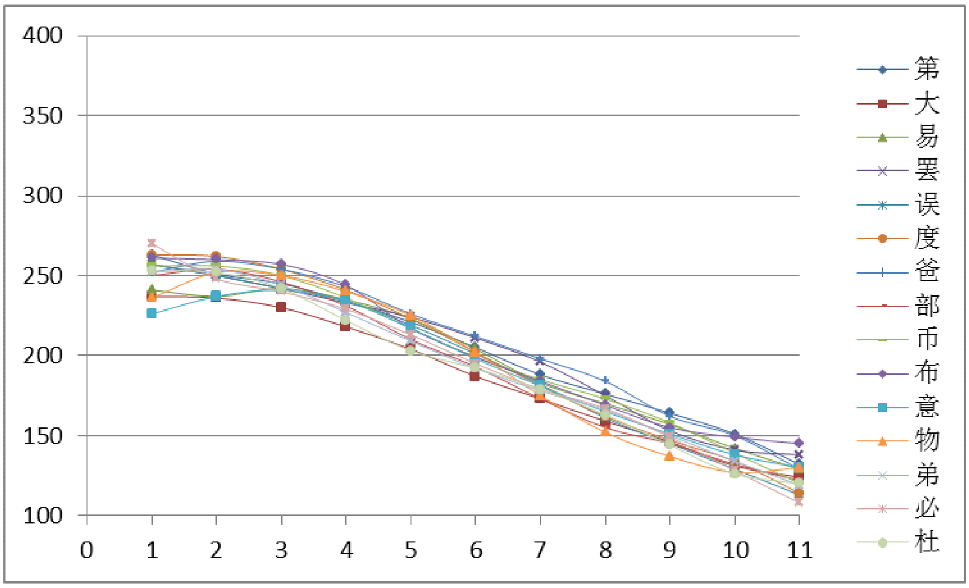


Figure 121. The F0 contour of 15 *Qusheng* syllables pronounced by FB

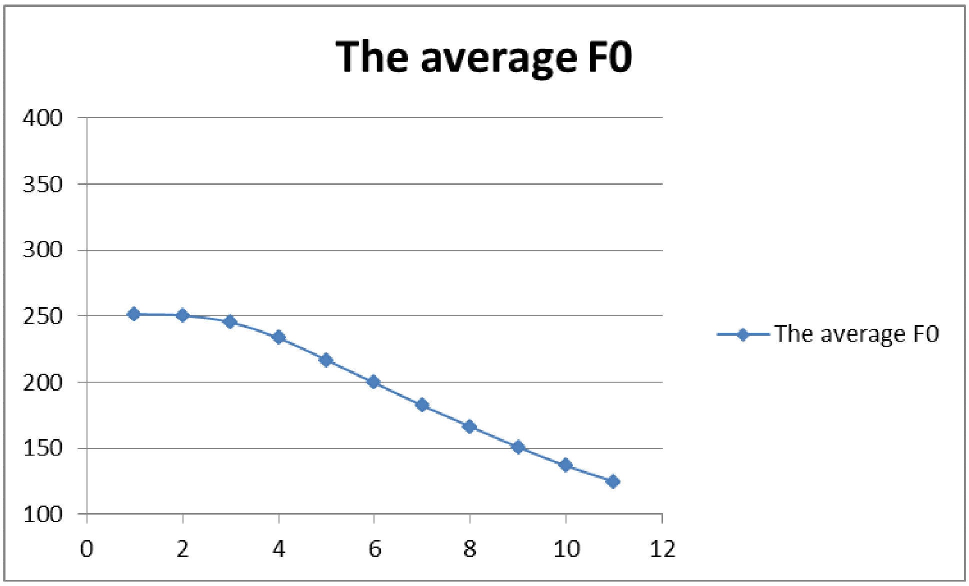


Figure 122. The average F0 contour of 15 *Qusheng* syllables pronounced by FB

Based on Chart 1, the following conclusions can be made. (1) The F0 range of FB is 100-270Hz, that is, about 170Hz. (2) The F0 contours of FB are highly concentrated. (3) Most F0 contours of FB have a considerably smooth start. (4) Overall, the falling parts of most F0 contours are noticeably steep, and almost all F0 contours of FB turn out as falling patterns.

From Chart 2, we can see that the mean F0 contour of FB presents a falling pattern with a delayed falling start. The value of the mean F0 contour ranges between 120 and 260Hz.

(9) MB

Table 86: *MB's F0 values for 15 Qusheng syllables at 11 measuring points*

Syllables	1	2	3	4	5	6	7	8	9	10	11
第 dì	154	153	148	138	126	114	103	97	94	92	83
大 dà	150	147	144	139	134	123	112	105	98	94	89
易 yì	145	153	155	152	147	144	137	125	109	100	92
罢 bà	154	148	143	141	138	132	122	114	107	101	96
误 wù	143	145	147	141	133	125	114	105	98	93	89
度 dù	160	160	159	154	146	132	118	107	97	93	87
爸 bà	149	152	150	149	145	139	131	121	110	105	97
部 bù	166	164	159	147	131	118	107	99	92	84	76
币 bì	156	152	147	138	122	109	98	93	90	84	80
布 bù	152	152	149	143	136	129	121	108	100	94	89

意 yì	158	157	153	146	138	127	116	102	96	90	82
物 wù	153	159	163	158	152	144	129	111	98	92	86
弟 dì	172	172	167	160	147	134	120	106	97	93	88
必 bì	150	149	147	141	126	115	106	100	94	89	86
杜 dù	145	139	135	129	120	110	101	95	91	87	87
Average	153.8	153.5	151.1	145.1	136.1	126.3	115.7	105.9	98.1	92.7	87.1

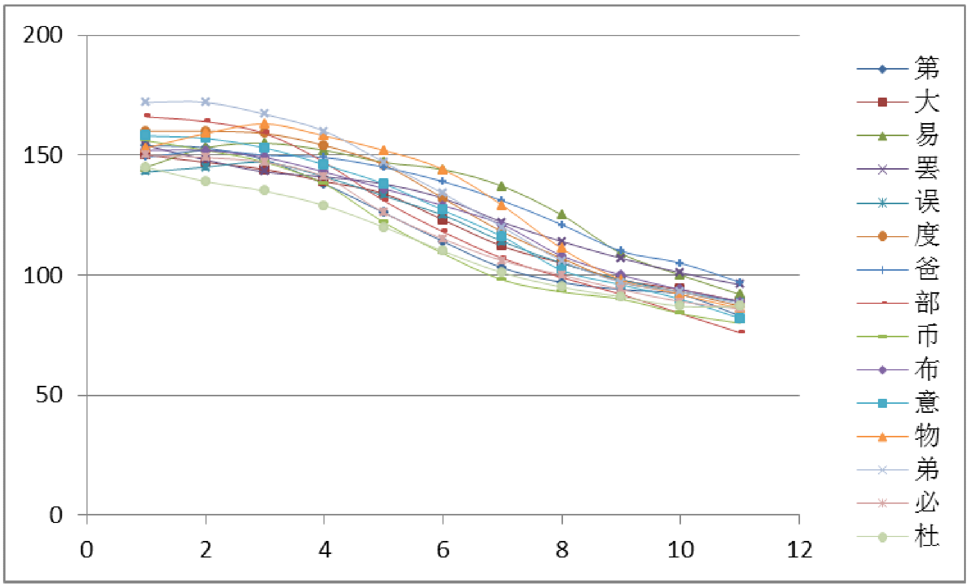


Figure 123. The F0 contour of 15 *Qusheng* syllables pronounced by MB

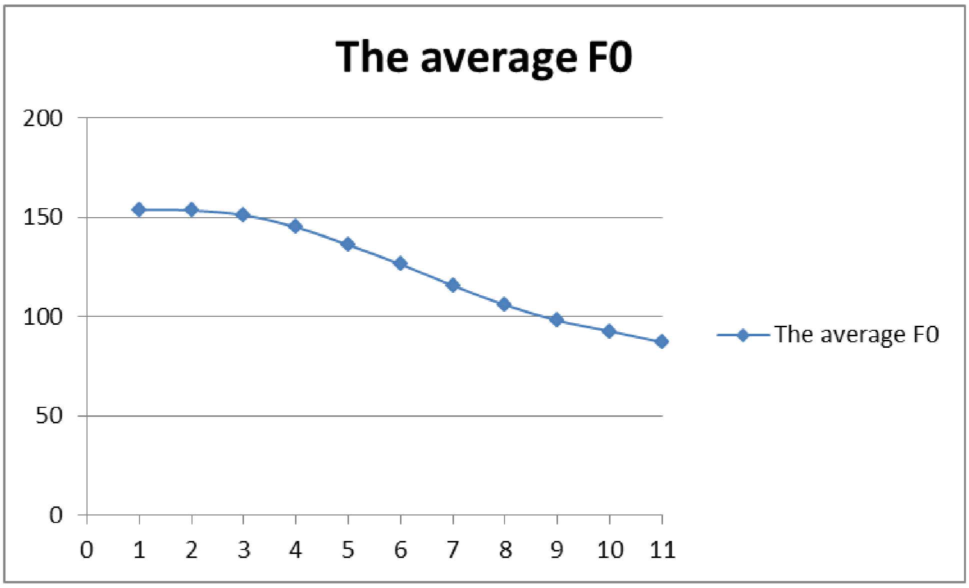


Figure 124. The average F0 contour of 15 *Qusheng* syllables pronounced by MB

Based on Chart 1, the following conclusions can be made. (1) The F0 range of MB is 80-180Hz, that is, about 100Hz. (2) The F0 contours of MB are relatively concentrated. (3) Most F0 contours of MB have a fairly smooth start. (4) Overall, the falling parts of most F0 contours are noticeably steep, and all F0 contours of MB turn out as falling patterns.

From Chart 2, we can see that the mean F0 contour of MB presents a falling pattern with a delayed falling start. The value of the mean F0 contour ranges between 80 and 160Hz.

5.4.2.2. General comments

From the tables and figures presented above, we can see that about 43% of the Hungarian subjects pronounce *Qusheng* syllables with an appropriate falling contour, especially F1 and M1, who can pronounce them perfectly. However, about 53% of the subjects produce deviant contours, mainly due to a considerably flat or rising end, a late inflection point, too gentle falling or “S” pattern.

In order to better compare the 7 Hungarian subjects with the 2 native speakers in general terms, we made a chart for the average F0 contours of *Qusheng* syllables pronounced by 9 subjects.

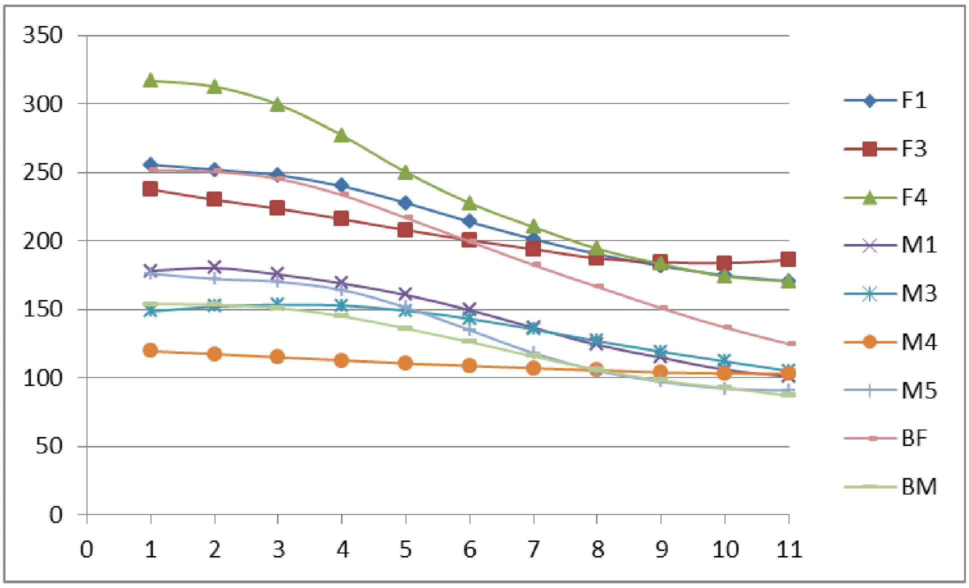


Figure 125. The average F0 contours of *Qusheng* syllables pronounced by 9 subjects

On the basis of the average F0 contours, the following conclusions can be made:

(1) All the Hungarian university students have a falling average F0 contour when pronouncing

Qusheng syllables;

(2) About 43% of the Hungarian subjects pronounce *Qusheng* syllables correctly; in other words, their average F0 contours are appropriate falling patterns;

(3) About 57% of the Hungarian subjects do not pronounce *Qusheng* syllables appropriately, even though most of them do manage to pronounce *Qusheng* syllables with a falling pattern. The deviant contours are as follows: F1's F0 contour displays a falling pattern, but contains an inappropriately flat end; M3 pronounces *Qusheng* syllables with a considerably smooth start, and the inflection point of falling comes inappropriately late; M4's F0 contour has a very gentle falling part followed by an inappropriately flat end; M5's F0 contour has an appropriate start but follows an appropriate end as well, in other words, it contains inappropriate inflection points. Therefore, the F0 contour of M5 turns out as a declining "S" pattern.

(4) The F0 values produced by females are considerably higher than males': nearly twice as high;

(5) Of the 12 Hungarian subjects, only 9 produced entire and regular F0 contours, 4 female and 5 male, which indicates that creaky voice can occur relatively often in *Qusheng* in Hungarian university learners. In what was admittedly a small sample, with data provided voice usually happens in *Shangsheng*. According to the voice samples which were recorded by 12 subjects, it turned out that the creaky voice happens relatively often in *Qusheng*. 3 subjects (one female and two male) in the sample pronounce *Qusheng* syllables with creaky voice frequently. In other words, 25% of these Hungarian university learners produce *Qusheng* syllables with creaky voice.

5.4.3. Intrinsic vowel fundamental frequency

The intrinsic vowel fundamental frequency, shortened for IF0, refers to Lehiste's (1970) observation that F0 of a high vowel is higher than that of a low one when it is pronounced under the same conditions[201]. Zhu Xiaonong believes that the IF0 in high areas of the individual's own frequency domain is more distinct than that in low areas, and the IF0 appears

[201] Lehiste, Ilse 1970: 68.

more obviously in female than in male speakers[202].

For the purpose of defining the relationship between the height of the F0 contour sand vowels, we classified all the *Qusheng* syllables produced by each subject in terms of /i/, /u/ and /a/, and averaged the F0 values of each category. The following table shows the relevant details and the average F0 of all the Hungarian female subjects and the Hungarian male subjects (labeled F, M, respectively).

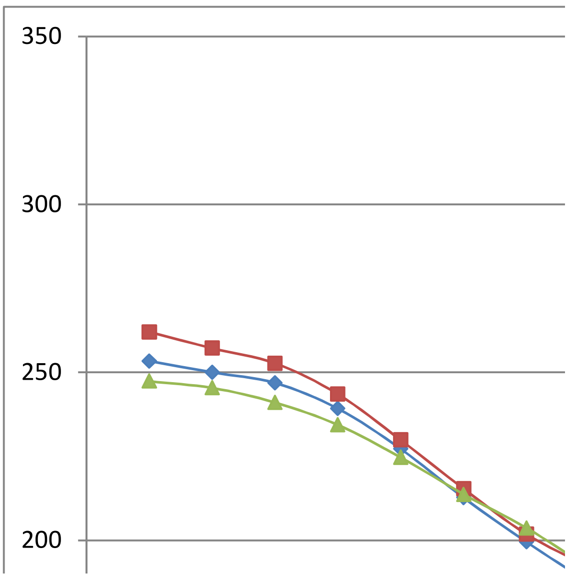
Table 87: *Qusheng F0 of each Hungarian female subject and Hungarian male subject, with averages for all female and all male subjects.*

		1	2	3	4	5	6	7	8	9	10	11	ms
F1	/i/	253.3	250	246.8	239.2	227.2	212.7	199.5	187.8	180.3	174.8	169.7	290
	/u/	262	257.2	252.7	243.5	229.8	215.3	201.8	192	182.7	174.5	170.8	312
	/a/	247.3	245.3	241	234.3	224.7	213.7	203.7	192	182.3	175.3	172.7	252
F3	/i/	230.7	227.3	223.2	217.8	211.2	205.2	200	192.8	189	185.8	185.5	247
	/u/	244.8	236.3	228.8	219	209.3	200.3	193.5	187.8	184.5	186.2	189.8	253
	/a/	237.3	222.7	214	205.7	198.3	191.3	183	175.3	175.3	174.7	179.3	235
F4	/i/	334.8	328.7	314.7	293.2	262.7	238.3	219.8	202	187.8	173.7	170.3	235
	/u/	307.3	308.2	297.8	275.5	250.2	225.5	206.2	189.2	178.7	174.7	172.3	240
	/a/	300	289.3	273	247	224.7	210	198.3	189	182.7	174.3	166.7	240
M1	/i/	184.8	186.5	181.7	174.5	166.2	155.8	142.2	128.2	118.8	110.2	106.2	206
	/u/	184.2	187	182.2	174.8	164.5	150.5	136.2	123.3	113	105.3	99.5	219
	/a/	153	154	150	146.3	141.7	134.7	126.3	117.7	110.3	99	94.33	187
M3	/i/	154.5	158.8	160.8	161	157	150.2	140	128.3	119	111.3	101.5	175
	/u/	147.7	153.5	155.8	155.2	151.8	146	138.7	131.3	124.3	116.5	108.5	168
	/a/	138	136.3	134	132.3	127.7	123.7	119.7	115	109	104.7	105.7	155
M4	/i/	118	116.5	114.3	112.2	109.5	107.8	106.2	104.8	102.7	102.7	103.8	265
	/u/	121.3	118	116.3	114.7	112.7	110.8	108.7	106.8	105.2	103.5	101.8	262
	/a/	119	118.3	113.7	110.3	108.3	106.7	105.7	104.7	104.3	104.7	103.7	277
M5	/i/	181.8	179.3	176.8	169.3	152.8	132	112.7	101.2	95.3	91	89.7	210
	/u/	172.8	172.5	172.8	171	162.8	148.8	130.7	113.2	100.5	93.3	91	223
	/a/	171.3	158	151.3	140.3	125.7	112.7	103.7	98	94.3	92.7	92.7	181
F	/i/	272.9	268.7	261.6	250.1	233.7	218.7	206.4	194.2	185.7	178.1	175.2	257
	/u/	271.4	267.2	259.8	246	229.8	213.7	200.5	189.7	181.9	178.4	177.7	268
	/a/	261.6	252.4	242.7	229	215.9	205	195	185.4	180.1	174.8	172.9	242

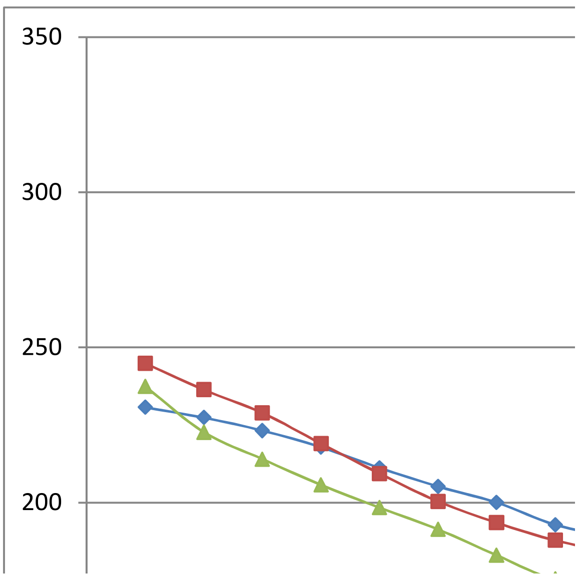
[202] Zhu 2005: 79.

M	/i/	159.8	160.3	158.4	154.3	146.4	136.5	125.3	115.6	109	103.8	100.3	214
	/u/	156.5	157.8	156.8	153.9	148	139	128.5	118.7	110.8	104.7	100.2	218
	/a/	145.3	141.7	137.3	132.3	125.8	119.4	113.8	108.8	104.5	100.3	99.1	200
FB	/i/	251.3	247.2	243.7	232.5	216.7	198.8	182.3	169.5	154.8	138.7	123	312
	/u/	253.3	255	248.5	235.3	216.2	198.2	179	160.3	145.5	132.8	124	280
	/a/	248.7	248.3	242	231.3	218	203.3	189	172.7	153.7	140.7	130.3	255
MB	/i/	155.8	156	152.8	145.8	134.3	123.8	113.3	103.8	96.7	91.3	85.2	303
	/u/	153.2	153.2	152	145.3	136.3	126.3	115	104.2	96	90.5	85.7	295
	/a/	151	149	145.7	143	139	131.3	121.7	113.3	105	100	94	215

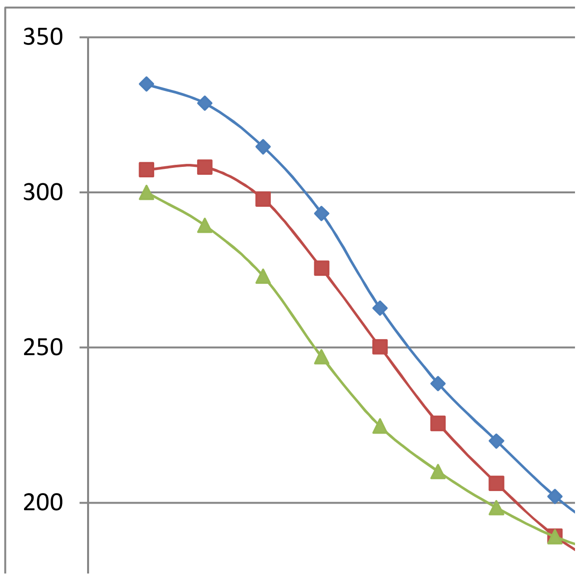
Based on the above table, we made charts of the F0 contours of /i/, /u/ and /a/ for each subject. In each chart the vertical (Y) axis shows the value of fundamental frequency, with Hz as the basic unit of measurement. The horizontal (X) axis presents the 11 measuring points. For the purpose of better comparison, the vertical axis covers a fixed value range based on the maximum and minimum frequencies recorded by the female and male subjects respectively. For the female subjects, the range is from 100Hz to 350Hz; for the male subjects, from 50Hz to 250Hz.



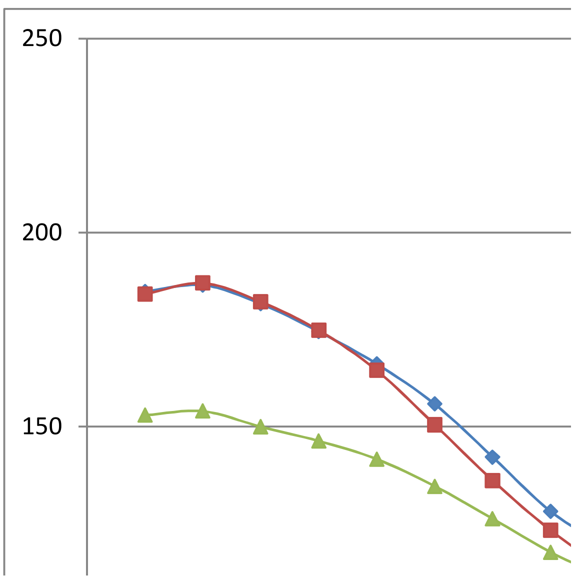
F1



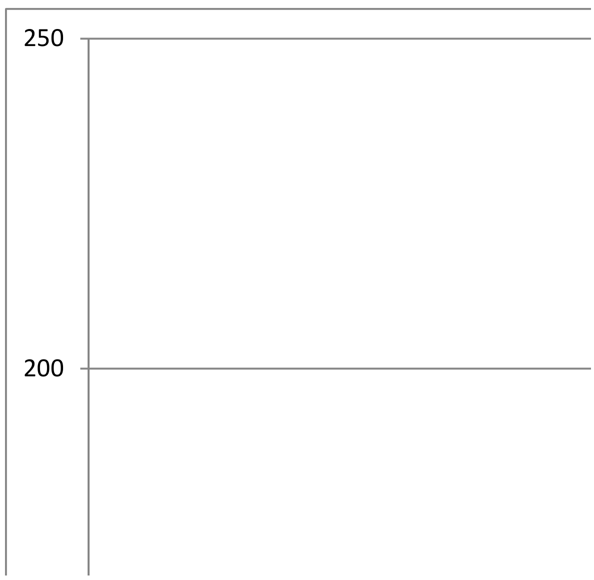
F3



F4



M1



M3

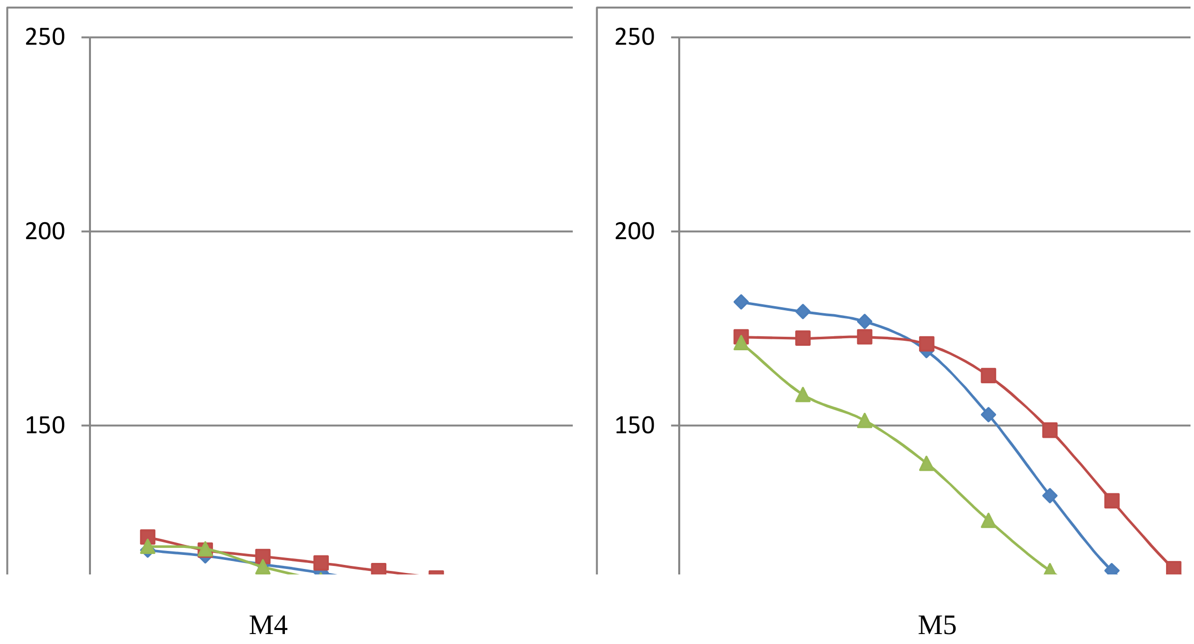


Figure 126. Average *Qusheng* F0 contours of /i/, /u/ and /a/ for each Hungarian subject

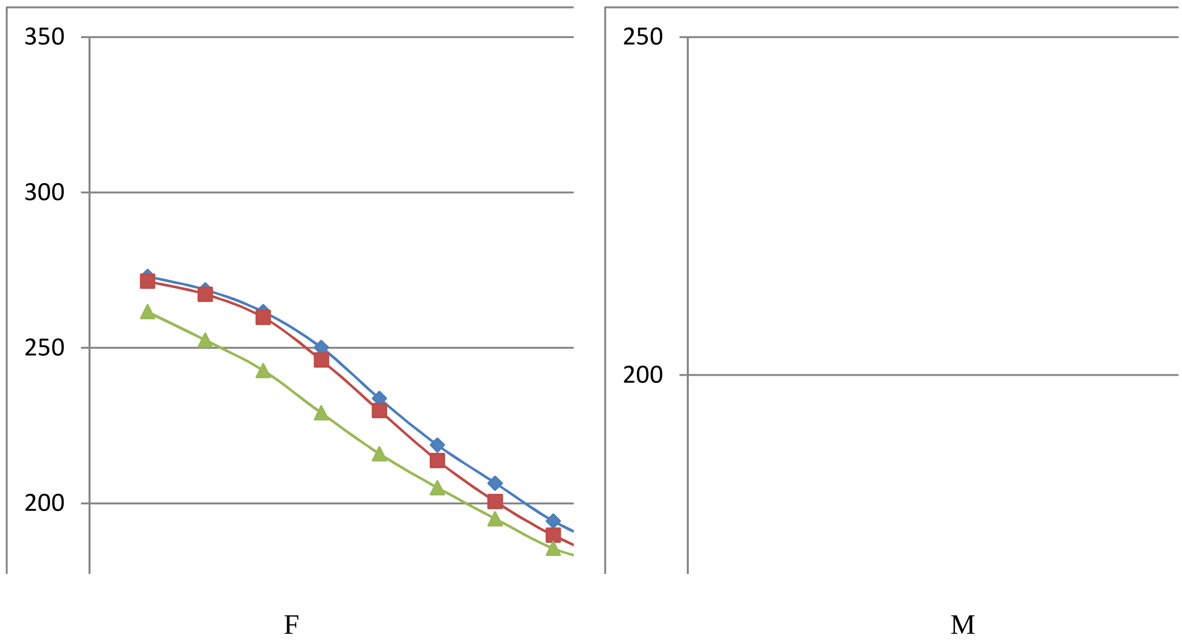
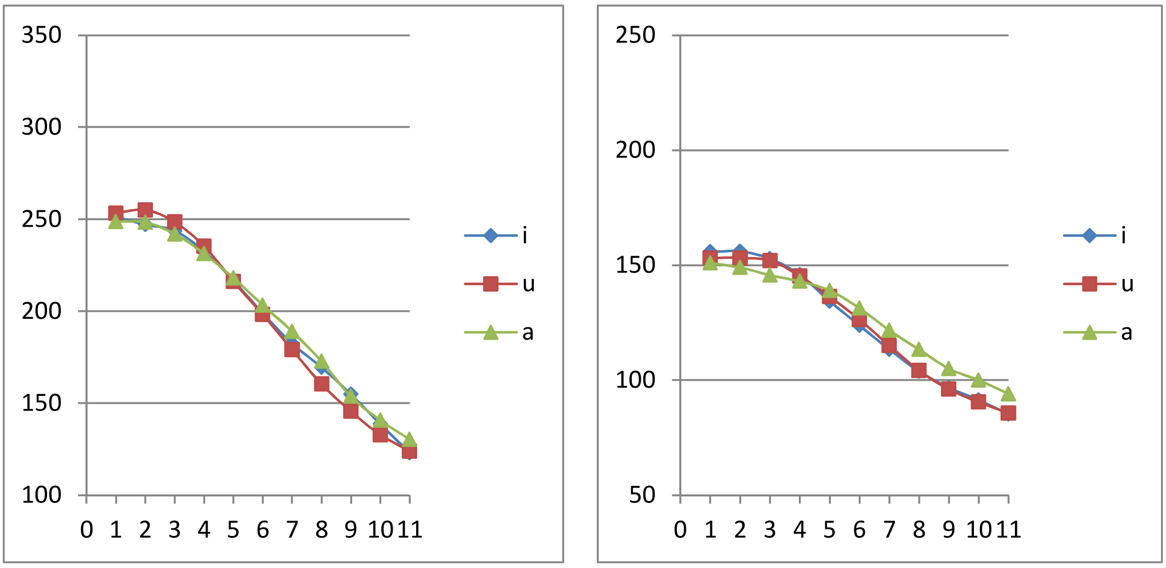


Figure 127. Average *Qusheng* F0 contours of /i/, /u/ and /a/ for Hungarian male and female subjects



FB MB

Figure 128. Average *Qusheng* F0 contours of /i/, /u/ and /a/ for the two Chinese subjects

From the charts presented above, we can see that: (1) The manifestations of the IF0 in *Qusheng* syllables produced by FB and MB are not very obvious, since their average F0 contours of /i/, /u/ and /a/ are quite concentrated. (2) The *Qusheng* syllables produced by all the female subjects and most of the male subjects (except for M4) have distinctly different F0 levels in terms of /i/, /u/ and /a/, with the order of /u/＞/i/＞/a/ or /u/≥/i/＞/a/, especially F4's, M1's, M3's and M5's average F0 contours of /i/ and /u/, which are quite far from their average contours of /a/. (3) The performances of the IF0 in M4 are not by any means obvious, since M4's average F0 contours of /i/, /u/ and /a/ are highly concentrated and almost overlapping. (4) The average F0 contours of /i/, and /u/ produced by all the Hungarian female subjects are almost overlapping and relatively higher that of /a/ produced by them, which indicates the considerably apparent performances of the IF0. And the average F0 contours of /i/, /u/ and /a/ produced by all the Hungarian male subjects are quite similar to those produced by all Hungarian female subjects. In other words, the performances of the IF0 are relatively obvious in F's and M's average F0 contours. (5) Based on those whose average F0 contours reflect the manifestations of IF0 in *Qusheng* syllables, we can see that the IF0 in high areas of the individual's own frequency domain is more distinct than that in low areas, which corresponds to Zhu Xiaonong's statement (See 5.1.3 above)[203].

All in all, the manifestations of the IF0 in the *Qusheng* tone produced by native speaker are not very significant. However, only one of the nine Hungarian subjects produced *Qusheng* tone without showing the obvious manifestations of the IF0, in other words, about 90% of Hungarian learners are likely to produce the *Qusheng* tone accompanied by the quite obvious existence of the IF0. What is more, the IF0 is reflected quite obviously in all the Hungarian female subjects and 75% of the male subjects, which is consistent with Zhu Xiaonong's view; that is, that the IF0 appears more obviously in female speakers than in male speakers.

5.4.4. Range

Range, the area of level or pitch at which sounds are produced, is an indispensable factor in analysis of learners' acquisition of tones. Chao Yuenren (1980) believes that it is in fact the range, and not the tone pattern, that causes the problems when learners try to produce tones[204]. Zhao Jinming (1988) agrees with Chao on this point, as he also found that the most common error concerns range[205]. In other words, learners may produce tone patterns which are appropriate in themselves, but which occur at an inappropriately high or low pitch-range in comparison with some or all of the other three tone patterns. Shen Xiaonan (1989) also showed that for American learners, errors tended to affect range rather than tone patterns. He further points out that errors of range can consist of the entire sound area of a tone occurring in too high or too low a pitch-range[206]. We will discuss this issue within the whole tone system later in this chapter.

For *Qusheng*, the range also indicates how clearly falling the contour is. At this point, the range means the difference between the maximal and minimum value of F0. Therefore, the smaller the range, the less steep the contour of *Qusheng*.

---

[203] Zhu 2005.

[204] Chao 1983.

[205] Zhao 1988: 10-14.

[206] Shen 1989.

Table 88: *The average range of 15 Qusheng syllables pronounced by 9 subjects*

Subjects	Maximal value of average F0	Minimum value of average F0	Range
F1	255.6	170.7	84.9
F3	237.7	186	51.7
F4	316.9	170.4	146.5
M1	180.2	101.1	79.1
M3	153.5	105.1	48.4
M4	119.5	103	16.5
M5	176.1	90.8	85.3
FB	251.6	124.9	126.7
MB	153.8	87.1	66.7

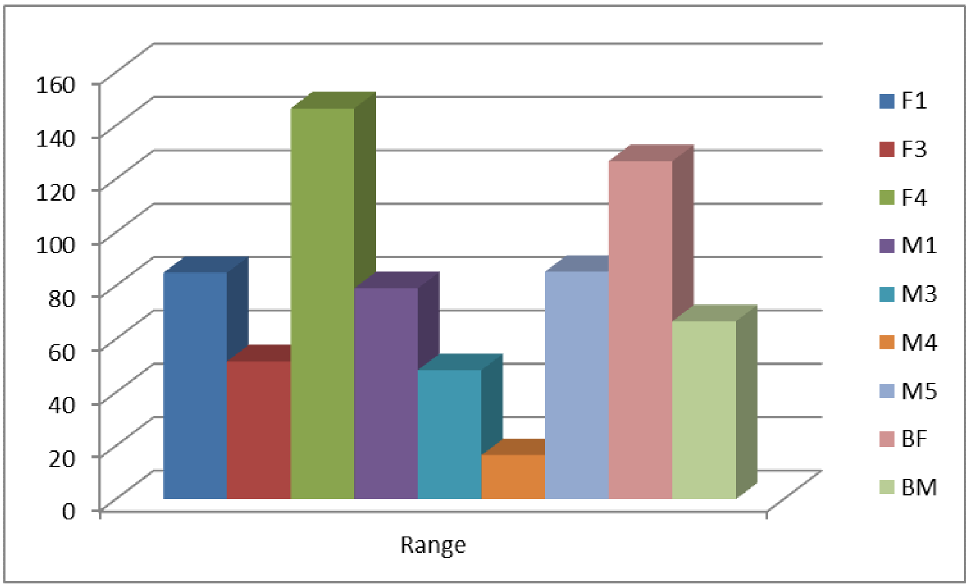


Figure 129. The average range of 15 *Qusheng* syllables pronounced by 9 subjects

From the table and the figure presented above, we can see: (1) the average ranges of FB and MB are 126.7Hz and 66.7Hz, but we should note that average range of FB's is relatively wider than normal due to the existence of the creaky voice; (2) the average ranges of F1, M1, M3, M5 are relatively close to FB's and MB's. However, due to the inappropriate patterns, only the steepness of the *Qusheng* contours pronounced by F1 and M1 subjects are really appropriate; (3) the average range of F4 is much wider than those of FB and MB, more than twice as wide as MB's, whose data were not affected by creaky voice, as FB's were; (4) the average ranges of F3, M4 are considerably narrower than FB's and MB's, which indicates that these four learners do not produce acceptably steep levels in *Qusheng*. What is more, the average range of M4's is very narrow and is only about one fourth of that of MB's. All in all, nearly 60% of Hungarian university learners do not appear to experience problems in pronouncing *Qusheng* in terms of range, but about 40% of them do find it difficult.

5.4.5. Summary

In this section, the recordings of 9 subjects are examined in terms of the duration, the fundamental frequency (F0) and the tone range of their *Qusheng* syllables. On the basis of measurement, extraction, analysis and comparison, the following conclusions can be drawn.

(1) In terms of the duration, about 57% of Hungarian university learners – mainly male subjects in the sample - differ considerably from the "model" native speakers in terms of the duration of *Qusheng* syllables, most often in producing much shorter durations. In other words, the male learners in this sample are more likely have problems with duration. What is more, 30% of Hungarian learners produce deviant orders of durations of /i/, /u/ and /a/, which indicates that the inappropriately produced vowels probably cause problems in duration.

(2) With regard to fundamental frequency (F0), about 43% of the Hungarian subjects pronounce *Qusheng* syllables correctly; in other words, their average F0 contours are appropriate falling patterns; in contrasts, about 57% of the Hungarian subjects have difficulty producing appropriately steep contours. The deviant contours they produce mainly appear as a contour with an inappropriately flat end, a late inflection point, a very gentle falling part, or even a declining "S" pattern. As to the pitch value of F0, it emerges that the male Hungarian subjects are closely grouped in terms of pitch around the native male, while most female Hungarian subjects were higher in pitch than the female native. Concerning intrinsic vowel fundamental frequency, it turned out that the effect of the IF0 is relatively obvious in Hungarian learners when they are producing the *Qusheng* tone, especially in female learners, with 100% in this sample.

(3) Summarizing the findings about pitch range we can say that nearly 60% of these

Hungarian university learners have a fairly appropriate range in *Qusheng*, that is, their ranges are quite close to those of native speakers. However, about 40% of them (one female and two male in the sample) can be considered deviant in the sense that they produce much wider or narrower ranges.

5.5 Individual's own tone system

In this section we will discuss the issues of individual subjects' tone duration, tone-distribution and tone value which are under the individual's own tone system.

5.5.1 Tone duration

As we mentioned before, the duration of syllables bearing different tones is a relative concept. It is hard to define standard or universally "desirable" values, because so much depends on the individual's general rate of speech, not to mention the nature of each specific utterance and the conditions under which it is made. It is, of course, perfectly possible to compare each subject's average syllable durations, for each tone-set of syllables, with the averages of all the other subjects, and in particular with the averages of the native-speaker "models". Therefore, it may be more useful to consider the duration of specific tones within the parameters of an individual's own tone system, paying particular attention to their relative durations. In fact, while it may not be possible to state precisely how long syllables carrying a particular tone must be, it should be possible to say, on the basis of the data provided by the native-speaker subjects, which kind of tones should be shorter or longer than which other kinds. This enables us to establish a sort of implicational scale, placing the tones in proper order of duration, so that each individual speaker, whether they speak slowly or fast, should always complete certain tones with shorter duration than others. It is probably worth mentioning at this point that there is a body of research (e.g. Howie 1976; Rose 1981)[207] which indicates that generally speaking the duration of a rising tone is the longest, while the duration of a falling tone is the shortest, and the duration of an even tone occupies an intermediate position

[207] Howie, John Marshall 1976; Rose, Phil 1981.

between the rising and the falling tone.

The average durations of each subject's pronunciation of each of the four tones have been presented in the relevant section above. In this section, we will further analyze and compare the average duration of each tone produced by 12 subjects on the basis of individual's own tone system. The following table shows the average durations of each tone syllable pronounced by 12 subjects. The unit of duration is the millisecond (ms). The last column in the table shows the subject's "duration order", with the longest tone first and the shortest tone last. Thus, for example, from F1's recorded data it emerges that her T2 (*Yangping*) syllables have the longest duration, her T3 (*Shangsheng*) syllables have the next longest duration, followed by her T1 (*Yinping*) syllables, and her T4 (*Qusheng*) syllables come at the end of the scale because they have the shortest duration.

Table 89: *Average duration of four tones produced by 12 subjects*

	Yinping (T1)	*Yangping* (T2)	*Shangsheng* (T3)	*Qusheng* (T4)	The order
F1	415	452	445	290	T2>T3>T1>T4
F2	492	462	574	312	T3>T1>T2>T4
F3	382	285	627	177	T3>T1>T2>T4
F4	515	367	697	238	T3>T1>T2>T4
M1	283	313	426	207	T3>T2>T1>T4
M2	348	437	463	191	T3>T2>T1>T4
M3	281	231	301	168	T3>T1>T2>T4
M4	352	330	481	267	T3>T1>T2>T4
M5	597	545	721	210	T3>T1>T2>T4
M6	345	322	310	388	T4>T1>T2>T3
FB	421	451	645	288	T3>T2>T1>T4
MB	479	521	599	282	T3>T2>T1>T4

On the basis of the above table, the following conclusions can be made:

(1) Within the parameters of FB and MB individual's own tone system, basically, the average durations of T4 are the shortest, while the average durations of T3 are the longest. The

average durations of T2 are longer than those of T1. The orders of the average durations of four tones produced by FB and MB are the same, that is, T3>T2>T1>T4, which suggests that this can be taken as the normal or desirable order.

(2) Among the Hungarian learners in the sample, only M1 and M2 have the same order of the average durations of four tones as the "model" native speakers, that is, T3>T2>T1>T4.

(3) Under the individual's own tone system, the average durations of *Qusheng* syllables pronounced by F2, F3, F4, M3, M4, and M5 are the shortest, while the average durations of *Shangsheng* syllables pronounced by these six subjects are the longest. At this point, the average durations of *Qusheng* and *Shangsheng* syllables pronounced by F2, F3, F4, M3, M4, and M5 are in accordance with those of *Qusheng* and *Shangsheng* pronounced by the "model" native speakers. However, the average duration of *Yinping* is longer than that of *Yangping* which is in contrast with those of FB's and MB's.

(4) M6's individual tone system is very special in terms of the average durations of four tones. The order of the average durations of four tones pronounced by M6 is T4>T1>T2>T3, which is completely in contrast with those of the "model" native speakers.

(5) The order of the average durations of four tones pronounced by F1 is T2>T3>T1>T4, in other words, the average duration of *Yangping* produced by F1 is longer than that of *Shangsheng* which is in contrast with those of the "model" native speakers.

All in all, under individual's own tone system, of 10 Hungarian subjects, only M1 and M2 produced appropriate orders of the average duration of four tones. In other words, about 80% Hungarian learners differ considerably from the "model" native speakers in terms of the durations of four tones, most often in producing *Yinping* longer than *Yangping*. This in itself represents a notable difference between our non-native learners and the native-speaker "models", sufficient in itself to warrant pedagogical attention. As we have seen, the absolute length of syllables is a less reliable indicator or predictor of problems; nonetheless, it seems worth noting that the percentage of this sample of Hungarian university learners who differ considerably from the "model" native speakers in terms of the absolute duration of *Yinping*, *Yangping*, *Shangsheng* and *Qusheng* syllables are 60%, 60%, 70%, 40%, respectively .

5.5.2 Tone-distribution

As we mentioned before, normalization serves to eliminate the random interpersonal differences so as to extract constant parameters, that is, eliminating personal characteristics and obtaining linguistically meaningful information (Rose 1993)[208]. Zhu Xiaonong (2010)[209] states that the primary role that normalization plays is to describe the tone based on standardized quantitative descriptions; the other role is to reduce the differences between pronunciation styles such as the formal, the casual and the nervous ones. Normalization makes it possible to find the constant in the interpersonal differences and find the commonalities in the interlanguage variations, thus making research on interpersonal and interlingual comparisons possible. For the normalization procedures followed in the present investigation, see 4.3.2.2.

Zhu Xiaonong (2010)[210] suggests that a LZ normalized chart of tones distribution can be made by normalizing each subject's data, then averaging the normalization results. Therefore, for the purpose of better comparison, we will present the normalization result of each subject, and make a chart of each subject's individual tone system based on the results of normalization. Each chart will be followed by some relevant comments. Note that in order to enable the optimal usage of the "model" native speakers' individual tone system as references, in this section, we will present FB's and MB's data before the Hungarian subjects' data, rather than afterwards. Although the creaky voice phenomenon occurred very frequently in *Shangsheng* syllables produced by FB, and as a result the central section of the relevant F0 contour could not be satisfactorily measured, the start and the end, which are the linguistic goals of *Shangsheng* could still be measured, therefore we keep FB's data in this section.

The concept and the nature of the linguistic goal of tone were introduced and discussed at some length previously. At this point it will suffice to recall that for *Yinping*, the start, the body and the end are the linguistic goals; for *Yangping*, the linguistic goal of tone lies in the peak point of the end tail; for *Shangsheng*, it is the start and the end; for *Qusheng*, is at the

[208] Rose, Phil 1981.
[209] Zhu 2010.
[210] Zhu 2010: 291.

beginning or 5%-20% in, at the adjacent peak point of the start and the body. Therefore, in this section, we focus on the linguistic goals when we are analyzing the individual's tone system.

With regard to tone value system, Chao divides an individual's F0 domain into five points, conventionally numbered from 1 to 5, which corresponds to a scale from the lowest pitch of the individual's F0 domain to the highest, and Chao describes *Yinping* as 55, *Yangping* as 35, *Shangsheng* as 214 and *Qusheng* as 51. (See 3.4.1) For the purpose of better description of the tone value, we divided each subject's F0 domain into five degrees, which correspond to Chao's tone value system. Note that due to the creaky voice phenomenon, there are only 6 Hungarian subjects who could be used in this section. The following tables and charts show the normalized versions of these 6 subjects. We also make a tone-distribution chart of the average data which are produced by all the Hungarian subjects.

1. The individual tone system of FB and MB

Table 90: *LZ-normalized version of FB's data*

	0%	10%	20%	30%	40%	50%	60%	70%	80%	90%	100%	ms
T1	0.820243	0.844948	0.832611	0.821619	0.794023	0.784327	0.77322	0.77461	0.795407	0.809221	0.816112	413.1333
T2	-0.63266	-0.61862	-0.64473	-0.71781	-0.76098	-0.60262	-0.29548	0.059154	0.476938	0.777983	0.961793	455.5
T3	-0.66593	-0.86299	-1.38382	undefined	undefined	undefined	undefined	undefined	-1.75054	-0.70862	-0.18187	644.875
T4	1.287478	1.253721	1.139413	0.891261	0.532236	0.137106	-0.2961	-0.73489	-1.20199	-1.72075	-2.29599	260.2667

On the basis of the above table, we made a tone-distribution chart of FB which is presented below.

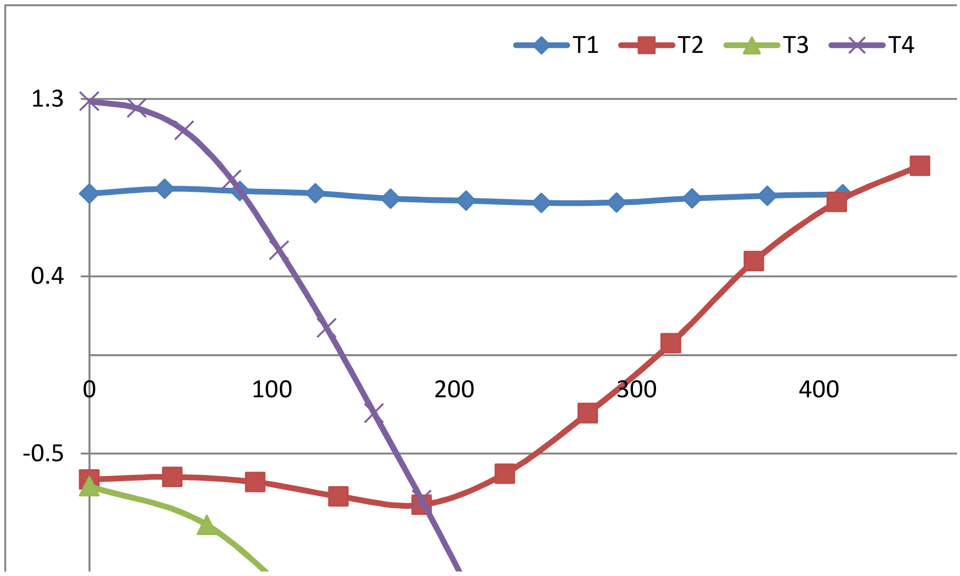


Figure 130. FB's individual tone system based on the results of LZ-normalization

Table 91: *LZ-normalized version of MB's data*

	0%	10%	20%	30%	40%	50%	60%	70%	80%	90%	100%	ms
T1	0.364997	0.359304	0.353604	0.350753	0.364997	0.336471	0.307799	0.307799	0.284756	0.267411	0.226728	479.2
T2	-0.85168	-1.00139	-1.01698	-0.9549	-0.79139	-0.56734	-0.215	0.173564	0.595734	0.952737	1.249432	520.8571
T3	-1.08962	-1.55451	-1.88337	-1.92336	-1.93139	-1.88735	-1.81603	-1.66072	-1.37335	-1.08962	-0.89923	607.75
T4	1.252555	1.2404	1.152093	0.925036	0.566206	0.150383	-0.34382	-0.83982	-1.26859	-1.58188	-1.93085	281.6667

On the basis of the above table, we made a tone-distribution chart of MB which is presented below.

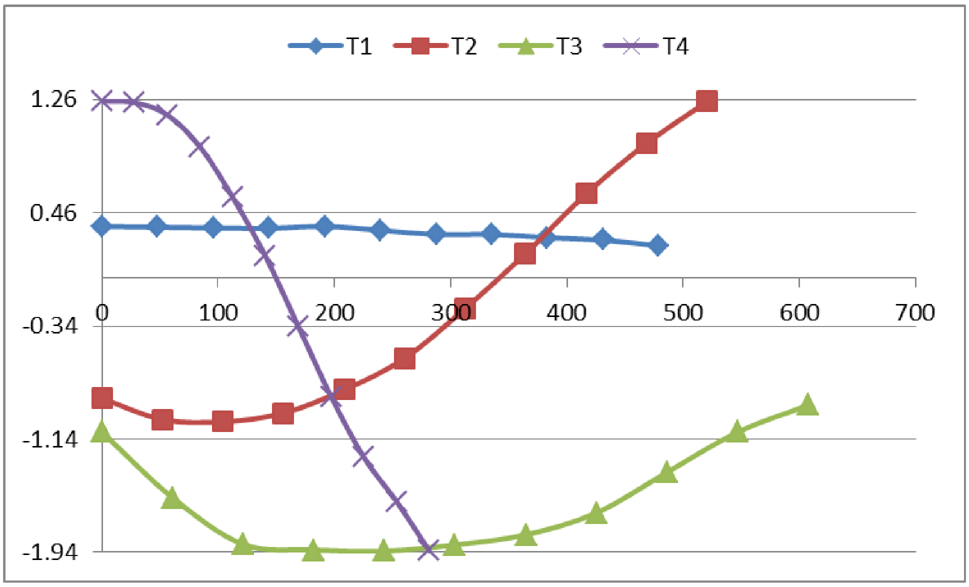


Figure 131. MB's individual tone system based on the results of LZ-normalization

On the basis of FB's and MB's LZ normalized individual tone system, the following conclusions can be made:

(1) The F0 domains of FB and MB are (-2.3, 1.3) and (-1.94, 1.26), respectively. In other words, the F0 domain of the female native speaker is relatively wider than that of the male native speaker.

(2) The start of the *Qusheng* is the highest point in both FB's and MB's individual F0 domain, which is in consistent with Chao's view; while the tone values of *Yinping* in "model" native speakers' F0 domain are considerably lower than those of *Qusheng* start. In other words, the tone value of *Yinping* is not as high as that of *Qusheng* start in both native speakers' F0 domain. At this point, we cannot help noting a striking contrast with Chao, who defines *Yinping* as 55[211].

(2) The tone value of the *Yangping* end is higher than the *Yinping* value and close to the value of the *Qusheng* start in both FB's and MB's individual F0 domain.

(3) The start of *Shangsheng* is quite close to that of *Yangping* in both FB's and MB's individual F0 domain, especially FB's which are almost overlapping. However, Chao indicates that the start of *Yangping* and *Shangsheng* are 3 and 2, respectively[212]. The tone value of the *Shangsheng* end in both "model" native speakers' F0 domain is considerably lower than 4, in other words, is considerably lower than their *Yinping*, especially MB's, which is even lower than mid pitch 3.

(4) FB's F0 of each tone is relatively higher than MB's.

2. The individual tone system of F1

Table 92: *LZ-normalized version of F1's data*

	0%	10%	20%	30%	40%	50%	60%	70%	80%	90%	100%	ms
T1	0.986487	0.900556	0.889747	0.870253	0.852884	0.85071	0.844185	0.831117	0.833297	0.859402	0.865915	415.0667
T2	-0.72118	-1.15856	-1.30182	-1.30797	-1.26499	-1.0954	-0.87394	-0.56857	-0.2211	0.119123	0.321294	451.7857
T3	-0.37149	-0.7471	-1.20116	-1.58164	-1.79289	-1.84199	-1.63541	-1.29526	-0.66204	-0.16052	0.202484	444.7333
T4	1.424631	1.312791	1.190993	0.935043	0.531115	0.047269	-0.42514	-0.85746	-1.21818	-1.51642	-1.69862	290

[211] Chao 1983.

[212] Chao (ibid).

On the basis of the above table, we made a tone-distribution chart of F1 which is presented below.

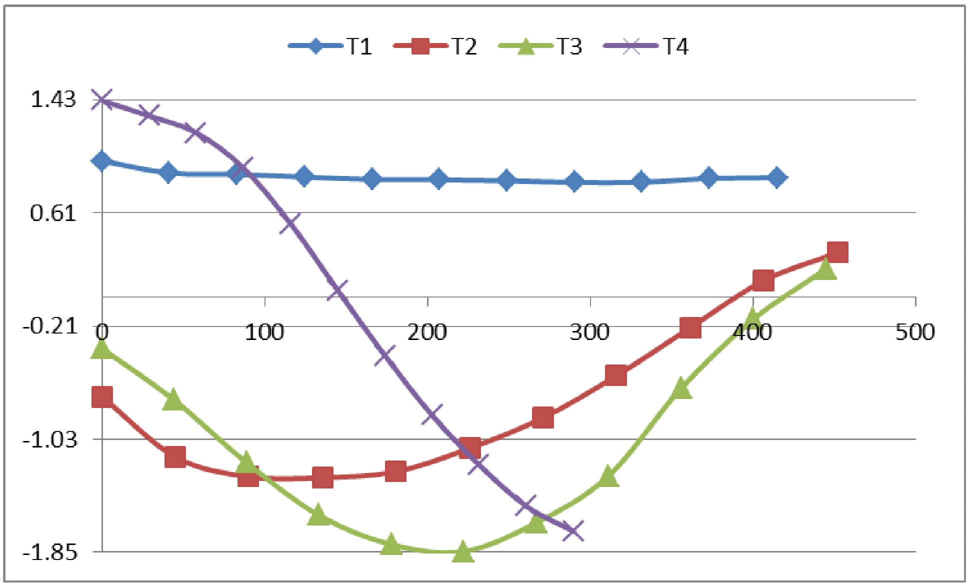


Figure 132. F1's individual tone system based on the results of LZ-normalization

Based on the chart presented above, we can see that:

(1) The F0 domain of F1 is (-1.85, 1.43) which is quite close to the "model" native speakers'.

(2) The start of *Qusheng* is the highest point in F1's individual F0 domain, while the tone value of *Yinping* is considerably lower than that of *Qusheng* start.

(3) The tone value of *Yangping* end is somewhat lower than that of *Yinping*.

(4) The start of *Shangsheng* is relatively close to that of *Yangping*, and the end of *Shangsheng* is considerably lower than *Yinping*.

3. The individual tone system of F4

Table 93: L*Z-normalized version of F4's data*

	0%	10%	20%	30%	40%	50%	60%	70%	80%	90%	100%	ms
T1	1.096169	1.034982	1.015754	1.024872	1.030941	1.042045	1.03902	1.049098	1.050105	1.067183	1.087193	514.7333
T2	-0.33469	-0.77383	-1.07029	-1.15192	-1.13268	-1.0514	-0.88481	-0.5597	0.029343	1.045141	1.754271	367.2143
T3	0.865807	0.246567	-0.6156	-0.96376	-1.30745	-1.31857	-1.29161	-1.09056	-0.61423	0.253321	1.213754	698.5
T4	1.292135	1.230652	1.038011	0.680122	0.218392	-0.20988	-0.57212	-0.92675	-1.19441	-1.42122	-1.52124	237.8667

On the basis of the above table, we made a tone-distribution chart of F4 which is presented below.

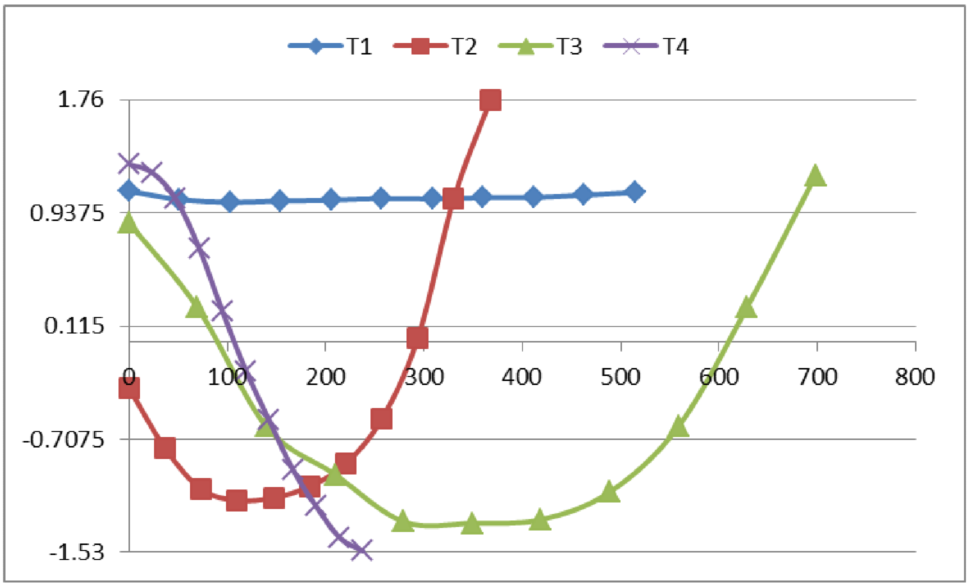


Figure 133. F4's individual tone system based on the results of LZ-normalization

Based on the chart presented above, we can see that:

(1) The F0 domain of F4 is (-1.53, 1.76) which is quite close to the "model" native speakers'.

(2) The start of *Qusheng* is not the highest point in F4's individual F0 domain, and the tone value of *Yinping* is relatively lower than that of the *Qusheng* start.

(3) The tone value of *Yangping* end is considerably higher than that of the *Yinping* and the *Qusheng* start; in other words, the end of Yangping is the highest point in F4's individual F0 domain.

(4) The start of *Shangsheng* is very close to the LZ-normalized F0 contour of *Yinping*, and the end of *Shangsheng* is considerably lower than *Yinping*.

4. The individual tone system of M1

Table 94: *LZ-normalized version of M1's data*

	0%	10%	20%	30%	40%	50%	60%	70%	80%	90%	100%	ms
T1	-0.23832	-0.21182	-0.23567	-0.23036	-0.21446	-0.22771	-0.22506	-0.20125	-0.20918	-0.17756	-0.13568	283.0667
T2	-1.34179	-1.5366	-1.70501	-1.77235	-1.76107	-1.59095	-1.45429	-1.09678	-0.70507	-0.2332	0.180246	312.6429
T3	-1.24948	-1.56688	-1.87684	-1.92742	-1.92742	-1.89365	-1.7311	-1.45696	-1.21367	-0.86783	-0.22042	425.8125

T4	1.173515	1.234802	1.090719	0.882431	0.602473	0.207958	- 0.28635	-0.81188	-1.2411	-1.67905	-1.93714	207.2

On the basis of the above table, we made a tone-distribution chart of M1 which is presented below.

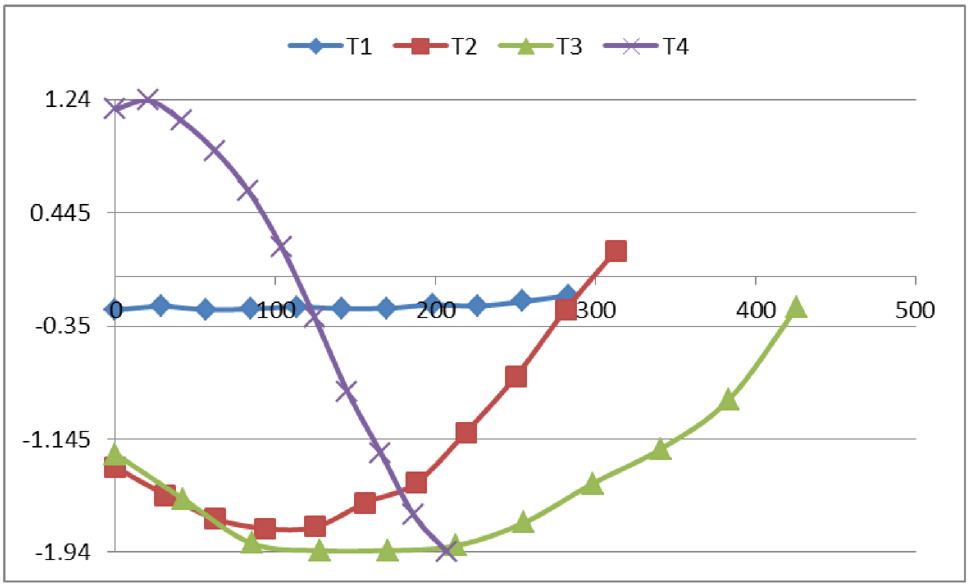


Figure 134. M1's individual tone system based on the LZ-normalization result

Based on the chart presented above, we can see that:

(1) The F0 domain of M1 is (-1.94, 1.24) which is quite close to the "model" native speakers'.

(2) The start of *Qusheng* is the highest point in M1's individual F0 domain, while the tone value of *Yinping* is considerably lower than that of the *Qusheng* start.

(3) The tone value of the *Yangping* end is relatively higher than that of *Yinping*, but is considerably lower than that of the *Qusheng* start.

(4) The start of *Shangsheng* is quite close to that of *Yangping* and almost overlapping, and the end of *Shangsheng* is very close to *Yinping*.

5. The individual tone system of M3

Table 95: *LZ-normalized version of M3's data*

	0%	10%	20%	30%	40%	50%	60%	70%	80%	90%	100%	ms
T1	-0.86813	-1.00384	-0.96972	-0.92604	-0.9212	-0.89221	- 0.82973	- 0.85851	-0.78672	-0.59802	-0.50975	280.8
T2	-1.44932	-1.68324	-1.74552	-1.73984	-1.68888	-1.46032	- 0.97285	- 0.29506	0.680312	1.724097	2.365478	230.6429
T3	-1.60459	-1.83761	-1.91326	-1.89808	-1.84766	-1.72778	-1.3346	-0.4358	0.937014	2.310421	3.18829	300.75
T4	0.695634	0.919281	0.993917	0.962567	0.731954	0.370376	-0.134	- 0.72022	-1.2866	-1.83727	-2.41239	168.0667

On the basis of the above table, we made a tone-distribution chart of M3 which is presented below.

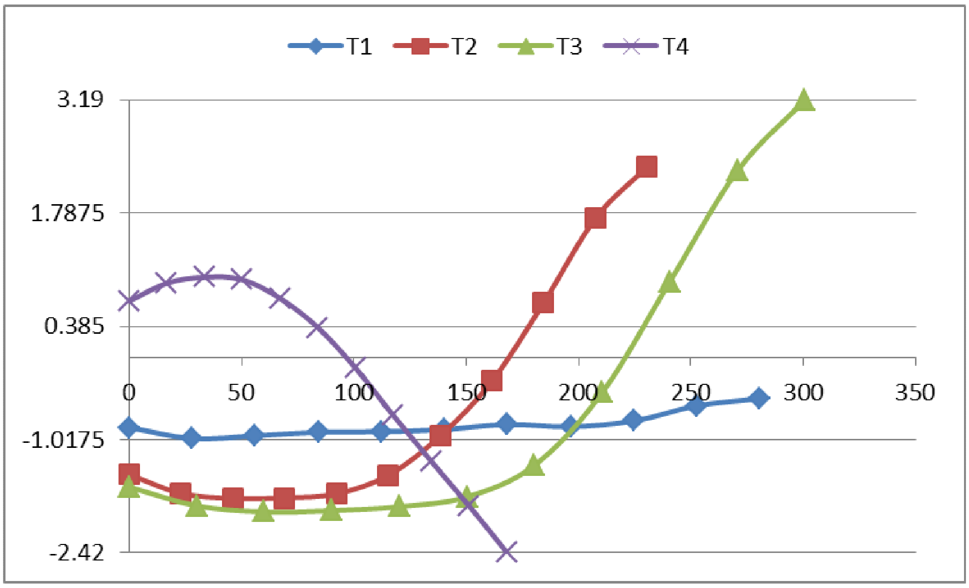


Figure 135. M3's individual tone system based on the result of LZ-normalization

Based on the chart presented above, we can see that:

(1) The F0 domain of M3 is (-2.42, 3.19), which is much wider than that of the "model" native speakers.

(2) The start of *Qusheng* is not the highest point in M3's individual F0 domain, and the tone value of *Yinping* is considerably lower than that of the *Qusheng* start.

(3) The tone value of the *Yangping* end is considerably higher than that of *Yinping* and indeed much higher than the start of *Qusheng*, in other words, the end of *Yangping* is the highest

point in M3's individual F0 domain.

(4) The start of *Shangsheng* is quite close to that of *Yangping* and almost overlapping, and the end of *Shangsheng* is very high and even higher than *Yangping* which is quite inappropriate.

6. The individual tone system of M4

Table 96: *LZ-normalized version of M4's data*

	0%	10%	20%	30%	40%	50%	60%	70%	80%	90%	100%	ms
T1	4.796673	4.359323	4.346712	4.447416	4.459975	4.422279	4.283563	4.245595	4.182185	4.232926	4.434851	352.3333
T2	3.636509	3.384848	3.286281	3.398897	3.552908	3.650415	3.580806	3.497015	3.398897	3.440995	3.747539	330.2857
T3	3.169604	2.503252	1.89687	1.059915	-0.18733	-1.02591	-1.23268	-0.47328	1.167784	3.268638	5.238562	481.125
T4	2.172236	1.741379	1.188405	0.754431	0.25317	-0.15229	-0.53375	-0.8744	-1.07793	-1.29901	-1.20402	266.4667

On the basis of the above table, we make a tone-distribution chart of M4 which is presented below.

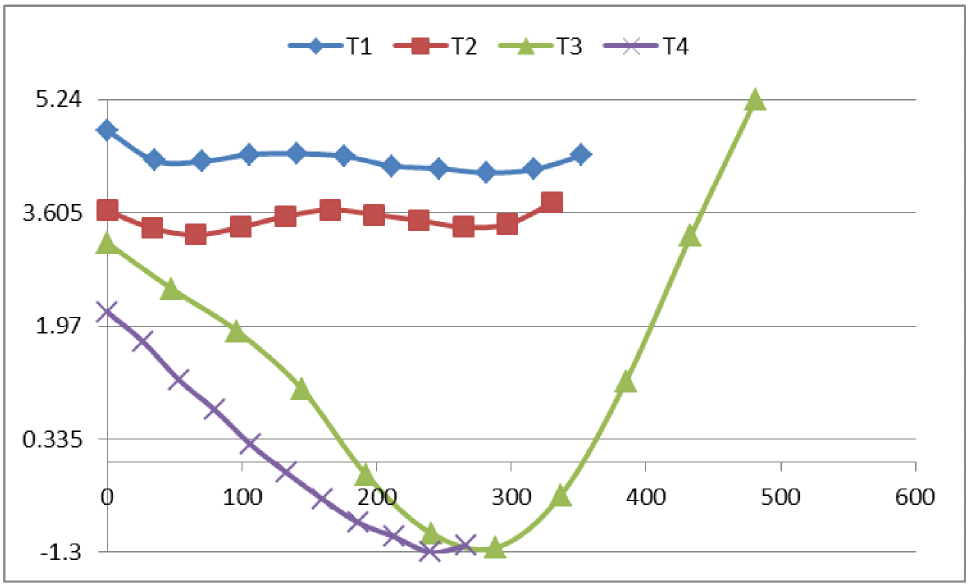


Figure 136. M4's individual tone system based on the results of LZ-normalization

Based on the chart presented above, we can see that:

(1) The F0 domain of M4 is (-1.3, 5.24), which is much wider than that of the "model" native speakers, indeed, almost twice as wide.

(2) The start of *Qusheng* is not the highest but actually the lowest of the starts in M4's individual F0 domain, and the tone value of *Yinping* is considerably higher than that of the

Qusheng start.

(3) The tone value of the *Yangping* end is very close to that of its start, and the LZ-normalized F0 contour is fairly flat and almost parallel to that of *Yinping*.

(4) The start of *Shangsheng* is quite high; higher than the mid pitch, and the end of *Shangsheng* is very high indeed: it is even higher than *Yinying*.

(5) To sum, the LZ-normalized tone system of this subject is very idiosyncratic and largely inappropriate.

7. The individual tone system of M5

Table 97: *LZ-normalized version of M5's data*

	0%	10%	20%	30%	40%	50%	60%	70%	80%	90%	100%	ms
T1	1.027544	1.119415	1.19206	1.235647	1.252642	1.2511	1.249558	1.26342	1.269566	1.275703	1.277236	596.9333
T2	-0.60095	-0.9079	-1.03696	-1.10455	-1.12237	-1.09567	-0.96761	-0.6882	-0.21701	0.356191	0.823212	544.8571
T3	-0.53126	-0.84698	-0.97048	-1.03878	-1.07208	-1.09013	-1.0618	-0.95046	-0.63608	-0.05446	0.552975	721.25
T4	1.292531	1.204562	1.152742	1.009571	0.68223	0.215826	-0.32075	-0.78103	-1.10514	-1.31522	-1.37985	209.5333

On the basis of the above table, we made the tone-distribution chart of M5 presented below.

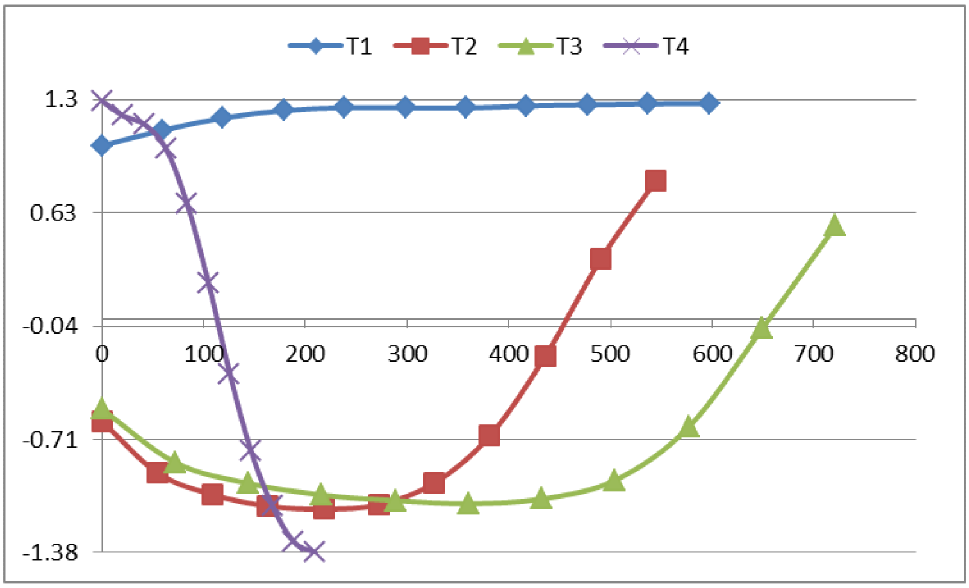


Figure 137. M5's individual tone system based on the results of LZ-normalization

Based on the chart presented above, we can see that:

(1) The F0 domain of M5 is (-1.38, 1.3) which is considerably narrower than that of the

"model" native speakers.

(2) The start of *Qusheng* is the highest point in M5's individual F0 domain, and the tone value of *Yinping* is very close to the *Qusheng* start.

(3) The tone value of the *Yangping* end is relatively close to that of *Yinping* and to the start of *Qusheng*.

(4) The start of *Shangsheng* is quite close to that of *Yangping* and almost overlapping, and the end of *Shangsheng* is relatively lower than that of *Yangying*.

8. The tone system of the average LZ-normalized data of 6 Hungarian subjects

On the basis of the above tables, we average the LZ-score normalized versions of 6 subjects' data, the result is presented in the following table.

Table 98: *the average Z-score normalized version of 6 subjects' data*

	0%	10%	20%	30%	40%	50%	60%	70%	80%	90%	100%	ms
T1	0.910988	0.825302	0.831244	0.861093	0.869229	0.862744	0.849578	0.846518	0.846663	0.902483	0.966149	479.2
T2	-0.13524	-0.44588	-0.59555	-0.61295	-0.56951	-0.44055	-0.26212	0.048117	0.494229	1.075391	1.532007	520.8571
T3	0.046432	-0.37479	-0.78008	-1.05829	-1.3558	-1.48301	-1.3812	-0.95039	-0.1702	0.791597	1.695941	607.75
T4	1.34178	1.273911	1.109131	0.870694	0.503222	0.079877	-0.37869	-0.82862	-1.18723	-1.51137	-1.69221	281.6667

Based on the above table, we made a tone-distribution chart of these six subjects, see the chart presented below.

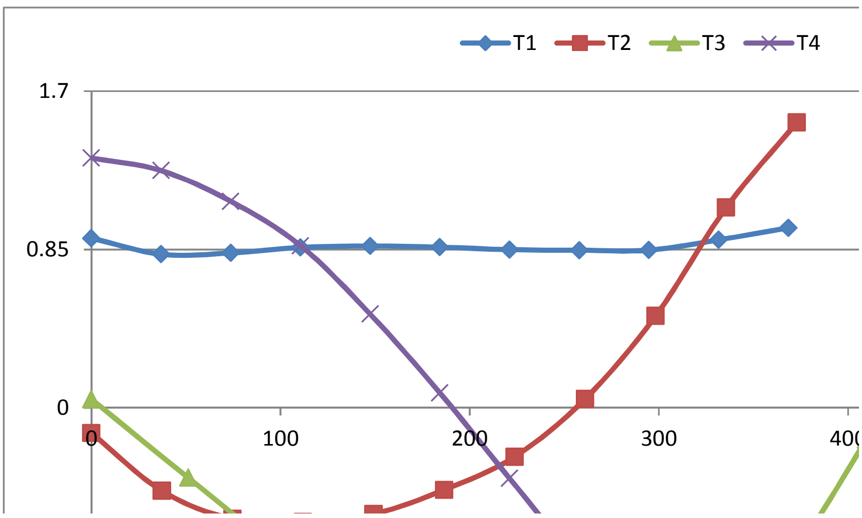


Figure 138. Tone-distribution of six Hungarian subjects

Based on the chart presented above, we can see that:

(1) The average F0 domain of all the Hungarian subjects covers the range (-1.7, 1.7), which is fairly close to the “model” native speakers’.

(2) The start of *Qusheng* is not the highest point in the average of the Hungarian subjects’ F0 domains, and the tone value of *Yinping* is considerably lower than that of the *Qusheng* start.

(3) The tone value of the *Yangping* end is considerably higher than that of *Yinping* and the start of *Qusheng*.

(4) The start of *Shangsheng* is quite close to that of *Yangping*, and the end of *Shangsheng* is higher than *Yinping*, the end of *Yangping* and the start of *Qusheng* and reaches the highest point in the average F0 domain.

All in all, concerning the individual’s F0 domain, compared with the “model” native speakers, F1’s, F4’s and M1’s are fairly close to the “model” native speakers’, M3’s and M4’s are much wider, especially M4’s, which is nearly twice as that of native speakers, while M5’s is somewhat narrower. In other words, 50% of Hungarian learners construct an appropriate F0 domain when they are producing the four tones. However, with regard to the distribution of the specific four tones, each subject has greater or lesser problems. In order to reveal these problems more precisely, the relatively exact tone value will be used to describe each subject’s four tones under their individual’s own tone system.

5.5.3 Tone value

To ensure the better description and comparison of the tone distribution, the chart of each subject’s F0 domain is divided into five points which correspond to the range from 1 to 5. 1 represnts the lowest point (the lower limit of a speaker’s normal voice range) and 5 the highest (the upper limit of a speaker’s normal voice range); 3 is mid pitch, 2 half-low and 4 half-high. From 1 to 5, means from the lowest to the highest. A tone can be described by indicating its beginning and ending point (If it is a falling-rising tone, the point which connects the falling section and the rising section is also indicated. The four tones of each subject are described in the following table. Note that a number accompanied by “+” indicates that the value is relatively higher than the number itself, and a number followed by "-" refers to a value that is relatively lower than the number itself. For example, the tone value of T1 produced by FB is 4^+4^+, which means FB's tone value of T1 is somewhat higher than 44. What is more, the middle part of T3 produced by FB could not be easily measured due to the existence of the creaky voice phenomenon, therefore the value of the inflection point remains undefined, and we use "~" to mark it. And the "FM" in the following table indicates all the female and male Hungarian subjects which are used in this section.

Table 99: *The tone value of each subject and all the Hungarian subjects*

	FB	MB	F1	F4	M1	M3	M4	M5	FM
T1	4^+4^+	4^-4^-	4^+4^+	4^+4^+	3^+3^+	2^+2^+	4^+4^+	5^-5	44
T2	35^-	2^+5	2^+4^-	2^+5	2^-4^-	2^-4^+	44	2^+4^+	25
T3	$3\sim3^+$	212^+	3^-14^-	41^+4^+	2^-13^+	2^-1^+5	4^-15	2^+1^+4	21^+5
T4	51	51	51^-	4^+1	51	3^+1	3^+1	51	4^+1

Based on the above table, the following conclusions can be made:

(1) In FB's and MB's individual tone systems, the tone value of the four tones can be described as follows: for T1, it is 4^+4^+ and 4^-4^-, respectively, which indicates that both the "model" native speakers' tone value of T1 are lower than 55; for T2, it is 35^- and 2^+5, in other words, the end of T2, which is considered as the linguistic goal, is at or nearly at the highest point in their individual tone system; for T3, it is $3\sim3^+$ and 212^+, the value of the start and the end of T3 which are the linguistic goals are 3, 3^+ and 2, 2^+, respectively, which indicates that the range between the start and the end of T3 is relatively narrow; for T4, it is 51 for both "model" native speakers, in other words, the tone value of T4 produced by both FB and MB covers the whole individual's F0 domain; that is, from the highest point 5 to the lowest point 1.

(2) In F1's individual tone system, the tone value of the four tones can be described as 4^+4^+, 2^+4^-, 3^-14^-, and 51^-, respectively. Compared with the "model" native speakers' individual tone system, F1's is quite appropriate, except for the tone value of the *Yangping* end which is considerably lower than the highest point.

(3) In F4's individual tone system, the tone value of the four tones can be described as 4^+4^+, 2^+5, 41^+4^+, and 4^+1, respectively. Compared with the "model" native speakers' individual tone systems, F4's tone value of the *Shangsheng* start is 4, which is inappropriately high. The tone value of the *Qusheng* start produced by F4 is 4^+, which is relatively lower than the native speakers'.

(4) In M1's individual tone system, the tone value of the four tones can be described as 3^+3^+, 2^-4^-, 2^-13^+, and 51, respectively. Compared with the "model" native speakers' individual tone systems, M1's tone value of *Yinping* is 3^+, which is inappropriately low, and the tone value of *Yangping* end is 4^-, which is considerably lower than the highest point.

(5) In M3's individual tone system, the tone value of the four tones can be described as 2^+2^+, 2^-4^+, 2^-1^+5, and 3^+1, respectively. Compared with the "model" native speakers' individual tone systems, M1's tone value of *Yinping* is 2^+, which is extremely low and inappropriate. The tone value of the *Shangsheng* start and end are 2^- and 5, in other words, the range between these two points is considerably wider than that of the native speakers'.

(6) In M4's individual tone system, the tone value of the four tones can be described as 4^+4^+, 44, 4^-15, and 3^+1, respectively. Compared with the "model" native speakers' individual tone systems, M4's tone value of *Yangping* is 44, which is inappropriately flat. The tone value of the *Shangsheng* start and end are 4^- and 5, which are considerably higher than those of native speakers'. Therefore, the range 1 (the distance between the start and the inflection point) and range 2 (the distance between the inflection point and the end) are inappropriately wide. What is more, the tone value of the *Qusheng* start is 3^+, which is inappropriately low.

(7) In M5's individual tone system, the tone value of the four tones can be described as 5^-5, 2^+4^+, 2^+1^+4, and 51, respectively. Compared with the "model" native speakers' individual tone system, M5's is quite acceptable, except for the tone value of the *Yinping* start which is relatively lower than its end.

(8) In FM's tone system, the tone value of the four tones can be described as 44, 25, $21^{+}5$, and $4^{+}1$, respectively. Compared with the "model" native speakers' individual tone systems, FM's tone value of the *Shangsheng* end is considerably higher, while the tone value of the *Qusheng* start is relatively lower.

All in all, under the individual's own tone system, all the Hungarian subjects have greater or lesser problems in terms of the tone value. The main problems can be summarised as follows: M1's and M3's tone values of *Yinping* are inappropriately low, especially M3's, which is lower than the mid pitch; the tone values of the Yangping end produced by F1 and M1 are considerably lower than the native speakers'; the tone values of the *Shangsheng* start produced by F4 and M4 are inappropriately high, while the tone values of the *Shangsheng* end produced by M3, M4 and FM are rather higher than the "model" native speakers'; the tone values of the *Qusheng* start produced by F4, M1, M4 and FM are considerably lower than FB's and MB's, especially M1's and M4's, which are inappropriately low.

5.5.4. Summary

The comparison of tone duration, tone distribution and tone value within the individual's own tone system clearly indicates the existence of discrepancies between Hungarian subjects and the "model" native speakers. On the basis of the tables and figures presented above, the following conclusions can be drawn.

(1) In terms of the tone duration, the order of the average durations of four tones produced by native speakers is T3>T2>T1>T4. Only 2 of the 10 Hungarian subjects produced the same duration order of the four tones, in other words, 80% of the Hungarian learners produce an inappropriate duration order of the four tones, most often in producing *Yinping* longer than *Yangping*. This in itself represents a notable difference between the Hungarian learners and the native-speaker "models", sufficient to warrant pedagogical attention.

(2) With regard to the tone distribution, only 50% of Hungarian learners construct an appropriate F0 domain when they are producing the four tones. What is more, when investigated with tone value, 100% of the Hungarian learners differ considerably from

the “model” native speakers, and the problems cover all the four tones.

5.6. Problems revealed by the data

In this section, we will summarize all the problems identified and described in detail in this chapter in terms of duration, fundamental frequency and range.

5.6.1. Duration

Concerning the absolute duration of each specific tone, it emerges from the analysis of the data that the Hungarian university learners in this study differ considerably from the “model” native speakers. In terms of the duration of *Yinping*, *Yangping*, *Shangsheng* and *Qusheng*, the percentages of the Hungarian learners who have problems are 60%, 60%, 70% and about 57%, respectively, most often in producing much shorter durations. However, this finding is somewhat mitigated when the concept of the intrinsic vowel duration is applied to the analysis: under these conditions the Hungarian university learners produce deviant orders of durations of /i/, /u/ and /a/ in *Yinping*, *Yangping*, *Shangsheng* and *Qusheng* with the percentages of 50%, 20%, 30% and 30%.

Leaving out the factor of the intrinsic vowel duration, the duration problems that existed in the four tones can be sorted as follows: *Shangsheng*, *Yinping*, *Yangping* and *Qusheng* (*Yinping* and *Yangping* are in the same level), with the tone which is most likely to cause problems first and the one which is least likely to cause problems last in the list.

Considering the duration under the individual’s own tone system, the “desirable” order of the average durations of four tones produced by native speakers is T3>T2>T1>T4. Only 2 of 10 Hungarian subjects produce the same duration order of the four tones, in other words, 80% of Hungarian learners produce the inappropriate duration order of the four tones, most often as a result of producing *Yinping* longer than *Yangping*. With regard to the tone distribution, 50% of Hungarian learners could not construct an appropriate F0 domain when they producing the four tones. However, when investigated with tone value, 100% of these Hungarian university learners differ considerably from the native speakers and their problems affect all the four tones.

Compared with the absolute duration of each tone, the duration problems appear even more serious when they are investigated under the individual's own tone system, which suggests that approaches to solving duration problems which are based on the individual's own tone system probably are likely to be more effective.

5.6.2. Fundamental frequency

With regard to fundamental frequency (F0), 20% of Hungarian university learners produce deviant contours of *Yinping*, most often as a result of a loose start or ending, while 75% of Hungarian female learners are likely to produce *Yinping* tone accompanied by the quite obvious existence of the IF0. 50% of the Hungarian subjects have difficulty producing appropriately steep contours of *Yangping*. The deviant contours they produce mainly appear as a contour with too steep or too gentle a rising pattern, a fall-rise pattern or a flat pattern, or an inappropriate rising point. And the effect of IF0 is fairly clearly present in 40% of Hungarian learners when they are producing *Yangping* tone, mainly appearing in the second half part of their average F0 contours. About 67% of the Hungarian subjects have difficulty producing appropriately fall-rise contours of *Shangsheng*. The deviant contours they produce mainly appear as a contour with an extremely sharp fall-rise pattern, an excessively steep slope, a rising pattern without a clear falling start, or an inappropriately located inflection point. And the effect of IF0 is quite obvious in Hungarian learners when they are producing *Shangsheng* tone, especially among female learners, with 100% in the sample showing IF0 effects. About 57% of Hungarian university learners have difficulty producing appropriately falling contours of *Qusheng*. The deviant contours they produce mainly appear as a contour with an inappropriately flat end, a late inflection point, a very gentle falling part, or even a declining "S" pattern. And the effect of the IF0 is relatively obvious in Hungarian learners when they are producing the *Qusheng* tone, especially in female learners, with 100% in this sample.

On the basis of the above conclusions, the percentages of Hungarian university learners with F0 problems in *Yinping*, *Yangping*, *Shangsheng* and *Qusheng* are 20%, 50%, 67% and about

57%, respectively. We can put them in the order of *Qusheng*, *Yangping* and, with the tone which is most likely to cause problems in F0 first and *Yinping* which is least likely to cause problems in F0 last in the list. The result suggests the conclusion that more effort should be put into “improving” *Shangsheng* when we are searching for the possible solutions for F0 problems.

5.6.3. Range

As to range, for *Yinping*, 20% of Hungarian university learners can be considered deviant in the sense that they produce much wider ranges. For *Yangping*, 60% of Hungarian university learners produce deviant results with much wider or narrower ranges. For *Shangsheng* and *Qusheng*, the percentages are 50% and about 40%, respectively, with much wider or narrower ranges.

Based on the percentages of Hungarian university learners who have range problems in *Yinping*, *Yangping*, *Shangsheng* and *Qusheng*, we can put them in the order of *Yangping*, *Shangsheng*, *Qusheng* and *Qusheng*. This indicates that *Yangping* is most likely to cause range problems, while *Yinping* is least likely to cause such problems.

To sum up, the acquisition of *Shangsheng* and *Yangping* can be considered more difficult for Hungarian university learners, while the acquisition of *Qusheng* and *Yinping* is relatively easy for them. Comprehensively evaluating all the data in terms of duration, fundamental frequency and range, we can say the acquisition of *Shangsheng* is most difficult for Hungarian learners, followed by *Yangping*. *Qusheng* is the third, and *Yinping* is least difficult. This suggests that teaching tones in the order of *Yinping*, *Qusheng*, *Yangping* and *Shangsheng* constitutes can form a relatively ideal sequence, starting from the easiest tone and concluding with the most difficult.

5.6.4. Tone distribution and tone value

Concerning the individual’s F0 domain, 50% of Hungarian learners construct an appropriate F0 domain for producing *Yinping*, *Yangping*, *Shangsheng* and *Qusheng*. However, all Hungarian subjects have greater or lesser problems in the four tones with regard to the tone value. Specifically, the percentages of the Hungarian learners who have the tone value problems in *Yinping*, *Yangping*, *Shangsheng*, and *Qusheng* are the same, that is 50%.

5.6.5. Summary

Based on the problems which were revealed by all the data in this chapter, we classified these problems and put them into the following table for the purpose of better comparison. An "X" in any cell indicates that the subject has problems in that particular aspect, and if the subject has no problems in a certain aspect, then the relevant cell remains blank. We have also used "CV" as the abbreviation of "creaky voice". Note that the intrinsic vowel duration and the intrinsic vowel fundamental frequency do not appear in this table; they will be discussed separately later in this section.

Table 100: *Problems revealed by the data for each Hungarian subject.*

		F1	F2	F3	F4	M1	M2	M3	M4	M5	M6
T1	Duration					X	X	X	X	X	X
	F0	X								X	
	Range		X							X	
T2	Duration			X	X	X		X	X		X
	F0			X			X		X	X	X
	Range	X		X	X		X		X		X
T3	Duration	X				X	X	X	X	X	X
	F0		CV	CV	X		CV	X	X	X	CV
	Range		CV	CV	X		CV	X	X	X	CV
T4	Duration			X			X	X			X
	F0	X	CV				CV	X	X	X	CV
	Range		CV	X	X		CV		X		CV
Individual's tone system	Duration	X	X	X	X			X	X	X	X
	Range		CV	CV			CV	X	X	X	CV
	Value	X	CV	CV	X	X	CV	X	X	X	CV

The above table clearly shows that the creaky voice phenomenon appears with considerable frequency in the production of these Hungarian university learners, mainly in *Shangsheng* and *Qusheng*, with the percentages of 50% of the female subjects and about 30% of the male subjects being affected. This indicates that the creaky voice phenomenon is more likely to appear in female subjects' vowels than in the males'. The data also show that if the creaky voice occurs during a subject's *Shangsheng*, it will occur during his/her *Qusheng* as well.

It is worth noting that although M1, M2 and M5 are all third-year students, and they have all studied Chinese in China for about one year, their longer and presumably deeper experience and study of the Chinese language does not appear to have conferred any advantage, compared with the other learners in the sample, with respect to their production of the four tones. This indicates that it would be mistaken to assume that the improvement of the pronunciation of tones will take place automatically or that correct pronunciation will be naturally and unconsciously acquired. On the contrary, it seems that deliberate, conscious and focused intervention is required, without which the learners' pronunciation of tones will not be better, no matter how long they spend in learning Chinese and where they learn.

The percentages of the Hungarian learners who experience duration problems in *Yinping*, *Yangping*, *Shangsheng* and *Qusheng* are 60%, 60%, 70% and about 57%, respectively; the percentages in which F0 problems occur in these four tones are 20%, 50%, 67% and about 57%; and the percentages in range problem in these four tones are 20%, 60%, 50% and about 40%. In other words, in terms of the four specific tones, the duration problem is likely to be the most serious, followed by the F0 problem, while the range problem is relatively less serious. However, when these four tones were investigated under the individual's own tone system, it turned out that the most serious problems occur in the tone value, affecting 100% of the learners; the duration problem comes second and is followed by the range problem.

With regard to the intrinsic vowel duration (IVD) and the intrinsic vowel fundamental frequency (IF0), the IVD exists in *Yinping*, *Yangping*, *Shangsheng* and *Qusheng* produced by the native speakers and the result corresponds to the "i>a" principle which is encapsulated in the formula IVD "i>a". Therefore, we could use IVD "i>a" as a "desirable" standard to evaluate the existence of IVD in *Yinping*, *Yangping*, *Shangsheng* and *Qusheng* produced by the Hungarian subjects, and it turned out that the Hungarian university learners produce deviant orders of durations of /i/, /u/ and /a/ in *Yinping*, *Yangping*, *Shangsheng* and *Qusheng* with the percentages of 50%, 20%, 30% and 30% respectively. This indicates that the inappropriate pronunciation of vowels probably causes problems in duration.

However, concerning the IF0, on the basis of the data obtained so far it is difficult to state definitively that to what extent this factor affects the native speakers when they are producing the four tones, since the manifestations of IF0 only occurred relatively clearly in the *Yangping* syllables produced by the "model" native speakers, but not in the *Yinping*, *Shangsheng* and *Qusheng* syllables produced by them. This phenomenon appears to contradict Zhu Xiaonong, who believes that IF0 in high areas of the individual's own frequency domain is more distinct than that in low area[213]. According to his statement, the performances of IF0 should have occurred in their pronunciations of *Yinping*, *Shangsheng* and *Qusheng* and should have been accompanied by evidence showing the effect of IF0 along the whole F0 contour of *Yinping*, the beginning and the ending parts of *Shangsheng* F0 contour, and the beginning part of the *Qusheng* F0 contour. Therefore, it is impossible to reach a final conclusion as to whether or not the existence of the IF0 in the F0 contours of *Yinping*, *Yangping*, *Shangsheng* and *Qusheng* produced by the Hungarian university learners caused problems.

[213] Zhu 2010.

6. Possible solutions

6.1. Current situation of tone teaching

The findings detailed and summarized in the previous chapter confirm the belief held by many experienced teachers of Chinese as a foreign language that the tone is one of the most difficult parts, and at the same time a key element, in the teaching of Chinese pronunciation. The reasons for this are not far to seek: on one hand, unlike the majority of languages spoken in western countries, where meaning at sentence-level is partly expressed through intonation but syllabic tone is absent, Chinese uses tones at syllabic level to carry essential lexical information as well as intonation at sentence level. On the other hand, as Zhao Jinming (2007) states, there are many tone languages in the world, but we find very few languages like Chinese language in which every syllable has a fixed tone; furthermore, the fixed tone range covers the area from high to low, rising and falling[214]. Lin Tao (1996) considers that the key of "speaking Chinese with a foreign accent" is mastery of correct tone, which implies a higher phonetic level than the ability merely to produce the initials and the finals properly[215]. And Lu Jianji (1984) believes that "the tone is particularly difficult for students whose native language is not a tone language"[216]. Therefore, tone is probably the greatest single obstacle facing the majority of foreign students in learning Chinese. That is why this thesis focuses on research into learners' production of tone rather than the research into the initials and the finals in the experimental phonetics part. And it further focuses on the study of monosyllabic tone within tone research, because the monosyllabic tone is the foundation of the Chinese tone. It is possible that learners cannot be sure of mastering the standard tone well even if they have already mastered monosyllabic tones accurately, but there is no doubt that proper mastery of monosyllabic tones is an essential prerequisite if learners are to have any chance at all of moving on to learn the standard tone well. In a word, there is a necessarily conditional relationship between the monosyllabic tone and standard tone.

[214] Zhao 2007: 369.
[215] Lin 1996.
[216] Lu 1984.

Tone teaching, then, undoubtedly plays an important part in the teaching Chinese as a foreign language. However it has been and it remains a weak link in the process. Zhao Jinming, editor in chief of the highly influential *Introduction to the Teaching of Chinese as a Foreign Language*, points out that the problems undermining successful tone teaching reflect systemic weaknesses that can be identified at every level of the process: in the teaching plan, the teaching material and teaching methodology. The weakness inherent in the overall teaching plan is that tone teaching lacks hierarchy and order: as we saw in Chapter 1.2 above, syllable tone is taught mainly at the very beginning of almost any course, after which it is largely ignored. This weakness stems partly from teachers' attitudes, but partly also from the arrangement of the available teaching materials, which are far from satisfactory in this respect. In fact the situation has changed very little since Guo Jinfu (1993) asked "why is the teaching of tone and intonation a weak link in the teaching of teaching Chinese as a foreign language?" and answered his own question "An important reason is that we lack a practicable, Chinese tone and intonation centered language learning material all the time. However, this kind of teaching material is extremely needful, and it is also the basic teaching material which highlights the characteristics of Chinese language"[217]. The weakness in teaching methods which is pointed out by Zhao Jinming (2007) is that, for the characteristics of the tone, currently only Chao's five-degree annotation is widely used or adapted to show the shapes of the four tones. Therefore, many teachers are reduced to using gestures as the only means help students in understanding the proper contours and characteristics of the tones during teaching[218].

Fortunately, the importance of tone teaching has been increasingly acknowledged, and teaching field of teaching Chinese as a Foreign Language has developed some relatively efficient teaching methods. Apart from the assistance of five-degree annotation system by Chao Yuan Ren and gestures, the imitation method is also a frequently-used method. It is almost certainly the most commonly-used method in tone teaching and phonetic teaching. Learners will imitate the model of correct pronunciation provided by their teachers or by the

[217] Guo 1993.

[218] Zhao 2007: 369.

available audio material, and the teachers will continue to demonstrate and correct constantly until they feel that they have reached a satisfactory result. Furthermore, there is also a listening & distinguishing method.

The very nature of pronunciation determines the important position of the pronunciation teaching in the teaching of Chinese as a foreign language. Yuen Ren Chao uses the *Mandarin Primer* as teaching material when teaching Chinese abroad[219]. "Most teaching is about pronunciation; then moves on to the teaching of other aspects". During his period, pronunciation teaching was highlighted. Along with the boom of the teaching Chinese as a foreign language, internally, this discipline has been unevenly developed. Considerable developments have been made in many aspects. However, some feel that pronunciation teaching has not been crowned with notable success. As Lin Dao said, "pronunciation teaching has not made any progress. On the contrary, it has greatly regressed.[220]" Therefore, research on pronunciation and pronunciation teaching is less than that in other fields. Even so, during the past several decades, there have been still many excellent theses and papers on pronunciation and pronunciation teaching. These theses and papers are of great significance for the theoretical research and teaching practices in pronunciation, and we will now proceed to consider some of the more notable examples.

6.1.1. General overviews of pronunciation teaching practice

All too often, when considering research into the teaching of pronunciation teaching, one comes upon purely theoretical research which is not informed or validated by no actual data. However, in recent years, research validated by concrete data has been on the increase and has begun to take up an increasingly large proportion in the entire body of literature on pronunciation research.

In his *Chuji Hanyu Shuiping Liuxuesheng de Disansheng Tingbian Fenxi*, Ma Yanhua designed a survey which is quite different from any such surveys that had been applied

[219] Chao 1948.

[220] Lin 1996.

previously in this field[221]. It included four aspects: contrastive analyses of listening to the third-tone monosyllables and the third tone contained in disyllables; contrastive analyses of the position of the third in a disyllable; whether familiarity or unfamiliarity with the content of listening tasks affects listening; whether the presence or absence of tone in the mother tongue has an effect on listening. Based on the results gained from the survey, the paper proposes a variety of teaching strategies designed to improve effective recognition of the third tone.

In his book *Yuyinxue Jiaocheng*, Lin Tao [222]discusses standards for the pronunciation teaching of Chinese as a foreign language with concrete examples. Lin Tao believes that the pronunciation teaching calls for strict pronunciation standards. However, since no completely detailed standard pronunciation for Chinese Mandarin that would cover the entire field has yet been established, he advocates that in the teaching of pronunciation, strict pronunciation standards should be imposed at the beginning of the process, as long as the basic, generally agreed areas are being covered, while gradually less rigorous attitudes should be adopted to deal with the ambiguous pronunciation phenomena in less well-known areas where full consensus has not been reached, which points are of great practical significance for the teaching practice.

In view of reasonableness of the *Hanyu Pinyin Fang'an* [223], in *Hanyu Pinyin Fang'an yu Shijie Hanyu Yuyin Jiaoxue*, Wang Lijia[224] points out that the internationally recognized *Scheme for the Chinese Phonetic Alphabet* is linguistically perfect. In the meanwhile, he analyzes problems such as the design philosophy, formulating principles of the *Scheme for the Chinese Phonetic Alphabet* and relations between the alphabet and its pronunciation and discusses the limitations of the *Scheme for the Chinese Phonetic Alphabet* from this point of view. Moreover, he also suggests that in pronunciation teaching, TCSL teachers should be flexible enough to choose different teaching means. Up to the present time, this paper is the

---

[221] Ma 2000.

[222] Lin Tao 1992.

[223] Hanyu Pinyin Fang'an 1958.

[224] Wang 2005.

one that makes the most detailed and comprehensive analyses to guide the use of the *Hanyu Pinyin Fang'an* in teaching Chinese as a foreign language. In *Hanyu Yuyin Jiaoxue de Xianzhuang yu Duice*, Zhang Baolin indicates that over the past ten years, the results of phonetics teaching have not come up to the expectations in TCSL[225]. The causes that he adduces for this disappointing state of affairs are related to its methodology, syllabus and correct understanding of the significance. He suggests that great importance should be attached to phonetics teaching and the teaching means: methods should be modernized. In the meanwhile, he recommends strengthening the basic study on the phonetics teaching. The paper also puts forward what the author believes to be an effective way for phonetic teaching: he suggests starting by making use of the International Phonetic Alphabet, and then using it as a basis for teaching learners to understand and use Chinese Pinyin.

In their *Liuxuesheng Rumen Jieduan Yuyin Jiaoxue Yanjiu* [226], Gu Zheng and Wu Zhong-wei state the target and principle of phonetics teaching at the initial stage, then analyze the three views of the phonetics teaching which they report as being popular at present in TCFL. They are phoneme teaching, tongue stream teaching and the combination of the previous two. Based on their analyses, the authors put forward a new approach to phonetics teaching, which they call syllable teaching. The authors claim that syllable teaching suits the features of Chinese phonetics and therefore can be expected to promote phonetics learning, in the meanwhile, the paper explores the effective practical steps involved in syllable teaching.

6.1.2. Research on errors connected to prosodic features of Chinese, and their treatment

Research on prosodic features of Chinese has drawn increasing concern in the field of TCSL, since it is increasingly clear that it is quite difficult to teach effectively simply focusing on consonants and vowels as phonemes of Chinese and to promote fundamentally correct accents in the speech production of foreign students. Research results in this respect are as follows.

In her *Chaoyinduan Chengfen de Biaoyi Gongneng Jiqi zai Duiwai Hanyu Jiaoxue zhong de Yingyong*, Dai Xiaoxue explores the constitutive characteristics of suprasegmentals and types

[225] Zhang 2005.

[226] Gu and Wu 2005.

of auxiliary ideographic functions and discusses applications of ideographic functions of supra-segmental phonemes in the teaching of Chinese as a foreign language[227]. She suggests that in the teaching process, TCSL teachers should consciously guide students in order to enable them to better understand the difference between the literal meaning and the expressional meaning so as to achieve the desired teaching aims. In his *Hanyu Ciyu Goucheng de Yinjie Yinsu he Jiezou Yugan de Peiyang*, Wu Huainan indicates that syllables are related to word formation and word collocation, and argues that many mistakes the learners make are due to the fact that they are unaware of the restrictions of word formation and word collocation, which includes their ignorance of syllabic factors in word collocation[228]. He advocates that prioritizing the acquisition of a good command of rhythmic principles is an effective way to improve the efficiency of teaching Chinese as a second language.

Taking a somewhat different focus, in his *Primary Investigation of the Structural Pattern of Chinese Rhythm*, Wu Weishan discusses the interaction between Chinese tone and rhythm as two sets of rhythmic systems and the structural pattern of Chinese rhythm, including the following two aspects: 1) building a basic framework of Chinese rhythmic structure; 2) uncovering typical characteristics of Chinese rhythmic structure, both of which must be seen as being of guiding significance for the teaching of pronunciation[229].

6.1.3. Research into the Teaching of Phonetic Elements

Research efforts in the field of the teaching of phonetic elements have been mainly devoted to the teaching of the tone system and of the syntax and rhythm systems. One particular issue concerned with tone teaching that has received a great deal of attention is the misunderstanding of the rising tone. In comparison to this, research on the teaching of consonants and vowels is relatively scarce.

In his *Tantan Xiandai Hanyu Shengdiao Jiaoxue*, Du Qinhuai discusses problems existing in students' learning the tone in the Chinese pronunciation learning process[230]. He focuses on

---

[227] Dai 2004.

[228] Wu 2003.

[229] Wu 2005.

[230] Du 1992.

analyzing the rising tone and the neutral tone and points out that the misunderstanding of the true nature and the proper pronunciation of the rising tone exercises very great influence on the accuracy of learners' pronunciation. Taking a somewhat controversial stance, Liu Chuanping's *Putonghua Shangsheng Benzhi Jiqi Jiaoxue de Zai Renshi* sets the tone pitch of the falling-rising tone as 211, owing to the low and level nature of the falling-rising tone and the high usage rate of the low and level tone[231]. 214, with a low usage rate, is identified as an alternative version of the falling-rising tone. In the meanwhile, he also puts forward methods for teaching the falling-rising tone. This paper is of great enlightening significance for the tone teaching.

In his *Duiwai Hanyu Shengdiao Jiaoxue Celüe Tansuo*, Song Yidan summarises the drawbacks that, in his opinion, bedevil the present teaching of Chinese tones, including the lack of explanations on the property of tones and the manner of articulation, the difficulties in level-division teaching, the relatively boring of teaching methods and so on[232]. On the basis of his investigations, Song Yidan puts forward various teaching strategies; interestingly, he proposes the use of PRAAT software for the purposes of feedback.

6.1.4. Theories and Principles related to the teaching of the pronunciation of Chinese as a foreign language

6.1.4.1. General Pronunciation Teaching Theories

In his *Yuyin Yanjiu he Duiwai Hanyu Jiaoxue*, Lin Tao points out that the tone and the pronunciation levels which are higher than the tone, rather than consonants and vowels are the main reason for the forming of foreign accents[233]. He believes that the pronunciation learning is not just about the imitating of the pronunciation of the word, but more about the imitation of the unstressed and the stressed characteristics of the tone, the high and the low points of the pitch and the intonation mode. Moreover, such kinds of imitations should be encouraged throughout the whole learning process. This paper makes a comparison between patterns of

[231] Liu 1996.
[232] Song 2009.
[233] Lin 1996.

the unstressed and the stressed tones of Chinese and English, which are of guiding significance for research on the teaching Chinese pronunciation as a foreign language and foreign accents existing in different countries.

In her *Hanyu Yuyin Yanjiu yu Hanyu Yuyin Jiaoxue Jiekou zhong de Ruogan Wenti,* Wang Yunjia integrates research with TCFL in practice, discussing problems existing in the Scheme for the Chinese Phonetic Alphabet, problems about the suffixation of a non-syllabic “r” to nouns, neutral tone, tones and intonation[234]. She believes that the most important problems that need to be better solved include those of how to absorb and treat achievements gained from research on pronunciation for the pronunciation teaching and how to better integrate theoretical research of pronunciation with their practical applications. Moving into a more direct application of principles, Mao Shizhen in his *Guanyu Duiwai Hanyu Yuyin Jiaoxue de Fansi* expresses his conviction that in the pronunciation teaching, “teaching students in accordance of their aptitude” as a principle should also be followed[235]. In accordance with the order “expressive- standardized-expressional-stylized”, students are encouraged and enabled to make constant progress. In the meanwhile, since the pronunciation input and output have strong correlation, Mao Shizhen recommends that the pronunciation training should not lay too much emphasis on the correctness of output, but provide massive pronunciation input opportunities for students via appropriate listening activities. Besides, this paper discusses the issue of whether it is more advisable to adopt phoneme teaching or speech flow teaching and concludes that on the whole, compared with phoneme teaching, the speech flow teaching is more suitable for the Chinese pronunciation teaching in view of the Chinese pronunciation teaching system and the rules that appear to govern the acquisition of Chinese pronunciation. However, he also puts forward two practical problems existing in the actual operation. On the one hand, phoneme teaching should not be completely abandoned in favour of conducting speech flow teaching. On the other hand, problems such as how to conduct the speech flow teaching, what kinds of things should be taught, how to allocate teaching time and how to arrange the teaching progress need further discussions.

---

[234] Wang 2002.

[235] Mao 2002.

In his *Duiwai Hanyu Yuyin Yanjiu yu Yuyin Jiaoxue Yanjiu Zongshu*, Zhou Fang reviews the achievements in the research and teaching of Chinese speech sounds in TCFL since the 1980s, then discusses the tendency to move from the unanalytical description of experience to practical research[236]. In the meanwhile, he draws particular attention to three existing problems: a limited body of contrastive research on speech sounds; a weak combination of the research with its application; a great disparity while compared with the main achievements in the rest of the world in this field due to the limitations of the research conditions and levels. Therefore, his main conclusion consists of a heartfelt call for professionals and academics in China to find out about and absorb the relevant research findings from abroad. Meanwhile, in a somewhat controversial paper, *Duiwai Hanyu Yuyin Jiaoxue Jige Jiben Wenti de Zai Renshi*, Lu Jianji discusses some basic issues concerning the teaching of Chinese phonetics to foreign learners, such as the position, the primary goal, the role of phonetic theory, the importance of stress and intonation rules and the application of the Scheme for the Chinese Phonetic Alphabet[237]. Then, he puts forward some suggestions on the principles that he believes should guide the teaching of phonetics. This paper gave rise to considerable debate in the circle of professionals involved in teaching Chinese as a foreign language, as the views that Lu Jianji expresses are different or partly different from the prevailing ones.

Research achievements in the teaching of specific phonetic elements are mainly in the field of tone teaching and syntax and rhythm teaching. The misunderstanding in the rising tone teaching is a particularly popular topic in tone teaching. In comparison, research into consonant and vowel teaching is relatively scarce.

In the *Hanyu Putonghua de Teshu Yunmu Jiqi Jiaoxue,* Zhou Cuilin discusses statuses of ê, –i（before the tongue tip）, -i（after the tongue tip）and er as four vowels in the *Scheme for the Chinese Phonetic Alphabet*, from three aspects: the Pronunciation theories, the textbook compilation and the classroom teaching[238]. In the meanwhile, he also puts forward methods to deal with these four vowels both in textbooks and in classroom teaching, which are of some

[236] Zhou 2006.
[237] Lu 2010.
[238] Zhou 1999.

significance for improving the pronunciation teaching and reducing mistakes made during the learning process. In his *Sentence Stress and Oral Language Teaching,* Ji Xiuqing discusses the role the sentence stress plays in oral expression and relevant factors affecting the proper utilization of the stress. Moreover, this paper also concludes that the teaching of the sentence stress can improve students' verbal expression skills, increase the grammatical knowledge of students, deepen students' understandings toward the methods for expressing themselves effectively in Chinese, cultivate students' language senses and increases students' listening skills.

In *Yuliu Jiaoxue Chutan*, Jiang Yiliang [239]analyzes errors that occur frequently in learning the speech flow, and explores possible ways to teach speech flow in the pronunciation teaching. According to him, so as to improve the pronunciation teaching, the tone teaching should be improved so as to ensure that students have a sense of Chinese tones. Moreover, Chinese language sense should also be fostered so as to cultivate the pronunciation habits of students. Besides, before students' start to form their own Chinese speech flow, sentences and paragraphs with distinctive Chinese prosodic features should be regularly practiced by students so as to enable them to gain the basic skills of language flow.

In their *Duiwai Hanyu Jiaoxue zhong Jianyuju de Yuyin Qiefen Yanjiu*, through the investigation of the phonetic segmentation of pivotal sentences, Wang Lixiang and Jiang Hailing find that in addition to the complexity of the predicate, the phonetic segmentations of most pivotal sentences are related to syntax rules, semantic expressions and rhythms and so on[240].

In her *Shengdiao Pianwu yu Duiwai Hanyu Shengdiao Jiaoxue Yanjiu Zongshu,* Mao Li [241]explores the notion that tones are the biggest obstacle in the mastering of Chinese phonetics to most of the Chinese learners. To those whose mother tongues are also tone languages, their tone errors vary according to the influences of the negative transfer from their mother tongues. While to those whose mother tones are non-tone languages, to some extent, their tone errors

---

[239] Jiang 1998.

[240] Wang and Jiang 2004.

[241] Mao 2007.

are quite similar. As to the errors and their causes, the paper puts forward six effective strategies on how to correct the tone errors in TCFL.

In their *Duiwai Hanyu Yuyin Jiaoxue zhong Shengmu Jiaoxue Yanjiu,* Shen Li'na and Yu Yanhua state that in teaching Chinese for overseas students, phonetic teaching runs through the primarylstage of Chinese teaching. And because of the negative transfer of native language, pronunciation errors often occur[242]. The paper discusses the errors that foreign students easily make in the initial teaching and puts forward possible solutions.

6.1.4.2. Country-specific Pronunciation Teaching

Since there are many differences between the phonological features of Chinese and language, can be argued that ideal teaching effects can be gained only by targeted teaching schemes which are based on the comparison between Chinese and particular foreign languages. Therefore, research on country-specific pronunciation teaching appears to be essential. Among current research achievements, theses analyzing the pronunciation of Japanese and Korean students in China account for the highest proportion of those submitted in Chinese institutions, and theses on similar topics by students from Europe and America and Southeast Asia in China account for the second highest proportion. Here we will survey only a small selection of work by Chinese scholars in this field.

It is hardly surprisingl that many researchers and authors have devoted themselves to issues related to the problems and possibilities of teaching Chinese pronunciation to speakers of particular languages. There follows what can only be a limited selection of some of the more notable examples. Meng Zhuyi's *On Shilun Yuyin Jiaoxue zhong de Wudao Wenti*, based on research into the actual situation of the teaching of Chinese as a foreign language in Korea, explores common and influential misleading phenomena or problems that need to be solved so as to improve the efficiency of pronunciation teaching[243]. Moreover, he also lists some misunderstandings in the teaching of falling-rising tones, the neutral tones, unstressed sounds, pronunciation and articulation and the order of the consonant and vowel teaching, and puts

[242] Shen and Yu 2009.

[243] Meng 2000.

forward aspects that need to be taken into account in order to avoid misleading students.

In his *Zhongji Hanyu Shuiping Hanguoren Shengmu Pianwu Diaocha*, Zhang Zhiyun divides his investigation into two parts, in the hope of discussing errors Koreans made in learning Chinese initials[244]. In one part of the investigation, recording is chosen as a method to investigate the Chinese pronunciation of Koreans. The author of this paper makes an analysis of the investigation results. The other part of investigation is conducted by distributing questionnaires, which investigate the background information of respondents, difficulties in learning Chinese and so on. The investigation results are of guiding significance for teaching the Chinese initials. Still with reference to leaners from Korea, Song Chunyang in his *Tan Dui Hanguo Xuesheng de Yuyin Jiaoxue---Nanyin ji Duice* analyzes difficulties existing in Korean students' learning of Chinese pronunciation, explores the causes of these difficulties and puts forward some relevant teaching solutions[245]. In similar vein Huang Zhenji[246]'s *Qiantan Dui Hanguo Liuxuesheng de Chuji Hanyu Yuyin Jiaoxue* discusses common problems existing in Korean students' learning Chinese pronunciation in terms of its tone, rhyme and intonation and puts forward some possible solutions.

In *Yinni Liuxuesheng Xide Hanyu Seyin he Secayin Shiyan Yanjiu*, Lin Yigao and Wang Gongping investigate the intermediate and advanced Indonesian learners' acquisition of stops and affricates in standard Chinese through experimental phonetics, including aspirate length, syllable duration, closure length and voice interval[247]. They establish that the aspirate length pronounced by the Indonesian learners is shorter than that of the Chinese, but the unaspirates and closure length and voice interval are longer. Besides, the initial-finial pattern is not so appropriate. The paper puts forward some teaching strategies on how to teach the Indonesian learners to pronounce the stops and affricates of standard Chinese: a fine example of the application of specific research findings to the teaching of Chinese to a particular group of learners.

---

[244] Zhang 2002.

[245] Song 1998.

[246] Huang 1999.

[247] Lin and Wang 2005.

In her *Research on the Acquisition of Chinese Vowels for Japanese Learners*, Wen Baoying investigates thirty Chinese learners whose native languages are Japanese. According to how long they have been learning, she divides them into three levels, beginner, intermediate and advanced, respectively. Then, she investigates the learning of seven first level vowels in Chinese by means of statistical analysis and pattern analysis on the hope of finding out the traits of the acquisition of vowels for Japanese learners. In the meanwhile, she tests the results with transfer theory and universal grammar theory.

In his *Ri, Han Yuyin Duibi Fenxi yu Hanyu Yuyin Jiaoxue* [248], Yu Wei adopts the contrastive approach to make a detailed analysis of the pronunciations of the Japanese and Chinese languages, and classifies Chinese pronunciation features according to their learning difficulty; he then explores issues related to teaching Japanese students Chinese pronunciation and various possible solutions. In addition to a comparison of consonants and vowels, this paper also puts forward a contrastive teaching method for Chinese monophthongs. Focussing on the same category of learners Dong Yuguo in his *Dui Riben Xuesheng Biyunmu Yin de Jiaoxue* indicates that Japenese learners have considerable learning problems with nasal finals of Chinese syllables[249]. In the meanwhile, this paper investigates and analyses the causes of such problems through auditory differentiation and articulation recordings. Then, the author proposes a set of teaching methods in for use pedagogical practice.

In his *Meiguo Yingyu Yudiao Dui Meiguo Xuesheng Xuexi Hanyu Putonghua Shengdiao de Ganrao*, Gui Mingchao mainly discusses problems existing in teaching Chinese tones to students in English-speaking countries[250]. On the basis of an in-depth study on common mistakes American students make during their learning of Chinese tones, he summarizes some tone types which cause many students to make mistakes when they are trying to learn them. The paper points out the somewhat obvious truth that in English the pitch is only used to refer to the intonation. However, the pitch in Chinese is used both to refer to the intonation and to make a distinction between different tones. Also with reference to native speakers of

[248] Yu 1995.

[249] Dong 1997.

[250] Gui 2000.

English, in *Yingyu Guojia Xuesheng Xuexi Hanyu Yuyin Nandian Fenxi,* through a small investigation, Ni Yan and Wang Xiaokui discusses regular difficulties existing in the learning of Chinese pronunciation of native English-speaking students[251]. This paper points out that difficulties for students' learning Chinese pronunciation in the English-speaking countries are basically the same. The great difficulties include: confusing u with ü, zh with j and the falling-rising tone and the rising tone.

In her *Taiguoren Xuexi Hanyu Putonghua Yuyin Nandian Fenxi,* Jiang Yinlian explores common problems experienced by Thais when learning Chinese pronunciation and mainly analyzes the difficulties in differentiating j, q, x and zh, ch, sh, and z, c, s as three groups of consonants, reasons for confusing them and teaching strategies to counter these problems[252]. She points out that the learners' propensity for making mistakes in learning pronunciation is closely related to the interference of the mother tongue and advocates the adoption of contrastive teaching method for pronunciation teaching. Finally, in *On Difficulties Existing in Cambodian Overseas Students' Learning Chinese Pronunciation*, Li Ai makes a contrastive analysis of the phonetic systems of Cambodian and Chinese, summarizes some common problems affecting Cambodian's learning of Chinese pronunciation and puts forward some relevant teaching strategies.

6.2. Possible solutions applicable to Hungarian learners

The main concern of this thesis has been to use experimental analysis of the nature and characteristics of Hungarian learners' production, but at this point, before concluding, we put forward the following corresponding teaching strategies, based partly on the available expert literature and partly on the practical experience of the teaching staff of the Department of Chinese at Eotvos Lorand University, Budapest.

[251] Ni and Wang 1992.

[252] Jiang 1997.

6.2.1. Establishing rigorous requirements and maintaining these throughout the learning process.

Zhang Baolin (2005) advises teachers that they should “establish high standards as a strict requirement of the students from the very beginning, and exercise the students in order to master the correct, standard, authentic pronunciation of Mandarin Chinese”[253]. Obviously, this starts with tone teaching. He warns that proper pronunciation cannot be expected at later stages if the speaking Chinese with a foreign accent is not corrected in a timely fashion at the very beginning. Lin Tao (1992) supports this point of view, saying that “Once speaking Chinese with a foreign accent is formed and ingrained, it’s quite hard to correct at a later stage”[254]. After the initial tone teaching stage, even though the learners have acquired related knowledge and understanding of the proper pronunciation method for tone, the tone system is still unstable, it can easily revert to “speaking Chinese with a foreign accent” due to the influence of negative transfer of native language. This means that even if initial teaching produces acceptable results, it will need constant revision, correction and consolidation throughout the learning process.

The investigation described in this thesis was not originally intended to focus on the differential effects of longer or shorter periods of study on the acquisition of correct tone, but it is nonetheless probably worth repeating that compared with the first-year students, the third-year students in the sample showed no noticeable advantages in their ability to produce correct tone even though they had had benefited from a longer period of study and even from the experience of studying Chinese while living in China. One might have expected the latter experience, in particular, to have favoured the natural, unconscious acquisition of tone, but the evidence indicates otherwise. Natural acquisition of certain language features may occur as a result of constant exposure and the existential need to communicate, but this does not seem to hold for syllabic tone. Indeed, it would appear that conscious attention and positive correction are required for the improvement of syllabic tone pronunciation. This suggestion is supported

[253] Zhang 2005.

[254] Lin and Wang 1992.

by the findings of Wang Youmin's tone annotation experiment (1999)[255]. Wang performed a tone annotation survey on 25 Hungarian students studying Chinese in Beijing, in which 12 of the students came from Eotvos Lorand University, like the learners involved in the present study, while 13 of the students came from Budapest College of Economics. The former group, from the ELTE department of Chinese, were majoring in Chinese and received 6-8 lessons a week; the latter group, not majoring in Chinese, received 4- 6 lessons a week. Both groups had studied Chinese for one and a half years in Hungary. Wang's survey shows that the Chinese-studying students who come from Hungary demonstrated the same annotation accuracy without significant difference in tone composite patterns. In other word, it seems that learners' tone pronunciation is not likely to improve automatically no matter how long the study time is or what a study environment it is. On the contrary, it is likely to become fossilized and consolidated in the learner's pronunciation if specific attention is not devoted to dealing with errors of tone on a long-term basis.

Besides practicing tone production, it is necessary to make sure that learners are aware from the very beginning of their studies of the important role played by tone in Chinese as a means of semantic distinction. An error in tone will quite probably result in misunderstanding or actual breakdown in communication. This positive awareness and attention on the learners' part should be combined the teacher's maintenance of rigorous requirements, and should to function as a basic principle throughout Chinese studies.

6.2.2. Giving due importance to imitation, supplemented by necessary tone knowledge.

Imitation and repetitive patter-practice is undoubtedly a highly effective way to teach tone. However, taking the characteristics of the adult-learners into consideration, far greater results are likely to be achieved with the expenditure of considerably less effort if behaviouristic pattern practice is complemented by the necessary theoretical knowledge and by explanation of the pronunciation of the tone. Physically and physiologically speaking, Cheng Meizhen and Zhao Jinming respectively explain that "the variation of tone is only dependent on the pitch"

[255] Wang 1999.

and the latter depends “only on variation of the tightness of vocal cords. The looser the vocal cords, the lower the voice; the tighter the vocal cords, the higher the voice”[256]. Guan Jian (2000) advocates that the teacher involved in tone teaching should focus on how to control the tightness of the vocal cords, guide students to pronounce the four tones correctly and be free of error of tones.[257] For example, the vocal cords shall be kept tight while pronouncing the steady high (*Yinping*) tone; the vocal cords should change from tightness to looseness while pronouncing the falling tone (*Qusheng*). Learners should be taught how to pronounce the rising tone (*Yangping*) and falling-rising tone (*Shangsheng*) on the basis of previous mastery of the level tone and falling tone. The vocal cords become tighter gradually while pronouncing the rising tone; the vocal cords should stay relaxed for a short period and then become tighter while pronouncing the falling-rising tone. Mao Li indicates that by performing a vocal cords-tightening action after the relaxation, the resulting tone can be distinguished from the rising tone and falling tone[258]. Chen Yu (2006) also puts forward his “key points on the assumption of the tone”, considering that the key point in tone teaching is to show students how to control the vocal cords and find out the key points of the tones[259]. As we mentioned before, Zhu Xiaonong proposed the concept of “linguistic goal of tone” (abbreviated to tone goal)[260]. He indicates that the tone goal is the *fēngdiǎn* or “peak point” of a tone, no matter whether it is a level tone, an oblique tone, a high tone or a low tone. The whole level tone itself constitutes a linguistic goal on the basis of the tone shape; the linguistic goal of the rising tone lies in the peak point of the end tail; the linguistic goal of the rising tone lies in the peak point of the end tail; the linguistic goal of a dipping tone or fall-rise tone lies in the start and the end; the linguistic goal of the falling tone is at the beginning or 5%-20% later, at the adjacent peak point of the start and the body. Undoubtedly, the linguistic goal of tone is the key point for mastering a tone. Moreover, the teacher should also help the students to establish “tone consciousness”. This is particular important to Hungarian learners

---

[256] Cheng and Zhao 1997.

[257] Guan 2000.

[258] Mao 2007.

[259] Chen 2006.

[260] Zhu 2005.

whose native language is not a tone language. The "tone consciousness" includes "duration consciousness", "fundamental frequency consciousness", "tone sphere consciousness" and "tone value". The experiment is this thesis shows that the "duration problem" is the most prevalent, and though it is hard to be sure, I may well be that this is a case of negative transfer from the native language. Since Hungarian is not a tone language this kind of duration issue does not arise: speakers of Hungarian do not apply it, consciously or otherwise. It is therefore not part of their repertoire and cannot be transferred as such to Chinese; it is unlikely to acquired naturally and will not be learned at all unless deliberately and overtly taught and consciously practiced.

6.2.3. Sequencing tone teaching from the easier elements to the more difficult

It would seem intuitively sensible to teach easier elements before attempting more difficult ones, but there are very considerable differences of opinion as to which elements should be considered easier or more difficult and therefore in what order they should be taught. Zhao Jinming (1986) suggests that the teaching sequence of tone should be the first tone, the fourth tone, the third tone and the second tone, basing his argument on the physical point of view[261]. Yu Aiqin (1986), quoting the findings of his own research, proposes that the teaching sequence of tones should be "the level tone, the falling tone, the rising tone, and the falling-rising tone"[262], while Shen Xiaonan (1989), on the basis of his research into the tone errors made by American learners of Chinese, claims that the best acquisition sequence would have "the rising tone and the falling-rising tone" coming prior to "the level tone and the falling tone"[263]. Wang Yunjia (1995) quotes experimental research findings in support of a conclusion which is the opposite of that reached by Shen Xiaonan, claiming that "the level tone and the falling tone" should be taught prior to "the rising tone and the falling-rising tone"[264]. To conclude, we can say that even though their favoured tone sequences are different, all the authors follow a same principle, namely that the proper sequence moves from the easier to the

---

[261] Zhao 2006.

[262] Yu 1986: 229-235

[263] Shen 1989.

[264] Wang 1995.

more different.

The research results described in the previous chapter indicate that, for this sample of Hungarian learners, the falling-rising tone and rising tone are difficult to acquire, while the falling tone and level tone is relatively easy. According to the principle of moving from the easier to the more different, teachers shall encourage their students to acquire the level tone and falling tone first, then teach the falling-rising tone and rising tone. One of the most interesting and ideas that has been proposed and is particularly relevant in this area is the "reverse bridge type" tone teaching strategy put forward by Hu Bingzhong in 1979[265]. The "reverse bridge type" teaching method abandons the traditional tone teaching sequence (from the first tone to the fourth tone), and introduces a different sequence which starts with the level *Yángpíng* tone, followed by the falling *Qùshēng* tone, then the falling-rising *Shǎngshēng* tone and ends with the rising *Yángpíng* tone. This sequence may be visualized as having the four tones linked from nose to tail (as shown in Figure 136), which contributes to the tone setting for students, therefore helping them to establish more reasonable individual tone systems.

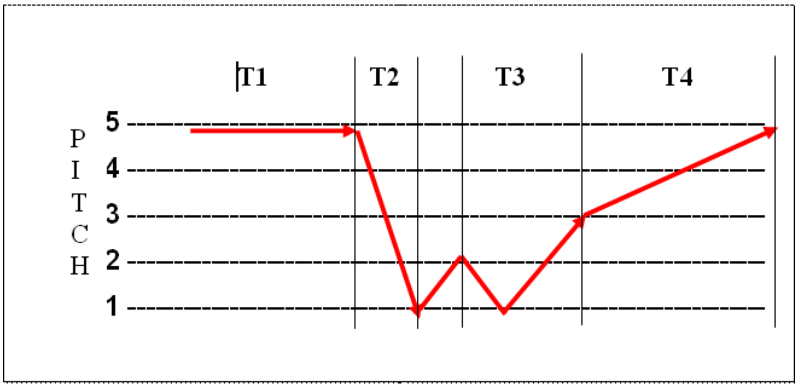


Figure 138. "Reverse bridge" pedagogical sequence of tones proposed by Hu Bingzhong.

6.2.4. Combination of listening and distinguishing

The final goal of tone teaching is for the learners to acquire and perform accurate pronunciation. Teaching experience shows that even for learners with the same native

[265] Hu Bingzhong 1979.

language in the background, the pronunciation of the same tone may result in different levels of difficulty. The same, of course is true of tone teaching. For the same tone, some learners can master it quickly, while others cannot. In the writer's experience of teaching Chinese as a foreign language the active process of articulation, the physical production of tones, is part of the problem, but the greater part of the difficulty can be attributed to learners' ability (or rather, inability) to listen and distinguish; in other words, there is a sort of "dead zone" in the listening and distinguishing of a certain tone. If this is the case, then to continue simply correcting the learner's pronunciation blindly, on a "listen and repeat" basis is a waste of time. Even if the learners succeed in pronouncing the tone accurately once or twice, it happens by accident and is highly unlikely that the correct form will be consolidated. This implies that on the contrary learners should start with extensive training and guided practice in listening and distinguishing, and should be given enough input to eliminate the learners' "dead zone" and gradually develop the ability to listen and distinguish.

6.2.5. Computer-based teaching strategies

There is little doubt that the above-mentioned approaches and methods can take an active role in oral teaching and can complement the traditional phonetic teaching model. However, with the gradual coming-of-age of the computer and multi-media technology, the updating and modernization of Chinese phonetic teaching approaches to include methods based in and taking advantage of information technology must be a matter of importance to any professional in the field, not to mention the learners themselves.

As we mentioned earlier, the PRAAT software, a system for computer-assisted phonetic studies, was made publicly available by its inventor Paul Boersma and David Weenink in 1995. It has been widely used in the field of experimental phonetics, mainly for the purpose of analyzing speakers' pronunciation in terms of their voice pitches, the formant, the fundamental frequency, shape of the fundamental frequency curve, sound intensity and other acoustic properties of their speech. Some applied linguists have put forward corresponding possible solutions based on the analyses of the data obtained by PRAAT and have proposed ways of using them in teaching practice[266]. Compared with the traditional listening and distinguishing method, which obviously depends heavily on auditory perceptive channels, the experimental phonetic method using PRAAT software on computers appears to offer considerable advantages in terms both of scientific accuracy and effective use of visual perceptive channels, not to mention the benefits of intelligent understanding of the issues involved. However, possibly as a result of limited resources available in schools and (according to, for example, Wilson (2008)) because the software may appear dauntingly difficult to use for most teachers and learners, PRAAT has been relatively little used in foreign-language teaching, and as far as we know it has never been applied to the teaching of Chinese to Hungarian learners.

Using PRAAT in the tone teaching process can enable the teacher to convert her abstract theoretical tone knowledge into intuitively understandable visual images by presenting the duration, fundamental frequency, range and other acoustic properties of a certain tone which is borne by syllables. And all that is required for the teaching procedure is a computer which contains the PRAAT software, linked to a digital projector for showing the relevant images.

What is more, there would appear to be no real reason why PRAAT should not also be used in the tone learning process on an individual basis, outside the classroom. It offers an effective way of self-regulation, with learners observing their own performance of tones by listening to recordings of themselves, and watching the instant-PRAAT-generated images and correcting themselves with regard to every aspect of tone error such as the duration error, the fundamental frequency error, the range error and so on. All that is needed for using PRAAT as a means for self-regulation is a computer which contains the PRAAT software. The details of how to use the PRAAT were explained and presented in Chapter 4: they are fairly complex, but can be learned with a modicum of intelligence and due application. It would probably be expedient for a teacher to devote a certain amount of class time to relevant explanation and practice, but today's young generation of "digital natives" are actually much more likely to

[266] Brett, D. 2004:103-113, or Wilson, I. 2008.

find this easily accessible than the older generation of "digital immigrants"[267] to which most of their teachers still belong. Once the learners have mastered the way of using PRAAT, they can correct their tone errors whenever they want. With the "model" native speakers' recordings provided by the teachers, the learners can investigate the acoustic properties of their own tone pronunciations by comparing with those of native speakers. Indeed, PRAAT can be used not only in analyzing and correcting the monosyllabic tone pronunciation, but also in the intonation, and the pronunciations of vowels and consonants as well. In other words, PRAAT can be used in the improvement of the learners' phonetic features at both segemental and suprasegmental levels.

6.3 Possible future directions

6.3.1. Ideal teaching and learning material

An ideal set of teaching materials for Chinese as a foreign language would certainly ensure, among other things, an appropriate hierarchy and order of tone teaching. A survey of the pronunciation teaching materials currently available makes it is obvious that high-quality tone-centered teaching resources are urgently needed. Such a set of teaching materials would cover all the acoustic properties of Chinese tones such as their duration, the fundamental frequency, the range and so on. And it would be arranged in such a way as to follow the teaching sequence from the easier tone to the difficult one, and focus on the establishing of the learner's individual tone system. Using an error-analytical approach such as the one described in this thesis would make it possible to design tone-centered teaching material specifically for Hungarian university learners on the basis of what we have found out, through empirical analysis, about the tone errors they are likely to commit.

However, although the data did enable us to map typical pronunciation patterns for Hungarian learners, it also revealed a certain amount of variation within the group of learners, both in terms of general levels of competence and with regard to specific tones. Moving beyond the

[267] The terms "digital natives" and "digital immigrants", now in common usage, are borrowed from their originator, Marc Prensky. See Prensky, M 2001.

notion of a standardized, generally applicable “package” for collective classroom use, it would also be possible to design individualized analytical systems to identify and analyse the weaknesses and strengths of particular learners. Depending on circumstances, teachers would be able perform this analysis and then choose appropriate materials from a universal resource bank, which they could then recommend to each learner for use in a tailored supplementary course of self-study.

Moreover, one might question how far the teacher ought to be in charge, or even involved at all, in this process. At primary school level the teacher’s impetus, skill and guidance would certainly be required at every stage and this may still be true at secondary level. At tertiary level, however, it would be reasonable to expect learners to be motivated, competent and confident enough to look after themselves in this respect. Indeed, a somewhat similar idea has been operationalised at the University of Minnesota, with students being offered online guidance in the use of PRAAT for self-study[268], though there appears to be no data on the actual success of the initiative. Something along these lines

6.3.2. The combination of PRAAT and other software

Undoubtedly, PRAAT itself constitutes an extremely valuable resource for use in both the teaching and the learning processes. However, as it stands it is somewhat cumbersome and certainly time-consuming to use it in the manner described in Chapter 4 above. In part this is due to its very broad range of possible uses – our research design involved a considerable number of successive choices among many available options, normalization and so on. Presumably a competent information technologist would be able to create a series of macros or a script which would reduce the need for so many operations, and would also apply the necessary calibrations automatically. If the PRAAT software were then associated with

[268] Baepler, P., Ekegren, C., Marsh, L., Rondeau, A. and Thompson, P. 2006.

another application that could provide an instantaneous graphic display of the results generated by PRAAT learners could receive immediate evidence of their own performance in a form that would be very easy to grasp. This visible evidence could be put to various uses; most obviously, it could serve for comparison of recordings made by learners with results generated from data provided by native-speaker models, and could provide opportunities for learners to evaluate their own performance.

This in itself is not a particularly revolutionary idea: the general principle of transforming auditory into visual data for teaching and leaning purposes via electronic means using, for example, oscillographs, has been discussed for half a century[269], and its application to F0 pitch contours for intonation also has a respectable tradition[270]. However, the practical solutions devised so far, though ingenious, are not fully satisfactory. With reference to the specific issue of how PRAAT can be used to compare learners' and teachers' pronunciation, Paul Boersma, one of the creators of PRAAT, stated that "this is one of the most asked questions. It is also one of the most difficult problems, and it has not generally been solved.[271]" In the same post, Boersma recommends an open-source script devised and published on the internet by Rob van Son[272]. This script is specifically designed to help learners of Mandarin Chinese, and includes material for practicing tones. There is a relatively simple user interface, including basic settings for male and female voices at pre-set pitch ranges, and rather than using live recordings of real speakers, the software provides synthetic reference tone(s) produced from the pinyin notation, using various ready-made sets of words and phrases which carry these tones. The user can listen to these tones and also see them in graphic form on the screen as contours; s/he can then attempt to reproduce both the sound and the visible contour, with his/her own contours superimposed on the "ideal" ones. The audible tones sound somewhat odd: indeed, an impromptu sample of five Chinese native speakers (all

---

[269] Vardanian, R. 1964: 109–118.
[270] Wichern, P., Boves, L. 1980: 53–63.
[271] Comment made on a PRAAT-users' forum, 5th January 2011, available at: http://uk.groups.yahoo.com/group/praat-users/message/5180 (accessed October 3rd, 2013)
[272] Van Op 2013.

guest students at ELTE, aged between 20 and 30) declared that they could neither identify nor understand more than half of the words and phrases that they heard. This suggests that the software should be employed with some caution, and teachers using this system in class might prefer a system that records their own speech and retains it for comparison. On the other hand, precisely because the tone contours are artificially produced, they are very clear and it must be admitted that a standard, unvarying model might be preferable from the point of view of the individual learner using the program for self-study purposes. Unfortunately the standard display does not include precise details about pitch, range or duration. The figure below is a screenshot of the display, with the green line being the software's ideal, artificially produced, contour and the red line being the learner's attempt to reproduce it.

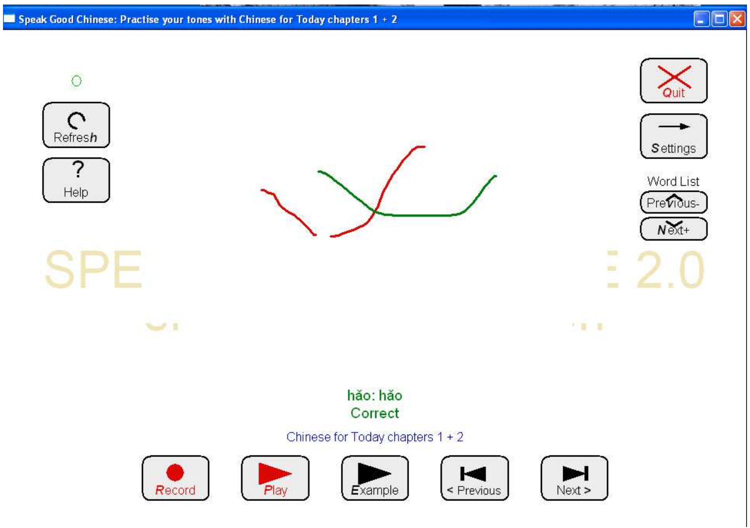


Figure 138. Screenshot of "Speak Good Chinese" display.

In the research described in this thesis we exported the results of data produced by PRAAT to Microsoft Excel for tabulation and in order to produce the graphic versions shown in Chapter 5 above. The results are clear and easy to interpret, and include not only tracings of tone contours but also precise information about pitch, duration and so on. They can also be displayed in different forms and combinations for various purposes. Again, the process of transferring PRAAT data into Excel and producing the tables involved a great many choices and operations which take some time to learn and are fairly slow to implement, but it should be possible to automate much of the procedure such that it can take place instantaneously, providing learners and teachers with real-time feedback as a basis for correction, further practice and development. The visual displays could be further enhanced by a system of automatically triggered warning messages, visual, auditory or both, which would appear when the learner makes a tone error in terms of the duration, the fundamental frequency, the range and other acoustic properties. Based on the above functions, the learners could correct their tone pronunciation through comparison with the “model” native speakers’ voice, thus, the acquisition of the proper tone pronunciation would become possible.

6.4 Further directions for research

The teaching of pronunciation in Chinese is a vast field, and ideally one would want to investigate all of the very wide variety of different errors made by learners. Apart from errors of monosyllabic tone, there are issues involving disyllabic tones and sentence intonation patterns. And going beyond tones, there are a great many other kinds of potential pronunciation errors concerning stress and rhythm, as well as vowel errors and consonant errors that deserve investigation. In other words, research intended to provide really comprehensive assistance to the Chinese as a foreign language teaching profession would need cover all the possible pronunciation errors at both segmental and suprasegmental levels.

Such an endeavour is clearly beyond the scope of a thesis such as this one: it was essential not only to narrow down the field to manageable proportions, but also to prioritise, and this largely explains the structure of the thesis itself. After a brief general introduction of Hungarian and Chinese, the thesis offers an in-depth comparison between the segmental and suprasegmental phonetic systems of the two languages, leading to a discussion of which of the identified differences are likely to cause problems both in terms of acquisition by learners and in terms of communicative efficiency. Applying these criteria for the purposes of selection, the comparative analysis concluded with the choice of monosyllabic tone as a focus for the rest of the thesis. Undoubtedly, the monosyllabic tone is important enough to be the main focus, since on the one hand, Chinese is a tonal language where the same syllable with different tone usually carries different meanings; thus inaccurate use of tones carries the risk of miscomprehension and communicative breakdown; on the other hand, the monosyllabic tone is the basic of disyllabic tone and suprasegmental intonation, and the latter two features cannot really be attempted until the former has been mastered.

As a result, the thesis continued with the detailed analysis and discussion of recorded data concerning learners' and native speakers' pronunciation of monosyllabic tones. Nonetheless, it is already possible to foresee some of the future directions of research to be undertaken in the same context. Indeed, data have already been collected from the same 12 Hungarian subjects, consisting of a substantial body of high-quality recordings of disyllables, sentences, vowels and consonants, all of which have been archived for the purposes of investigating the whole range of pronunciation features of the Hungarian learners. These recordings will be used for the further studies. The disyllables, sentences, vowels and consonants will be investigated based on the experimental phonetics, mainly by using the same PRAAT software as before.

One possible criticism of the research described in this thesis is that it focusses exclusively on single syllables (individual monosyllabic lexical items) pronounced in isolation, rather than on syllables which occur in the course of longer stretches of speech, which entails two notable drawbacks. The first is the fact that although this approach makes it possible for syllabic tone to be clearly measured (except when the "creaky voice" phenomenon occurs) without interference by suprasegmental intonation and other prosodic features, it does not reveal how these learners pronounce tone when they actually use Chinese to communicate. Thus the present research should be seen only as a baseline study which eventually needs to be continued using data collected first through scripted recordings of longer utterances (designed to contain target syllable-linked tones), and then through recordings of unscripted utterances occurring in real communicative situations. Secondly, the focus on isolated syllables means effectively excluding the sandhi (*biàndiào*) effect described in 3.4.3. above. Sandhi refers to the way in which single tones affect each other when juxtaposed in words, phrases and sentences. A common example of sandhi is found in the greeting phrase 你好 *nǐhǎo*. In isolation, both of the words in the phrase carry the third (S*hăngshēng*) tone, but when they are combined in the greeting phrase the sandhi effect causes the word *nǐ* to be pronounced *ní* instead. A valuable intermediate stage between the "isolated single syllable" approach applied in the present research and the more authentic “complete utterance” or “unscripted communication” approaches would be to select bisyllabic expressions (like *nǐhǎo*) where sandhi occurs in native-speaker language and use the these expressions to assess the learners’ “use” of sandhi.

Undoubtedly, there is still much to be done, and the present thesis cannot claim to be more than a first step towards the establishment of a solid body of research covering many aspects of the pronunciation of Chinese by Hungarian leaners. Eventually, it is to be hoped that the findings concerning these aspects will eventually be available to form the basis of teaching and learning methods for Hungarian university learners. What is more, in order to enable the better analysis of the possible effects of negative transfer from the native language in terms of vowels and consonants, the author also made up some “Hungarian” words which correspond to the Chinese monosyllables in the test paper although they do not really exist in Hungarian: the latter fact in itself should not cause problems because (unlike English) Hungarian has a very straightforward spelling system which nearly always corresponds directly to spoken pronunciation, with very few exceptions. Because the focus of this thesis was more on identifying and prioritizing errors with a view to correcting them, than on establishing the precise reasons for their occurrence, the “fake” Hungarian words were not used in this thesis, but they will be used in the further studies, focusing not only on what sort of errors characterize Hungarian learners’ speech, but also on the underlying reasons for those errors. This may seem an ambitious goal, and indeed it is: only when all the acoustic properties of the Chinese phonetic system have been adequately investigated with regard to their learning and use by Hungarian learners will the entirely successful acquisition of proper pronunciation become possible for our students.

List of References

1958 *Scheme for the Chinese Phonetic Alphabet* 汉语拼音方案. Hanyu Pinyin Fang'an Weiyuanhui Zhiding.

Adjemian, C. (1976): "On the nature of interlanguage systems." *Language learning* 26: 297-320.

Baepler, P., Ekegren, C., Marsh, L., Rondeau, A. and Thompson, P. (2006): Praat Language Lab. available at: http://www.tc.umn.edu/~parke120/praatwebfiles/intro.htm (accessed September 23, 2013)

Brett, D. (2004): "Computer generated feedback on vowel production by learners of English as a second language." *ReCALL* 16: P.103-113, or Wilson, I. (2008): "Using Praat and Moodle for Teaching Segmental and Suprasegmental Pronunciation." In: *Proceedings of the 3rd International WorldCALL Conference: Using Technologies for Language Learning* (WorldCALL 2008).

Boersma, P. & Weenink, D (2013): PRAAT: Doing phonetics by computer [Computer program: version 5.3.51]. Available at: http://www.praat.org/ (retrieved 30 May 2013)

Boersma, P. & Weenink, D (2013): PRAAT: Doing phonetics by computer [Computer program: version 5.3.51]. Available at: http://www.praat.org/ (retrieved 30 May 2013).

Bowen, T. & Marks, J. (1992): *The pronunciation book: Student-centered activities for pronunciation work*. Harlow, Longman.

Brière, E. (1968): *A psycholinguistic study of phonological interference*. The Hague, Mouton.

Cao Jianfen 曹剑芬 (2002): "Hanyu Shengdiao yu Yudiao de guanxi" 汉语声调与语调的关系. *Zhongguo Yuwen* 中国语文 3, 195-203.

Catford, John C. (1977): *Fundamental Problems in Phonetics*. Edinburgh, Edinburgh University Press. P. 152-153

Celce Murcia, M. (1996): *Pronunciation teaching: History and scope*. Available at:

http://web.ntpu.edu.tw/~language/workshop/method.pdf (accessed June 20, 2013)

Chao, Yuen Ren 赵元任 (1983): *Tongzi Fang'an* 通字方案 (*A Project for General Chinese*). Beijing, Shangwu Yinshuguan

Chen Yu 陈彧 (2006): "Sugelan Liuxuesheng Hanyu Putonghua Danziyin Shengdiao Yingao de Shiyan Yanjiu" 苏格兰留学生汉语普通话单字音的实验研究. *Shijie Hanyu Jiaoxue* 世界汉语教学 2, 99-111.

Cheng Meizhen 程美珍 (1997) and Zhao Jinming 赵金铭 (1997): "Jichu Hanyu Yuyin Jiaoxue de Ruogan Wenti" 基础汉语语音教学的若干问题. In: Zhao Jinming 赵金铭 *Yuyin Yanjiu yu Duiwai Hanyu Jiaoxue* 语音研究与对外汉语教学. Beijing, Beijing Yuyan Wenhua Daxue Chubanshe, 1997.

Cheng Tang 程棠 (2000): *Duiwai Hanyu Jiaoxue Mudi Yuanze Fangfa* 对外汉语教学目的原则方法. Beijing: Huayu Chubanshe.

Cheng Tang 程棠 (2000): *Duiwai Hanyu Jiaoxue Mudi Yuanze Fangfa* 对外汉语教学目的原则方法. Beijing: Huayu Chubanshe.

Cheng Tang 程棠 (1996): "Duiwai Hanyu Yuyin Jiaoxue zhong de Jige Wenti" 对外汉语语音教学中的几个问题. *Yuyan Jiaoxue yu Yanjiu* 语言教学与研究 3, 4-17.

Chomsky, N. (1957): A Review of B. F. Skinner's Verbal Behavior. In Jakobovits, L & Miron, M. (eds.) *Readings in the Psychology of Language*, Prentice-Hall, 1967, 142-143.

Corder, S. P. (1967): "The significance of learners' errors'." *International Review of Applied Linguistics* 5: 161-170.

Corder, S. P. (1967): "The significance of learners' errors". *International Review of Applied Linguistics* 5: 161-170.

Dai Xiaoxue 戴晓雪 (2004): "Chaoyinduan Chengfen de Biaoyi Gongneng Jiqi zai Duiwai Hanyu Jiaoxue zhong de Yingyong" 超音段成分的表意功能及其在对外汉语教学中的应用. *Diqijie Guoji Hanyu Jiaoxue Taolunhui Lunwen Xuan* 第七届国际汉语教学讨论会论文选. Beijing, Beijing Daxue Chubanshe.

Dong Yuguo 董玉国 (1997): "Dui Riben Xuesheng Biyunmu Yin de Jiaoxue" 对日本

学生鼻韵母音的教学. *Shijie Hanyu Jiaoxue* 世界汉语教学 4, 66-70.

Du Qinhuan 杜秦还 (1992): “Tantan Xiandai Hanyu Shengdiao Jiaoxue” 谈谈现代汉语声调教学. *Yuyan Jiaoxue yu Yanjiu* 语言教学与研究 1, 75-81.

Dulay, H. C., & Burt, M. K. (1974): “Natural sequences in child second language acquisition. *Language Learning*” 24: 37-53.

Ellis, N. (2006): “Selective attention and transfer phenomena in L2 acquisition: Contingency, cue competition, salience, interference, overshadowing, blocking, and perceptual learning”, *Applied Linguistics* 27/2:164-194.

Ellis, R. (1997): Second Language Acquisition. Oxford, Oxford University Press.

Feng Shengli 冯胜利 (1998): “Lun Hanyu de ‘Ziran Yinbu’” 论汉语的“自然音部”. *Zhongguo Yuwen* 中国语文 1, 40-47.

Feng Zhichun 冯志纯 (2008): *Xiandai Hanyu* 现代汉语. Chongqing, Xinan Shifan Daxue Chubanshe.

Fisiak, J. (1981): *Contrastive linguistics and the language teacher*. New York, Pergamon. P.1

Fraser, H. (2001): *Teaching pronunciation: A handbook for teachers and trainers, Australian Department of Education Training and Youth Affairs*. p.18. Available at:

Fries, C. (1945): *The teaching and learning of English as a foreign language*. Ann Arbor, MI, University of Michigan Press.

Gass, S. & Selinker, L. (1994): *Second language acquisition: An introduction course*. New Jersey, Lawrence Erlbaum Associates.

George, H (1972): *Common errors in language learning*. Rowley, Mass, Newbury House.

Gu Zheng & Wu Zhongwei 顾筝，吴中伟 (2005): “Liuxuesheng Rumen Jieduan Yuyin Jiaoxue Yanjiu” 留学生入门阶段语音教学研究. *Yunnan Shifan Daxue Xuebao* 云南师范大学学报 第 3 卷第 2 期, 1-6.

Guan Jian 关键 (2000): “Shengdiao Jiaoxue Gaige Chutan” 声调教学与改革初探. *Yuyan Jiaoxue yu Yanjiu* 语言教学与研究 4, 51-54.

Gui Mingchao 桂明超 (2000): “Meiguo Yingyu Yudiao Dui Meiguo Xuesheng Xuexi Hanyu Putonghua Shengdiao de Ganrao” 美国英语语调对美国学生学习汉语普通话声调的干扰. *Shijie Hanyu Jiaoxue* 世界汉语教学 1, 89-93.

Guo Jinfu 郭锦桴 (1993): *Hanyu Shengdiao Yudiao Chanyao yu Tansuo* 汉语声调语调阐要与探索. Beijing, Beijing Yuyan Xueyuan Chubanshe.

Harley, B. (1980): “Interlanguage units and their relations.” *Interlanguage Studies Bulletin* 5: 3-30.

Hartley, B. & Viney, P. (1978): *Streamline Departures*. Oxford, Oxford University Press.

Hornby, A. & Zhang, F. (1984): *The Oxford Advanced Learner's Dictionary of Current English with Chinese Translation*. Oxford, Oxford University Press.

Howie, John Marshall (1976): *Acoustical Studies of Mandarin vowels and tones*. Cambridge, MA, Cambridge University Press; Rose, Phil (1981): *An Acoustically based Phonetic Description of the Syllable types in the Zhenhai Dialect*. PhD dissertation. The University of Cambridge.

http://helenfraser.com.au/downloads/HF%20Handbook.pdf (accessed May 23 2012).

http://www.uwo.ca/sogs/WJGR/2005/WJGR2005_v12_p113_Zhang.htm (retrieved June 20 2013) p

Hu Bingzhong 胡炳忠 (1979): “Sisheng Liandu yu ‘Biandiao Daibiaozi’——Jiaoxue Biji” 四声连读与“变调代表字”——教学笔记. *Yuyan Jiaoxue yu Yanjiu* 语言教学与研究 1, 23-27.

Hu Mingyang 胡明扬 (1987): *Beijing Hua Chutan* 北京话初探. Beijing, Shangwu Yinshuguan.

Huang Zhenji 黄贞姬 (1999): “Qiantan Dui Hanguo Liuxuesheng de Chuji Hanyu Yuyin Jiaoxue” 浅谈对韩国留学生的初级汉语语音教学. *Shuangyu Shuangwen Hua Luncong (Disi Ji)* 双语双文化论丛. Jinlin, Yanbian Daxue Chubanshe.

Ji Xiuqing 季秀清 (1998): “Yuju Zhongyin yu Kouyu Jiaoxue” 语句重音与口语教学.

Yuyan Wenhua Jiaoxue Yanjiu Jikan (Di'er Ji) 语言文化教学研究集刊，第二辑. Beijing, Huayu Jiaoxue Chubanshe.

Jiang Xinlian 蒋印莲 (1997): "Taiguoren Xuexi Hanyu Putonghua Yuyin Nandian Fenxi" 泰国人学习汉语普通话语音难点分析. *Diwujie Guoji Hanyu Jiaoxue Taolunhui Lunwen Xuan* 第五届国际汉语教学讨论会论文选. Beijing, Beijing Daxue Chubanshe

Jiang Yiliang 蒋以亮 (1998): "Yuliu Jiaoxue Chutan" 语流教学初探. *Yuyan Jiaoxue yu Yanjiu* 语言教学与研究 4, 95-103.

Jiao Liwei 焦立为 (2001): "Xiandai Yuyinxue de Fanfxiang---Fang Lin Tao Jiaoshou" 现在语音学的方向——访林焘教授. *Yuyan Jiaoxue yu Yanjiu* 语言教学与研究 2, 41-46.

Jin Song 劲松 (1992): "Beijinghua Yudiao de Shiyan Tansuo" 北京话语调的实验探索. *Yuyan Jiaoxue yu Yanjiu* 语言教学与研究 2, 71-96.

Kálmán, L. & Nádasdy, Á. (1994): "A hangsúly stress." In Kiefer, F. & É. Kiss, K. (eds.) *The syntactic structure of Hungarian. Syntax and Semantics 27*. San Diego and New York, Academic Press, 393-467.

Kenworthy, J. (1987): *Teaching English pronunciation*. London and New York, Longman. P15

Kiefer, F. & É. Kiss, K. (eds.) (1994): *The syntactic structure of Hungarian. Syntax and Semantics 27*. San Diego and New York, Academic Press.

Kingdon, R. (1958): *The Groundwork of English Intonation*. London, Longman

Kornai, A. (1994): "On Hungarian Morphology" (*Linguistica, Series A: Studia et Dissertations*, 14. Budapest, Linguistics Institute of HAS.

Krashen, S. & Terrel, T. (1983). "The natural approach: Language acquisition in the classroom." In: Richards, J. & Rodgers, T. (eds.) *Approaches and methods in language teaching* (2nd ed.). Cambridge, New York, Cambridge University Press.

Krashen, S. (1981): *Second language acquisition and second language learning*. Oxford, Pergamon

Lado, R. (1957): *Linguistics across cultures: Applied linguistics for language teachers*.

Ann Arbor MI, University of Michigan Press. P.vii.

Lado, R. (1957): *Linguistics across cultures: Applied linguistics for language teachers*. Ann Arbor MI, University of Michigan Press.

Larsen Freeman, D. (2000): *Techniques and principles in language teaching*. Oxford, Oxford University Press.

Larsen-Freeman, D. & Long, M. (1991): *An introduction to second language acquisition research*. London, Longman.

Lehiste, Ilse (1970): *Suprasegmentals*. Cambridge, MA, MIT Press. P.68

Lenneberg, J. (1967): *Biological Foundations of Language*. New York: John Wiley & Sons.

Lisker, L. & Abramson, A.S. (1964). A cross-language study of voicing in initial stops: acoustical measurements. Word, 20, 384-422

Li Fang-Kuei (1972). “Language and dialects in China”. *Free China Review* 22,5. Norman, J. (1988): *Chinese*. Cambridge, Cambridge University Press

Li Ming 李明 (1996): “ABshi Shuangyinjie Xingrongci Chongdieshi de Duyin Kaocha” AB式双音节形容词重叠式的读音考察. *Yuyan Jiaoxue yu Yanjiu* 语言教学与研究 1, 73-82.

Lin Maocan 林茂灿|(2004): “Hanyu Yudiao yu Shengdiao” 汉语语调与声调. *Yuyan Wenzi Yingyong* 语言文字应用 3.

Lin Tao, Wang Lijia 林焘，王理嘉 (1992): *Yuyinxue Jiaocheng* 语音学教程. Beijing, Beijing Daxue Chubanshe.

Lin Tao 林焘 (1996): “Yuyin Yanjiu he Duiwai Hanyu Jiaoxue” 语音研究和对外汉语教学. *Shijie Hanyu Jiaoxue* 世界汉语教学 3, 18-21.

Lin Yigao and Wang Gongping 林奕高，王功平 (2005): “Yinni Liuxuesheng Xide Hanyu Seyin he Secayin Shiyan Yanjiu” 印尼留学生习得汉语塞音和塞擦音实验研究. *Yuyan Jiaoxue yu Yanjiu* 语言教学与研究 4, 59-65.

Liu Chuanping 刘川平 (1996): “Putonghua Shangsheng Benzhi Jiqi Jiaoxue de Zai Renshi” 普通话上声的本质及其教学的再认识. *Waiyu yu Waiyu Jiaoxue* 外语与外语教学 6, 60-63.

Liu Juan 刘娟 (1997): “Qingsheng de Benzhi Tezheng” 轻声的本质特征. *Yuyan Jiaoxue yu Yanjiu* 语言教学与研究 1, 143-151.

Liu Xun 刘珣 (2000): *Duiwai Hanyu Jiaoyuxue Yinlun* 对外汉语教育学引论. Beijing, Beijing Yuyan Wenhua Daxue Chubanshe.

Liu Xun 刘珣 (2002): *Xin Shiyong Hanyu Keben* 新实用汉语课文 (*New Practical Chinese Textbooks*). Beijing, Beijing Yuyan Wenhua Daxue Chubanshe.

Lü Bisong 吕必松 (2007): *Hanyu he Hanyu Zuowei Di’er Yuyan Jiaoxue* 汉语和汉语作为第二语言教学. Beijing, Beijing Daxue Chubanshe.

Lu Jianji 鲁健骥 (1993): “Zhongjieyu Yanjiu Zhong de Jige Wenti” 中介语研究中的几个问题. *Yuyan Wenzi Yingyong* 语言文字应用 2, 21-25.

Lu Jianji 鲁健骥 (1984): “Zhongjieyu Lilun yu Waiguoren Xuexi Hanyu de Yuyin Pianwu Fenxi” 中介语理论与外国人学习汉语的语音偏误分析. *Yuyan Jiaoxue yu Yanjiu* 世界汉语教学 3, 44-56.

Lu Jianji 鲁健骥 (2010): “Duiwai Hanyu Yuyin Jiaoxue Jige Jiben Wenti de Zai Renshi” 对外汉语语音教学几个基本问题的再认识. *Dalian Xueyuan Xuebao* 大连学院学报 5, 1-4.

Lu Wo 卢偓 (2001): “Xiandai Hanyu Yinjie de Shuliang yu Goucheng Fenbu” 现代汉语音节的数量与构成分布. *Yuyan Jiaoxue yu Yanjiu* 语言教学与研究 6.

Luo Changpei and Wang Jun, 罗常培、王均 (2002): *Putong Yuyanxue Gangyao (Xiuding Ben)* 普通语音学纲要 (修订本). Beijing, Shangwu Yinshuguan. 156-157.

Ma Yanhua 马燕华 (2000): “Chuji Hanyu Shuiping Liuxuesheng de Disansheng Tingbian Fenxi” 初级汉语水平留学生的第三声听辨分析. *Beijing Shifan Daxue Xuebao* 北京师范大学学报 6, 110-116.

Mao Li 毛丽 (2007): “Shengdiao Pianwu yu Duiwai Hanyu Shengdiao Jiaoxue Yanjiu Zongshu” 声调偏误与对外汉语声调教学研究综述. *Hunan Diyi Shifan Xuebao* 湖南第一师范学报 2, 119-122.

Mao Shizhen 毛世桢 (2002): “Guanyu Duiwai Hanyu Yuyin Jiaoxue de Fansi” 关于对外汉语语音教学的反思. *Duiwai Hanyu Jiaoyan Luncong* 对外汉语教研论

从 2. Shanghai, Huadong Shifan Daxue Chubanshe.

Meng Zhuyi 孟柱亿 (2000): “Shilun Yuyin Jiaoxue zhong de Wudao Wenti” 试论语音教学中的误导问题. *Diliujie Guoji Hanyu Jiaoxue Taolunhui Lunwen Xuan* 第六届国际汉语教学讨论会论文选. Beijing, Beijing Daxue Chubanshe

Meng Zimin 孟子敏 (1997): “Duiwai Hanyu Yuyin Jiaoxue Zhong Shiyong Hanyu Pinyin Fang’an de Jige Wenti” 对外汉语语音教学中使用《汉语拼音方案》的几个问题. *Yuyin Yanjiu yu Duiwai Hanyu Jiaoxue* 语音研究与对外汉语教学.Beijing, Beijing Yuyan Wenhua Daxue Chubanshe.

Nemser, W. (1971): “Approximative systems of foreign language learners.” *International Review of Applied Linguistics* 9: 115-24.

Ni Yan and Wang Xiaokui 倪彦，王晓葵 (1992): “Yingyu Guojia Xuesheng Xuexi Hanyu Yuyin Nandian Fenxi” 英语国家学生学习汉语语音难点分析. *Hanyu Xuexi* 汉语学习 2, 47-50.

Norman, J. (1988): *Chinese*. Cambridge, Cambridge University Press.

Oller, J. & Ziahosseiny, S. (1970): “The contrastive analysis hypothesis and spelling errors.” *Language Learning* 20: 183-189.

Osgood, C. (1953): *Method and theory in experimental psychology*. New York, Oxford University Press.

PRAAT-users’ forum, 5th January 2011, available at: http://uk.groups.yahoo.com/group/praat-users/message/5180 (accessed October 3rd, 2013)

Porter, D. and Garvin, S. (1989): “Attitudes to pronunciation in EFL”. Speak Out! 5: 8-15. Dalton, C. & Seidlhofer, B. (1994): *Pronunciation*. Oxford: Oxford University Press

Prator, C. (1967): “Guidelines for planning classes and teaching materials.” *Working papers in English as a Second Language. Matter, methods, materials.* Los Angeles, Department of English, University of California.

Richards, J. & Rodgers, T. (1986): *Approaches and Methods in Language Teaching*. Cambridge, Cambridge University Press.

Richards, J. (1971): “Error analysis and second language strategies.” *Language*

Sciences 17: 12-22

Roach, P. (1982): “On the distinction between stress-timed and syllable-timed languages”. In Crystal, D. (ed.) *Linguistic Controversies*. London, Edward Arnold 73-9; Crystal, D. (1995). “Documenting rhythmical change”. In Lewis, J. (ed.), Studies in general and English phonetics: Essays in honour of Professor J. D. O’Connor. London, Routledge 174-9.

Rogers, C. (1969): *Freedom to learn: A view of what education might become*. Columbus, OH, Charles E. Merrill.

Rose, Phil (1981): *An Acoustically based Phonetic Description of the Syllable types in the Zhenhai Dialect*. PhD dissertation, The University of Cambridge.

Ronald J. Baken (1987): Clinical measurement of speech and voice. Boston, College-Hill Press

Schmidt, R. (1990): “The role of consciousness in second language learning”. *Applied Linguistics*, 11, 129-158.)

Selinker, L. (1972): “Interlanguage.” *International Review of Applied Linguistics* 10: 209-31.

Shen Jiong 沈炯 (1994): “Hanyu Yudiao Gouzao he Yudiao Leixing” 汉语语调构造和语调类型. *Fangyan* 方言 3, 221-228.

Shen Jiong 沈炯 (1995): “Hanyu Yingao Xitong de Youshengxing he Qubiexing” 汉语音高系统的有声性和区别性. *Yuyan Wenzi Yingyong* 语言文字应用 2, 13-18.

Shen Li’na and Yu Yanhua 沈丽娜，于艳华 (2009): “Duiwai Hanyu Yuyin Jiaoxue zhong Shengmu Jiaoxue Yanjiu” 对外汉语语音教学中声母教学研究. *Changchun Ligong Daxue Xuebao (Shehui Kexue Ban)* 长春理工大学学报(社会科学版) 3, 507-508.

Shen Xiaonan 沈晓楠 (1989): “Guanyu Meiguoren Xuexi Hanyu Shengdiao” 关于美国人学习汉语声调. *Shijie Hanyu Jiaoxue* 世界汉语教学 3, 158-168.

Sheng Yan 盛炎 (2006): *Yuyan Jiaoxue Yuanli* 语言教学原理. Chongqing, Chongqing Chubanshe.

Siptár, P. & Törkenczy, M. (2000): *The phonology of Hungarian*. Oxford, Oxford

University Press.

Huang Borong and Liao Xudong 黄伯荣、廖旭东 (1991): *Xiandai Hanyu* 现代汉语. Beijing, Gaodeng Jiaoyu Chubanshe

Skinner, B. (1957): *Verbal behaviour*. New York, Appleton-Century-Crofts.

Song Chunyang 宋春阳 (1998): "Tan Dui Hanguo Xuesheng de Yuyin Jiaoxue---Nanyin ji Duice" 谈对韩国学生的语音教学——难音及对策. *Nankai Xuebao* 南开学报 3, 72-76.

Song Yidan 宋益丹 (2009): "Duiwai Hanyu Shengdiao Jiaoxue Celüe Tansuo" 对外汉语声调教学策略探索. *Yuyan Jiaoxue yu Yanjiu* 语言教学与研究 3, 48-53.

Stenson, N. (1974): "Induced errors." In Schumann, J and Stenson, N (eds.) *New frontiers in second language learning*. Newbury House, Rowley, Mass.

Stockwell, R., Bowen, J. & Martin, J. (1965): *The grammatical structures of English and Spanish*. Chicago, University of Chicago Press.

Swan, M. (2010): "Striking a balance". *English Teaching Professional* 70.p.4 37. Richards, J. (2005):"Second thoughts on teaching listening". RELC 36.1. 85-92

Tarone, E. (1983): "On the variability of interlanguage systems." *Applied Linguistics* 4: 142-162. P.144

Prensky, M (2001): "*On the Horizon*." MCB University Press, 9,5.

Trask, R. (1996): *Dictionary of phonetics and phonology*. Taylor & Francis (Routledge).

Van Op (2013) SpeakGoodChinese at http://www.speakgoodchinese.org/articles.html, where the complete script is also available for download. (accessed October 30, 2013)

Vardanian, R. (1964): "Teaching English intonation through oscilloscope displays." *Language Learning* 14:109–118

Wang Lijia 王理嘉 (2005): "Hanyu Pinyin Fang'an yu Shijie Hanyu Yuyin Jiaoxue" 《汉语拼音方案》与世界汉语语音教学. *Shijie Hanyu Jiaoxue* 世界汉语教学 2, 5-12.

Wang Lixiang and Jiang Hailing 王丽香，姜海陵 (2004): “Duiwai Hanyu Jiaoxue zhong Jianyuju de Yuyin Qiefen Yanjiu” 对外汉语教学中兼语句的语音切分研究. *Yunnan Shifan Daxue Xuebao* 云南师范大学学报 4.

Wang Li 王立 (1998): “Cong ‘Danyin Dongci + Danyin Mingci’ Jiegou de Pinxie Kan Zhengcifa de Keguan Yiju” 从“单音动词+单音名词”结构的拼写看正词法规则的客观依据. *Yuyan Wenzi Yingyong* 语言文字应用 1, 19-24.

Wang Youmin 王又民 (1998): “Xiongyali Xuesheng Hanyu Shuangshengci Shengdiao Biaozhu Lianghua Fenxi” 匈牙利学生汉语双声词声调标注量化分析. *Shijie Hanyu Jiaoxue* 世界汉语教学 2, 90-97.

Wang Yunjia 王韫佳 (1995): “Ye Tan Meiguoren Xuexi Hanyu Shengdiao” 也谈美国人学习声调. *Yuyan Jiaoxue yu Yanjiu* 语言教学与研究 3, 126-140.

Wang Yunjia 王韫佳 (1996): “Qingsheng Yingao Suoyi” 轻声音高琐议. *Shijie Hanyu Jiaoxue* 世界汉语教学 3, 50-53.

Wang Yunjia 王韫佳 (2002): “Hanyu Yuyin Yanjiu yu Hanyu Yuyin Jiaoxue Jiekou zhong de Ruogan Wenti” 汉语语音研究与汉语语音教学借口中的若干问题. *Duiwai Hanyu Luncong (Di’er Ji)* 对外汉语论丛，第 2 辑. Shanghai, Shanghai Waiyu Jiaoyu Chubanshe

Wardhaugh, R. (1970): “The contrastive analysis hypothesis.” *TESOL Quarterly* 4: 123-30.

Wen Xiaohong 温晓虹 (1992): “Xide Diyi Yuyan he Di’er Yuyan Bijiao” 习得第一语言和第二语言比较. *Yuyan Jiaoxue yu Yanjiu* 语言教学与研究 3, 49-65.

Whorf, B. (1941): “Language, mind and reality.” In: Carroll, J. (ed.) (1956). *Language, thought, and reality: Selected writings of Benjamin Lee Whorf*. Cambridge, Mass., Technology Press of Massachusetts Institute of Technology.

Wichern, P., Boves, L. (1980): “Visual feedback of F0-curves as an aid in learning intonation-contours.” *Proceedings of the Institute of Phonology*, Nijmegen, Vol. 4, 53–63.

Wu Huainan 吴淮南 (2003): “Hanyu Ciyu Goucheng de Yinjie Yinsu he Jiezou Yugan de Peiyang” 汉语词语构成的音节因素和节奏语感的培养. *Nanjin Shifan*

Daxue Xuebao (Zhexue Ban) 南京师范大学学报（哲学版）3, 142-148.

Wu Jiemin 吴洁敏(2000): “Hanyu Jidiao de Jiugong Diaomo Juzhen Jiqi Canshu”汉语基调的九宫调模及其参数. *Shijie Hanyu Jiaoxue* 世界汉语教学 4, 146-156.

Wu Weishan 吴为善 (2003): “Hanyu Jielü de Ziran Tezheng” 汉语节律的自然特征. *Shanghai Shifan Daxue Xuebao* 上海师范大学学报（哲学社会科学版）第 32 卷第 2 期 , 100-106.

Wu Zongji 吴宗济 (1997): “Cong Shengdiao he Yuelü de Guanxi Tichu Putonghua Yudiao Chuli de Xinfangfa” 从声调和乐律的关系提出普通话语调处理的新方法. *Qingzhu Zhongguo Shehuikexueyuan Yuyanyanjiusuo Jiansuo 45 Zhounian Xueshu Lunwenji* 庆祝中国社会科学院语言研究所建所 45 周年学术论文集. Beijing, Shangwu Yinshuguan

Xiandai Hanyu Cidian 现代汉语词典 (2002). Beijing, Shangwu Yinshuguan

Xiandai Hanyu Guifan Zidian 现代汉语规范字典 (2011). Beijing, Waiyu Jiaoxue yu Yanjiu Chubanshe.

Xinhua Zidian 新华字典 (1998). Beijing, Shangwu Yinshuguan

Yu Aiqin 余蔼芹 (1986): “Shengdiao Jiaofa de Shangque” 声调教法的商榷. *Diyijie Guoji Hanyu Jiaoxue Taolunhui Lunwen Xuan* 第一届国际汉语教学讨论会论文选. Beijing, Beijing Yuyan Xueyuan Chubanshe, P. 229-235

Yu Wei 余维 (1995): “Ri, Han Yuyin Duibi Fenxi yu Hanyu Yuyin Jiaoxue” 日、汉语音对比分析与汉语语音教学. *Yuyan Jiaoxue yu Yanjiu* 语言教学与研究 4, 123-141.

Yuen RenChao 赵元任 (1983): *Tongzi Fang’an* 通字方案 (*A Project for General Chinese*). Beijing, Shangwu Yinshuguan.

Zhang Baolin 张宝林 (2005): “Yuyin Jiaoxue de Xianzhuang yu Duice” 语音教学的现状与对策. *Yunnan Shifan Daxue Xuebao (Duiwai Hanyu Jiaoxue Lilun Yanjiu)* 云南师范大学学报（对外汉语教学理论研究）第 3 卷第 6 期.

Zhang Hesheng 张和生 (2006): *Hanyu Keyi Zheyang Jiao---Yuyan Yaosu Pian* 汉语可以这样教——语言要素篇. Beijing, Shangwu Yinshuguan.

Zhang Lanlin, (2013): Language differences and communication: Contrastive analysis revisited. P.45. Available at:

http://www.uwo.ca/sogs/WJGR/2005/WJGR2005_v12_p113_Zhang.htm (retrieved June 20 2013)

Zhang Weijia 张维佳 (2001): "Hanyu Pinyin Xide zhong de Hanzi Yin Jieru" 汉语拼音习得中的汉字音介入. *Shanxi Jiaoyu Xuyuan Xuebao* 陕西教育学院学报 4, 55-60.

Zhang Yajun 张亚军 (1990): *Duiwai Hanyu Jiaoxuefa* 对外汉语教学法. Beijing, Xiandai Chubanshe.

Zhang Zhiyun 张志云 (2002): "Zhongji Hanyu Shuiping Hanguoren Shengmu Pianwu Diaocha" 中级汉语水平韩国人声母偏误调查. *Duiwai Hanyu Jiaoxue Yanjiu* 对外汉语教学研究. Shanxi, Shanxi Renmin Chubanshe

Zhao Jinming 赵金铭 (2006): *Duiwai Hanyu Yuyin ji Yuyin Jiaoxue Yanjiu* 对外汉语音及语音及语音教学研究. Beijing, Shangwu Yinshuguan

Zhao Jinming 赵金铭 (1985): "Jianhua Duiwaihanyu Yinxi Jiaoxue de Keneng he Yiju" 简化对外汉语音系教学的可能和依据. *Yuyan Jiaoxue yu Yanjiu* 语言教学与研究 3, 62-75.

Zhao Jinming 赵金铭 (1988): "Cong Yixie Shengdiao Yuyan de Shengdiao shuodao Hanyu Shengdiao" 从一些声调语言的声调硕导汉语声调. *Di'erjie Guoji Hanyu Jiaoxue Taolunhui Lunwen Xuan* 第二届国际汉语教学讨论会论文选. Beijing, Yuyan Daxue Chubanshe.

Zhao Jinming 赵金铭 (2007): *Duiwai Hanyu Jiaoxue Gailun* 对外汉语教学概论. Beijing, Shangwu Yinshuguan.

Zhao Yuanren 赵元任 (1948): *Mandarin Primer, An Intensive Course in Spoken Chinese*. Cambridge, MA, Harvard University Press

Zhou Cuilin 周翠林 (1999): "Hanyu Putonghua de Teshu Yunmu Jiqi Jiaoxue" 汉语普通话的特殊韵母及其教学. *Hanyu Sucheng Jiaoxue Yanjiu (Di'er Ji)* 汉语速成教学研究，第二辑. Beijing, Huayu Jiaoxue Chubanshe

Zhou Fang 周芳 (2006): "Duiwai Hanyu Yuyin Yanjiu yu Yuyin Jiaoxue Yanjiu

Zongshu” 对外汉语语音研究与语音教学研究综述. *Yuannan Shifan Daxue Xuebao* 云南师范大学学报（对外汉语教学与研究版）第 4 卷第 2 期。

Zhu Xiaonong, Zhang Ting and Yi li 朱晓农，章婷，衣莉 (2012): “A classification of dipping tones: Jianlun Beijinghua Shangsheng de Yinjiexue Xingzhi” 凹调的种类——兼论北京话上声的音节学性质. *Zhongguo Yuwen* 中国语文 5, 420-418.

Zhu Xiaonong 朱晓农 (2005): *Shanghai Shengdiao Shiyang Lu* 上海声调实验录 (*An Experimental Study in Shanghai Tones*). Shanghai, Shanghai Jiaoyu Chubanshe.

Zhu Xiaonong 朱晓农 (2010): *Yuyin Xue* 语音学. Beijing, Shangwu Yinshuguan.

Appendix:

Syllable sets 1 & 2, used for collecting recorded data from Hungarian learners and Chinese native speakers.

The test paper of monosyllabic tones 单字调测试卷 1

Bà 爸	Dī 滴	Wū 乌	Dǎ 打	Bì 必	Dū 督	Dà 大	yí 姨	Dú 读	Dī 低
Bì 币	Wǔ 午	Dǐ 底	Bǔ 捕	ā 阿	Dì 第	Dú 独	Bī 逼	Bá 拔	Dù 度
Dā 答	Bǔ 哺	Wǔ 五	Dǐ 抵	Dū 都	Bí 鼻	Yì 易	dì 弟	Wú 吴	Bù 部
Wū 屋	Bà 罢	Dá 达	Yǐ 已	Bù 布	Yì 意	Dǔ 堵	Bǐ 比	Wù 误	Dí 迪
Dǔ 赌	Bǐ 笔	Yí 宜	Wù 物	Dí 笛	Bā 巴	Yī 衣	Dú 毒	Bǎ 把	Dā 搭
Yī 一	Bā 八	Dí 敌	Wú 无	Dù 杜	Yǐ 以	Bū 逋	Yí 咦	Dǎ 沓	Bǔ 补

The test paper of monosyllabic tones 单字调测试卷 2

Bǔ 哺	Dū 都	Wū 乌	Dì 第	Dǎ 沓	Dū 督	Dà 大	yí 姨	Dú 读	Yì 易
Bà 罢	Bǐ 笔	Dǐ 底	Bǔ 捕	ā 阿	Wù 误	Dú 独	Dǐ 抵	Bá 拔	Dù 度
Dā 答	Bà 爸	Wǔ 五	Dǎ 打	Dī 滴	Yī 衣	Dī 低	Bí 鼻	Wú 吴	Bù 部
Wū 屋	Bì 币	Bù 布	Yǐ 已	Dá 达	Yì 意	Dǔ 堵	Bǐ 比	Bī 逼	Bǎ 把
Dǔ 赌	Wǔ 午	Yí 宜	Wù 物	Yī 一	Bā 巴	Bǔ 补	Yí 咦	Dí 迪	Dā 搭
Dì 弟	Bā 八	Dí 敌	Wú 无	Dí 笛	Yǐ 以	Bū 逋	Dú 毒	Bì 必	Dù 杜