

國立交通大學

外國語文學系外國文學與語言學碩士班

碩士論文

送氣與華語聲調

Prevocalic Aspiration and Mandarin Tones



研究生：陳米琪

指導教授：賴郁雯 教授

中華民國一百年六月

送氣與華語聲調

Prevocalic Aspiration and Mandarin Tones

研究生：陳米琪

Student：Mi-Chi Chen

指導教授：賴郁雯

Advisor：Yuwen Lai

國立交通大學

外國語文學系外國文學與語言學碩士班



Submitted to Department of Foreign Languages and Literatures
College of Humanities and Social Sciences
National Chiao Tung University
in partial Fulfillment of the Requirements
for the Degree of
Master
in
Foreign Languages and Literatures

June 2011

Hsinchu, Taiwan, Republic of China

中華民國一〇一年六月

送氣與華語聲調

學生：陳米琪

指導教授：賴郁雯

國立交通大學

外國語文學系外國文學與語言學碩士班

摘 要

華語的送氣分調議題曾被討論過，然而過去研究中發現，送氣聲母對基頻所造成的結果並不一致、甚至相互矛盾。因此本文不僅針對相互矛盾的研究結果提出更縝密的實驗方法，也對相關問題略做紓解。

實驗中使用了四十組最小辨意詞組，這四十組最小辨意詞組由三個不同發音部位所組成，搭配上華語的四個聲值，並且測量其母音基頻。測量基頻方式分成三種，分別是：正規化韻母基頻(normalized F0)、韻母起始點基頻(onset F0)、前一百毫秒韻母基頻(first 100 ms F0)。雖然有三種不同基頻萃取方式，但本文著重於前一百毫秒韻母基頻，而其相關結論亦在文中探討。

影響基頻變化的變量包括送氣與否(送氣聲母和非送氣聲母)、性別差異(男性與女性)、語速(語速快和語速慢)。實驗結果顯示：無論在送氣與否、性別及語速快慢上都有顯著性差異。基頻在送氣聲母後較高，女性的基頻也顯著於高於男性基頻，但基頻高低在語速快慢上則無一致表現。有趣的是：送氣與性別及與語速之間的互動也有顯著性差異。

關鍵字：送氣，基頻，性別，語速

Prevocalic Aspiration and Mandarin Tones

Student: Mi-chi Chen

Advisor: Dr. Yuwen Lai

National Chiao Tung University

Department of Foreign Languages and Literatures
National Chiao Tung University

ABSTRACT

The perturbation effect of prevocalic aspiration on fundamental frequency (F0) was investigated in Mandarin. Previous researches have shown conflicting results regarding the effect of aspiration on F0. The present study aims to develop a rigorous experimental design to test possible modulating factors which gave rise to the controversy. Forty minimal pairs contrasting in prevocalic aspiration across three places of articulation from 4 Mandarin tones were recorded in a carrier sentence. F0 of the following vowel was measured at the onset F0, normalized F0, and the first 100 ms F0 of the tonal contour. The results put more emphasis on the first 100 ms F0, and the relative discussion is illustrated in the discussion section. Effect of independent variables including Aspiration (aspirated and unaspirated prevocalic stops), Gender (male and female), and Speaking Rate (fast and slow) on F0 were evaluated. Results from 20 native Mandarin speakers (10 female, 10 male) reveal significant main effects of Aspiration (higher F0 after aspirated stops), Gender (higher F0 in female), and Speaking rate (higher F0 in fast and slow mode). More interestingly, interactions of Aspiration and Gender as well as Aspiration and Speaking Rate are also significant. The effect of Aspiration and its interaction with intrinsic F0 of speakers and speaking rate are discussed as well.

Keywords: Aspiration, Fundamental frequency (F0), gender, speaking rate

致 謝

終於，這份論文完成了。

如果把論文產生的過程比喻成母親懷胎，真是再貼切也不過了。一開始的論文架構、內容、實驗設計，就像是寶寶的骨肉長成；而每次 meeting 也像是戰戰兢兢的產檢，深怕一個不注意，論文(or 寶寶)就不能安穩茁壯。

現在回頭看研究生生活，像是一眨眼的事情。而這一眨眼裡，有著滿滿充實和獲得。首先要感謝我論文的主治大夫，喔！其實是我的指導老師賴郁雯教授，不僅在論文中提供許多指導與協助，更在生活上供應許多歡笑來源。並且在研究所最後一年裡，和老師出國參加會議，這趟旅程將我的眼界變得更寬廣。感謝老師這兩年來的包容和諄諄教誨，無論是學術上指導、或者人生經驗。感謝兩位口試委員：歐淑珍老師、何延光老師，謝謝他們對論文的建議與提醒。也感謝研究所修課期間的諸位老師的指導，謝謝你們的付出，讓我在學術路上成長著。

俗云：在家靠父母，出外靠朋友。感謝研究所的同儕以及學長姊們，如果研究生涯少了你們，肯定少了很多回憶。還有梅竹團契的大家，感謝你們情義相挺幫我錄音、還要忍受我每個禮拜在代禱事項發牢騷。每一次聚會、午聚、看感人片到最後都會變搞笑片的電影時間、以及出遊，都為我研究生活增添幾筆色彩。

也要感謝我最親愛的家人，我的父親陳海清先生、母親王芳蓮女士，謝謝他們總是讓我每次回家都當個茶來伸手、飯來張口的大小姐，雖然我爸每次都會嘮叨：「還要念多久，等你畢業都沒人要娶你」的玩笑話，但我知道他們深深以我這個女兒為豪。還有我兩個妹妹，謝謝你們總是無厘頭的搞笑。

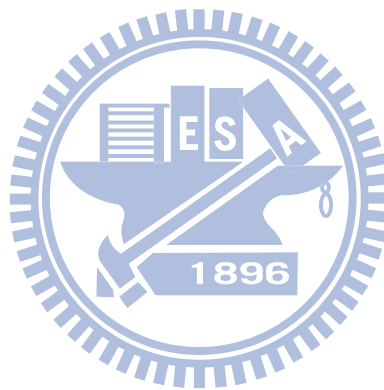
想像過很多次，當我論文完成，寫致謝時是什麼樣的感受。而此時此刻的我，正在細細體會這樣的心情。也許多年後我會記不起論文裡頭的細節，但論文完成那瞬間的興奮，我想我會永遠記得的。最後，我要感謝神，謝謝祂把我帶到新竹來，謝謝祂讓我在這三年所經歷、學習的一切。

Content

摘 要.....	i
ABSTRACT.....	ii
致 謝.....	iii
Content.....	iv
Figure.....	vii
Table.....	x
Chapter I Introduction.....	1
Chapter II Literature Review.....	4
2.1 Segmental interaction.....	4
2.1.1 Consonants affect vowels.....	4
2.1.2 Consonants affect Tones.....	4
Voicing and F0.....	4
Aspiration and F0.....	7
Non-tonal languages.....	8
Tonal Languages.....	9
Mandarin.....	10
2.2 Determinants of F0.....	12
2.2.1 Physiological Factors of F0 control.....	12
The anatomy of larynx.....	12
Vocal Folds.....	14
Laryngeal Mechanisms.....	16
Cricothyroid Muscle (CT).....	16
Other intrinsic muscles.....	17
Extra-laryngeal Mechanisms.....	18
Laryngeal Height.....	18
Extrinsic muscles.....	19
2.2.2 Aerodynamic factors.....	20
Air Pressure.....	20
Airflow.....	21
2.2.3 Speaking rate.....	23

2.3 Introduction of Mandarin	24
2.3.1 Initials	24
2.3.2 Finals.....	25
2.3.3 Mandarin tones	25
Chapter III Methodology.....	28
3.1 Subject.....	28
The first stage participant	28
The second stage participant.....	29
3.2 Stimuli	29
3.3 Instrument.....	30
3.4 Procedure.....	30
3.5 Measurements.....	31
Voice Onset Time (VOT).....	31
Fundamental Frequency (F0).....	31
3.6 Statistical analysis	34
Chapter IV Results	35
4.1 Voice Onset Time (VOT).....	35
4.2 Fundamental Frequency (F0) Measurement.....	39
Tone 1	39
Tone 2.....	43
Tone 3.....	47
Tone 4.....	50
Chapter V Discussion.....	55
5.1 Summary of results.....	55
5.2 VOT.....	56
5.3 F0.....	57
Aspiration.....	58
Aspiration and Rate.....	60
Aspiration and Gender	61

5.4 Further research	64
5.5 Conclusion	66
References	67
Appendix	73
1. Stimuli in the methodology	73
2. The overall results of measurements of F0	75
3. Results of a 3-way ANOVA test on VOT	76
4. Results of a 3-way ANOVA test on Normalized F0	79
5. Results of a 3-way ANOVA test on Onset F0	82



Figure

Figure 1. Anterior view of larynx	13
(adopted from http://en.wikipedia.org/wiki/File:Larynx_external_en.svg)	13
Figure 2. The larynx seen from the back and right side (adopted from Lai (2009))	14
Figure 3. The intrinsic muscles of larynx (adopted from Honda (2004)).....	17
Figure 4. The extrinsic muscles of larynx (adopted from http://www.rci.rutgers.edu/~uzwiak/AnatPhys/APFallLect14.html).....	20
Figure 5. Oral airflow and air pressure from Seoul and Cheju speakers (adopted from Cho et al., 2002, p210)).....	23
Figure 6. Mean F0 contour of four Mandarin tones in the monosyllable /ma/ (adopted from Xu, 1997, p67).....	27
Figure 7. The sample <i>kha</i> was pronounced by a male speaker in Tone 1 shown in Praat screen. The upper panel displays the waveform while the lower panel shows the spectrogram of VOT (128 ms) the tonal contour (289 ms). The x-axis is time in seconds and the y-axis is frequency in Hz.	32
Figure 8. This picture shows the normalized F0 in the rising tonal contour.	33
Figure 9. This picture shows the every 10 ms F0 in the rising tonal contour.	34
Figure 10. The main effect of Aspiration on VOT.....	36
Figure 11. The main effect of Rate on VOT.....	36
Figure 12. The main effect of Gender on VOT	37
Figure 13. The interaction between Gender and Rate on VOT	37
Figure 14. The interaction between Aspiration and Rate on VOT	38
Figure 15. The main effect of Aspiration on 100 ms F0 in Tone1	40
Figure 17. The main effect of Rate on 100 ms F0 in Tone 1	41
Figure 18. The interaction between Aspiration and Gender on 100 ms F0 in Tone 1	41
Figure 19. The interaction between Aspiration and Rate on 100 ms F0 in Tone 1 ..	42
Figure 20. The interaction between Gender and Rate on 100 ms F0 in Tone 1.....	42
Figure 21. The main effect of Aspiration on 100 ms F0 in Tone 2.....	44
Figure 22. The main effect of Gender on 100 ms F0 in Tone 2.....	44
Figure 23. The main effect of Rate on 100 ms F0 in Tone 2	45
Figure 24. The interaction between Aspiration and Gender on 100 ms F0 in Tone 2	45
Figure 25. The interaction between Aspiration and Rate on 100 ms F0 in Tone 2..	46
Figure 26. The interaction between Gender and Rate on 100 ms F0 in Tone 2.....	46
Figure 27. The main effect of Aspiration on 100 ms F0 in Tone 3.....	48
Figure 28. The main effect of Gender on 100 ms F0 in Tone 3.....	48

Figure 29. The main effect of Rate on 100 ms F0 in Tone 3	49
Figure 30. The interaction between Aspiration and Gender on 100 ms F0 in Tone 3	49
Figure 31. The interaction between Gender and Rate on 100 ms F0 in Tone 3.....	50
Figure 32. The main effect of Aspiration on 100 ms F0 in Tone 4.....	51
Figure 33. The main effect of Gender on 100 ms F0 in Tone 4.....	52
Figure 34. The main effect of Rate on 100 ms F0 in Tone 4	52
Figure 35. The interaction between Aspiration and Gender on 100 ms F0 in Tone 4	53
Figure 36. The interaction between Aspiration and Rate on 100 ms F0 in Tone 4..	53
Figure 37. The interaction between Gender and Rate on 100 ms F0 in Tone 4.....	54
Figure 38. Male-female comparisons of dimensions of the larynx (a) Sagittal view of the thyroid cartilage and (b) horizontal section showing difference in membranous length (Adopted by Kahane, 19198)	62
Figure 39. First 100 F0 ms of aspirated and unaspirated stops between male and female (high tones).....	65
Figure 40. First 100 ms F0 of aspirated and unaspirated stops between male and female (low tones).....	65
Figure 41. The main effect of Aspiration on VOT across four tones.....	76
Figure 42. The main effect of Gender on VOT across four tones	76
Figure 43. The main effect of Rate on VOT across four tones	77
Figure 44. The interaction between Aspiration and Gender on VOT across four tones	77
Figure 45. The interaction between Aspiration and Rate on VOT across four tones	78
Figure 46. The interaction between Gender and Rate on VOT across four tones ...	78
Figure 47. The main effect of Aspiration on normalized F0 across four tones	79
Figure 48. The main effect of Gender on normalized F0 across four tones	79
Figure 49. The main effect of Rate on normalized F0 across four tones.....	80
Figure 50. The interaction between Aspiration and Gender on normalized F0 across four tones.....	80
Figure 51. The interaction between Aspiration and Rate on normalized F0 across four tones.....	81
Figure 52. The interaction between Gender and Rate on normalized F0 across four tones	81
Figure 53. The main effect of Aspiration on onset F0 across four tones	lxxxii
Figure 54. The main effect of Gender on onset F0 across four tones	lxxxii
Figure 55. The main effect of Rate on onset F0 across four tones	83

Figure 56. The interaction between Aspiration and Gender on onset F0 across four tones 83

Figure 57. The interaction between Aspiration and Rate on onset F0 across four tones 84

Figure 58. The interaction between Gender and Rate on onset F0 across four tones 84



Table

Table 1. Mandarin initial consonants	24
Table 2. Mandarin finals	25
Table 3. Mandarin character differ in tones	25
Table 4. Result of a 3-way ANOVA test on VOT	35
Table 5. Result of an ANOVA test on first 100 ms F0 (Tone 1)	39
Table 6. Result of an ANOVA test on first 100 ms F0 (Tone 2)	43
Table 7. Result of an ANOVA test on first 100 ms F0 (Tone 3)	47
Table 8. Result of an ANOVA test on first 100 ms F0 (Tone 4)	50
Table 9. Overall results of normalized F0 measurement by factors (aspiration, rate, gender) and their interaction	75
Table 10. Overall results of first 100 ms F0 measurement by factors (aspiration, rate, gender) and their interaction	75
Table 11. Overall results of Onset F0 measurement by factors (aspiration, rate, gender) and their interaction	75



Chapter I Introduction

Adjacent speech sounds are known to affect each other, which is known as coarticulation. The effect could be anticipatory (sounds affected by the following sounds) or carry-over (sounds affected by the preceding sounds) (Gandour et al., 1992). These effects have been found in several previous studies. For example, in English, vowels are generally nasalized when followed by a nasal segment (Ladefoged and Johnson, 1975). Homorganic nasal assimilation is the assimilation which assimilates a nasal consonant to the feature of place of articulation with the conditioning sound (Ohala, 1990). In other terms is that the nasal assimilated to the next consonant in place, becoming homorganic with the following consonant, such as in • possible → im • possible.

Specific phonetic manifestation has also been found in several studies. The vowels duration, for instance, is longer when followed by voiced consonants than followed by voiceless consonants (Ladefoged and Johnson, 1975). Another example is that the voiceless stops /p, t, k/ are unaspirated after /s/ in words such as *stew*, *skew* (Ladefoged and Johnson, 1975).

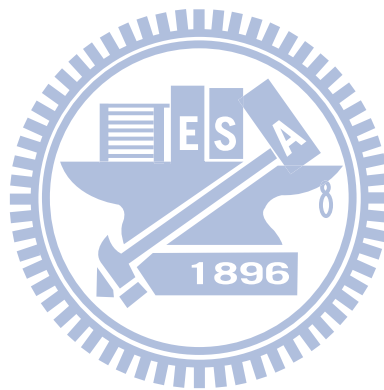
In addition to the interaction between segments, segments and supra segmentals were also found to affect each other. The most well-known effect being the voicing of the prevocalic consonants affect the fundamental frequency (F0) of the following vowel. It has been shown that F0 is lower when followed by voiced stops than followed by voiceless stops. This perturbation has been found in both tonal and in non-tonal languages (Ohde, 1984; Hombert & Ladefoged, 1977; Whalen, Abramson, Lisker, & Mody, 1990; Lehiste and Pererson, 1961; Mohr, 1971; Fromkin, 1978).

In the middle age of Chinese, the tones are departed because the prevocalic voicing. This gives rise to a famous theory--tonogenesis. More description will be illustrated in the following section. The effect of voicing is well established, but the effect of aspiration less clear. The present study aims to investigate the effect of aspiration on Mandarin tones and new factors investigated are speaking rates and gender.

In the present study, how aspiration affects F0 of the following F0 is concerned. Furthermore, the determinant of the source of the effect interests us as well. Presumably the aerodynamic condition related to aspirated stops and unaspirated stops are different, and may result in different F0 values of the vowels. The respiratory system generates a constant subglottal pressure during closure for all stops. At the release of an aspirated stop, high rate of airflow runs through the glottis and pressure decrease during the aspirated pronouncing (Ohala and Ohala, 1972). Isshiki (1964) noted that F0 decreased when the subglottal pressure increased. Aspirated stops thus should give rise to higher F0 of the following vowel than unaspirated stops. Besides the aerodynamic factor of aspiration itself, Dromey and Ramig (1998) indicated that the different speaking rate resulted in different F0 performance. Inspired by the result of Dromey and Ramig, speech rate is one of concerns in the present study.

Lai (2004) indicated that the significant rising effect of aspirated stops only was found only in female. Compared with the result of Xu and Xu (2003), there were only seven female subjects in the study and F0 was lower after aspirated stops. According to gender differences from these two opposite results, gender will be one of the factors in the present study as well.

The aim of this paper is to provide a more consistent measurement and to help clarify the effect of consonant aspiration on the F0 of following vowels in Mandarin. Stimuli adopted in the experiment are composed of three places of articulation, four vowels, and four tones. Apart from these, Gender and Rate are also included in order to test their effect on the perturbation phenomenon. Based on the results, further discussion on our concerns will be addressed.



Chapter II Literature Review

2.1 Segmental interaction

2.1.1 Consonants affect vowels

Speech segments are known to affect each other. The interaction between consonants and vowels has been investigated widely. Many papers have reported the segmental interaction between consonants and vowels (Peterson & Lehiste, 1960; House, 1961; Umeda, 1975). Peterson & Lehiste, (1960) pointed out that all syllable nuclei in English are significantly affected by the nature of the consonants that follow the syllable nuclei. For example, the syllable nucleus is shorter when it is followed by a voiceless consonant, and longer when followed by a voiced consonant.

Another statement where consonants affect vowels is during vowel nasalization. Vowels tend to become nasalized before nasal consonants ($VN \rightarrow \tilde{V}$) (Ladefoged and Johnson, 1975). The nasalized vowels are produced primarily by lowering the velum, resulting in opening a side passage for air flow through the nasal cavity. In addition, the cross-language phonological evidence indicates that vowels are more likely to be affected by nasal assimilation when followed by a syllable-final nasal consonant than when preceded by a nasal consonant (Krakow, 1999; Clumeck, 1976).

2.1.2 Consonants affect Tones

Voicing and F0

In addition to the interaction between segments, segment and supra segmental were also found to affect each other. The effect of the prevocalic consonants, particularly voicing, on F0 of the following vowel has been discussed in many studies. It is well-known that F0 is higher when followed by voiceless stops than by voiced stops.

Previous studies show a common consensus when the F0 of following vowels is lower after voiced consonants in both tonal and non-tonal languages (tonal: Matisoff (1971), Gandour, (1975); nontonal: Ohde (1984), Hombert & Ladefoged (1977), Whalen, Abramson, Lisker, & Mody (1990), Lehiste and Pererson (1961) Mohr (1971), Fromkin (1978)).

House and Fairbanks (1953) investigated the variation of vowels varying in different consonantal environments in English. The general plan was to place vowels in various consonant environments (CV syllables which C=/p, t, k, f, s, b, d, g, v, z, m, n/ and V=/i, e, a, o, æ, u/). The CV syllables were produced by 10 male subjects. The duration and F0 of vowel were measured. The duration measurement showed that the vowel duration was longer in the voiced environments. Furthermore, results revealed that F0 was higher after voiceless stops than after voiced stops.

Fromkin (1978) conducted a similar experiment on how voicing affects F0 in American English. Five subjects are asked to produce six CV nonsense word (C = /p, t, k, b, d, g/, and V = /i/) in the frame “Say __ again”. With reference point at the onset of the vowel, F0 was measured at onset and at 20, 40, 60, 80, and 100 ms after the onset. The result showed that F0 of vowels after the voiced stops is lower than after voiceless stops. Fromkin explained the phenomenon in the following terms. “After the closure of voiced consonant, voicing continues, but since the oral pressure increases, the pressure drop decreases, leading to a lower frequency. The F0 then rises after the release until it reaches the ‘normal’ value of the vowel which is being realized.”

In addition to make the different performance on duration and F0, voicing also causes another phonetic phenomenon in Mandarin—tonal split. During middle age of

Chinese, people found that the tone is lower when producing voiced stops, and the tone is higher when producing voiceless stops. Thus, the tone is departed after voiced obstruents, and the differential way in which voiced stop and affricates devoiced in Mandarin. For instance, each of the Middle Chinese tonal categories splits to high and low registers—known as *ying* and *yang* in Chinese phonology—yielding a perfectly symmetrical eight-tone system. In each case, the *yang* register with a voiced onset has a lower pitch values than the corresponding *yin* register with higher pitch values (Chen, 2000). This gives rise to a theory—tonogenesis—in which development of contrastive tones are due to the loss of voicing distinction in prevocalic obstruents (Pulleyblank, 1986).

The rising effect of voiceless stops may be explained by physiological mechanisms—the vocal folds tension. In making the voiced vs. voiceless distinction on stops, vocal folds tension is changed so as to affect the F0 of adjacent vowels (Hombert et al., 1979; Fromkin, 1978). Halle and Steven (1971) suggested that these intrinsic variations are the result of horizontal vocal folds tension: the vocal folds are presumably slack in order to facilitate voicing during voiced stops and stiff in order to inhibit voicing during voiceless stops. These vocal folds states spread to adjacent vowels, affecting their F0. Another variant of the vocal folds tension hypothesis is that which suggests that it is the vertical tension of vocal folds which is affected by the voiced vs. voiceless distinction (Ohala, 1973, Ewan 1979, Steven 1975).

Furthermore, the aerodynamic hypothesis may explain the rising effect of voiceless stops. When producing a voiced stop, oral pressure gradually builds up, this decreases the pressure drop across the vocal cords—which in turn decreases the F0. Upon the release of the stop, the pressure drop returns to normal, producing an initially low and

rising F0 contour after voiced stops (Ladefoge, 1967). In the case of voiceless stops, the airflow past the vocal cords is very high upon release, creating a high-than-normal Bernoulli force—which will draw the vocal cords together more rapidly, and so increase the rate of their vibration at vowel onset. As the airflow returns to normal, the F0 will too. Thus, after voiceless stops, the F0 contour will be initially high and falling (Ohala, 1970; Ohala and Ewan, 1973; Abramson 1974; Hombert and Ladefoged, 1977).

Aspiration and F0

Although the effect of voicing on F0 is widely regarded, research about the effect of aspiration on F0 has received less scholarly attention. The possible effect of aspiration on F0 is particularly interesting when it induces possible phonetic contrast. A series of documentations regarding aspirations as a cause of tone splitting have been published.

Tonal split caused by prevocalic aspiration was documented in some Chinese languages, such as Wu, Gan, Xiang, and Miao (Ho, 1990; Shi, 1998). Although no instrumental research was conducted, both Ho and Shi argued that F0 is lower after aspirated stops due to the lowering of the larynx when producing aspirated stops.

There is no conducting regarding the effect of aspiration on F0. Three kinds of results regarding this perturbation effect have been provided. One is that F0 is lower after aspirated stops, data was found in Korean (Kagaya, 1974), Cantonese (Francis et al., 2006) and Mandarin (Xu & Xu, 2003). Another is that there is no difference in F0 after aspirated or unaspirated stops, for instance, Hombert and Ladfoged (1977). Finally, F0 is higher after aspirated stops. Examples can be found in Cantonese (Zee, 1980), Korean (Kenstowicz & Park, 2006), and Taiwanese (Lai, 2004).

Form above, there exists no agreement and different results within the same languages were found (in Korean Kenstowicz & Park (2006) vs. Kagaya (1974)). Related researches conducted on non-tonal and tonal languages will be reviewed respectively in the following section; in addition to studies on aspiration and F0 in Mandarin as well.

Non-tonal languages

$C^h < C$

In non-tonal languages, Kagaya (1974) studied the laryngeal gestures of three types of consonants in Korean. Two native speakers of Seoul dialect were recorded producing /CV/ and /VCV/ in isolation. F0 was measured for each sample, by averaging values for the first three fundamental periods from voice onset. The results showed that F0 at voice onset of the aspirated type is lower than ones of the unaspirated stops.

$C^h = C$

Hombert and Ladefoged (1977) investigated the two series of voiceless stops in English and French (the English series is voiceless aspirated as opposed to the French series which is supposed to be voiceless unaspirated). Two American English speakers (1male, 1female) and 2 French speakers (1male, 1female) were asked to produce 6 CV nonsense words (consonants = /p, t, k, b, d, g/ and vowel = /i/) in the frame—“Say ___ louder”. Results indicated that these two series of voiceless consonants (English voiceless aspirated stops and French voiceless unaspirated stops) had very similar effect on the F0 contours of the following vowels.

$C^h > C$

Kenstowicz and Park (2006) investigated the three-way contrasts in Korean Kyungsang dialect and how F0 was utilized to implement the tonal and laryngeal contrast in Kyungsang. Seven speakers (2 males, 5 females) were recorded producing 48 words in a sentential frame. A number of measurements were taken including VOT and F0 at four points (the onset, the mid-point of the first and of the second vowels). Results showed that VOT of aspirated stops was the longest and that F0 of vowels following tense and aspirated consonants have higher F0 than those following a lax consonant. The data indicated that the laryngeal category of the onset consonant has a systematic effect on the F0 value of the following vowel that is highly significant at both the onset and middle of the vowel.

There is no consistent agreement in the non-tonal languages from above investigations. The few studies on non-tonal languages suggest that aspiration has both rising effect (Kenstowicz & Park, 2006) and lowering effect (Kagaya, 1974), or the similar effect on aspiration (Hombert & Ladefoged, 1977).

Tonal Languages

C^h < C

Tonal languages such as Thai (Gandour, 1974), and Cantonese (Zee, 1980; Francis, 2006), for many years have been investigated to uncover the relationship between aspiration and the following vowel for many years. Gandour (1974) investigated the effect of preceding consonants on tone in Thai. A male speaker was asked producing CV₁V₂ syllable where C=/p, p^h, b, t, t^h, d, s, n/, V₁=V₂= /a, i, u/ with five tones and F0 was measured. Gandour found that the initial F0 value after the release of an unaspirated stop is higher for a voiceless aspirated stop.

$C^h > C$

Zee (1980) studied the difference between the effect of /p^h/ and /p/ on the F0 onset of the following diphthong /ei/ in Cantonese. Three male participants were asked to read the two Cantonese words (/p^hei/, /pei/) in a sentence frame at a normal rate of speech. A F0 measurement for each test word was obtained every 10ms. The result showed that for all the tokens the F0 onsets associated with aspirated stops were higher than those associated with unaspirated for all three speakers.

Francis et al. (2006) investigated the effect of aspiration differences in Cantonese initial stops on the F0 of the following vowels, as well as the interaction of this effect with tone contour. 16 native speakers of Cantonese (8 males, 8 females) were asked to produce CV syllables with six tones. F0 of vowels after aspirated and unaspirated stops was measured over the first 100 ms. Results showed that the onset F0 is higher after unaspirated stops than after aspirated ones. In addition, there is a falling F0 contour over the first 100 ms, consistent with the voiceless status of both aspirated and unaspirated stops.

Mandarin

Tonal splits are triggered by aspiration in some Chinese dialects such as Wu and Tanyang (Chao, 1967). Chao (1967) pointed out that the most characteristic feature of Wu dialects is the tripartite division of initial stops into voiceless unaspirated, voiceless aspirated and voiced aspirated. However, not only Chao mentioned the phenomenon of tonal splitting by aspiration, Shi (2007) remarked as well on it in the northern Wu dialect. Shi (2007) indicated that the characteristic of Wu dialect is that aspirated tones are realized as lower register tones.

Besides the voicing parameter, aspiration is another important aspect in differentiating the relationship between F0 and the prevocalic consonants. More investigations were conducted by Xu and Xu (2003) for Mandarin. Xu and Xu (2003) investigated the effect of consonant aspiration on the following vowels. Seven females native speaker were asked to read the stimuli (/ma/, /ta/, /t^ha/, and /ʂa/ with four tones) in two carrier sentences—*wo3 lai2 shuo1 ____ zhe4 ge4 ci2* ('I say the word ____') and *wo3 lai2 zhao3 ____ zhe4 ge4 ci2* ('I look for the word ____'). F0 of these targets words were measured by an automatic vocal detection and manual rectification. The results indicated that the onset F0 is higher following unaspirated consonants than following aspirated consonants. Xu and Xu (2003) indicated that during the closure of the stops, pressure builds up to a constant level irrespective to the aspiration feature of the consonants. At the release of /t^h/, pressure decreases markedly and at the release of /t/, however, pressure remains at a high level and gradually returns to normal. Pressure should be lower at the voice onset for /t^h/ than for /t/. These differences should lead to lower onset F0 in /t^h/ than in /t/.

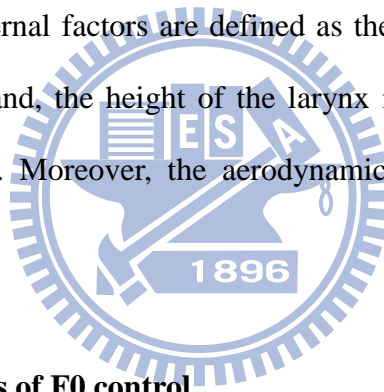
In addition, Lai (2004) studied whether and how aspiration influences the tones in Taiwanese. Four participants (2 males, and 2 females) were asked to produce 56 CV(O) syllables, consisting of the stops (/p/, /p^h/, /t/, /t^h/, /k/, /k^h/) and alveolar affricates (/ts/, /ts^h/) followed by a vowel (/i/, /ɛ/, /a/, /u/, and /o/) with seven tones. VOT, and F0 (onset F0, end of F0, and mean F0) were measured. Results showed that onset F0 and mean F0 are significantly higher after aspiration stops than after unaspirated ones. F0 after aspirated stops is higher than after voiceless unaspirated stops due to the faster airflow rate and higher larynx position. Further analysis in Lai (2004) showed that this rising effect was only significant in females and not in males. Compared with Lai (2004) and Xu and Xu (2003), the aspiration rising effect is

opposite in their result. Inspired by the gender difference, the present study recruited more subjects and greater efforts were made to balance the gender of the participants.

The results of tonal languages are similar to that of non-tonal languages. No agreement can be reached with respect to the effect of the aspirated stops on F0. The disagreement certainly requires further research on this aspect.

2.2 Determinants of F0

Next, the interaction between segment and F0 is investigated, the physiological factor which affects F0 are reviewed in this section. Two aspects of laryngeal mechanisms are discussed here: the internal factors are defined as the coordination of muscles in the larynx; on the other hand, the height of the larynx is defined as external factor (Hirano and Ohala, 1969). Moreover, the aerodynamic and speaking rate will be reviewed as well.



2.2.1 Physiological Factors of F0 control

The anatomy of larynx

A basic function of the larynx and the mechanics of voice production are necessary to be understood before discussing the production of F0 can be discussed. The larynx plays the role of breathing, speaking and swallowing in the human body. The epiglottis is closed to ensure that food will pass through the pharyngeal cavity into the esophagus. In speech, the larynx is important as an articulator and a source of sound.

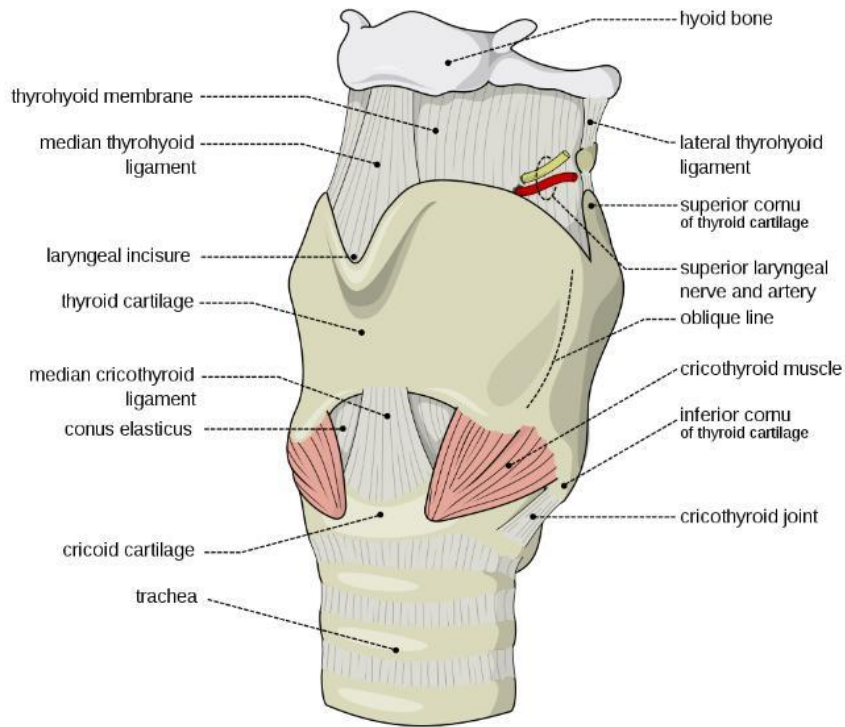


Figure 1. Anterior view of larynx

(adopted from http://en.wikipedia.org/wiki/File:Larynx_external_en.svg)

The larynx (Fig. 1) has a skeletal frame formed by a series of cartilages. There are two main cartilages—the upper thyroid cartilage and the lower and smaller cricoids cartilage. The epiglottis lies superiorly; it protects the larynx during swallowing and prevents the inspiration of food.

The most prominent laryngeal cartilage is called the thyroid cartilage. It consists of two plates which are arranged in a wedge-like shape. The hyoid bone is found above the thyroid cartilage. It is connected to the larynx by the thyrohyoid membrane. The U-shaped hyoid bone serves as an attachment point for the tongue muscles. Beneath the thyroid cartilage is the cricoids cartilage, which forms the base of the larynx. The anterior part of the cricoids cartilage is narrow and referred to as the arch. The

posterior part, which is called lamina, is much broader and forms much of the larynx's back wall. The cricoids cartilage supports the thyroid cartilage and the arytenoids (see Fig. 2). Its upper edge from four articulatory surfaces: two at the side for the thyroid and two at the back for the arytenoids. Above to the lamina are the arytenoids cartilages, which attach to the vocal folds. A pair of triangle-shaped arytenoids cartilages is located along the upper edge of cricoids lamina. On the top of each arytenoid cartilage is a small corniculate cartilage. Each arytenoid cartilage attaches itself to the posterior end of a vocal ligament (Marchal, 2009).

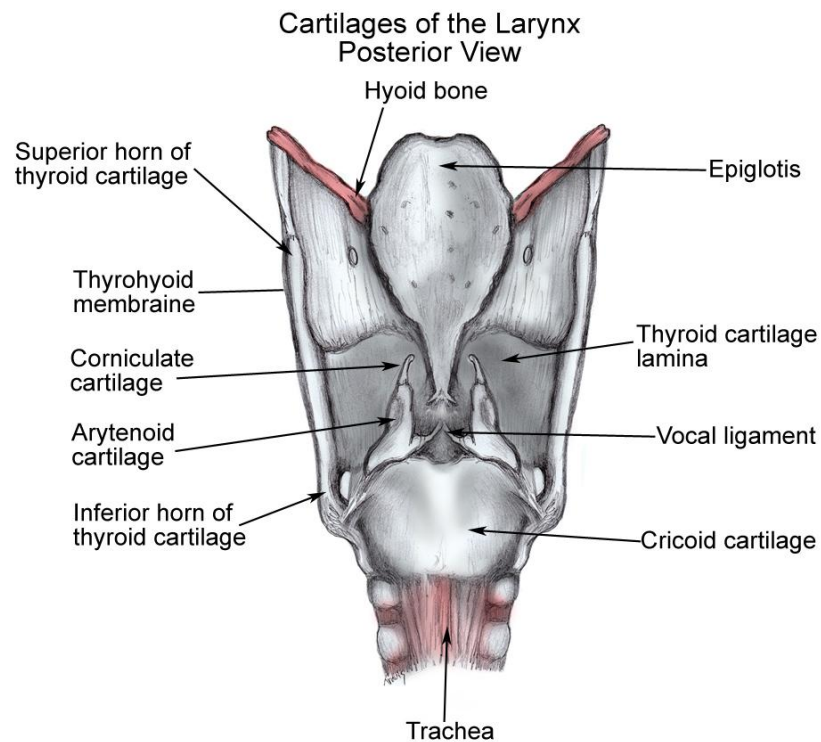


Figure 2. The larynx seen from the back and right side (adopted from Lai (2009))

Vocal Folds

F0 control at the larynx is considered to be achieved by adjusting the effective length and the stiffness of the vocal folds (Hirose, 1997). The studies of vocal folds have been investigated for many years (Hirano, 1981; Hirano, 1983), with the objective of

understanding the behavior of the folds during speech. The vocal folds are twin infoldings of membranes and muscular fibers stretched horizontally across the larynx. They are located below the epiglottis. They are attached at the back to the processes of the arytenoids cartilage and at the front to the thyroid cartilage. Above the vocal folds is a similar structure known as the false vocal cords. There are no significant contributions of the false vocal cords to the normal vocal folds. The vocal folds and the space between them are referred to as the glottis. The glottis expands into a triangular-shape opening while breathing. This allows oxygen to enter the trachea and lungs. When people hold their breath, for example, the vocal folds are closing. When humans breathe, the vocal folds are opening and vibrating as air passes through the larynx, which also occurs when people speak or sing (known as phonation). To make sound, the laryngeal muscles adduct the size of the opening to a narrow slit.

A first factor influencing the rate of vocal fold vibration is the length of vocal folds. Long vocal folds oscillate at a slower rate than shorter ones. The length of the vocal folds is 18-24 mm in a man and 14-19 mm in a woman. Since the vocal folds are longer for men than for women, male voices are usually lower than female voices. A second factor determining the rate of vocal fold vibration is the mass, which is also linked to their thickness (Reetz and Jongman, 2009). Thick vocal folds oscillate at a lower rate than thin vocal folds. In addition, the rate of vocal fold vibration depends on the elastic tension: tense folds vibrate faster than slack folds, because they are pulled back to the rest position with more force. These three effects relate the vibration of the muscular vocal folds.

Except for the factors mentioned above, the vocal folds are different at different ages. Hirano et al. (1989) previously described several structural changes in the vocal folds

tissues associated with aging. Some of these included: a shortening of the membranous vocal folds in males, a thickening of the vocal folds mucosa and cover in female, and edema development in the superficial layer of the lamina propria in both sexes. Taking both gender and age into consideration, these intrinsic factors cause a different performance in F0.

Laryngeal Mechanisms

Cricothyroid Muscle (CT)

The framework of the larynx consists of four different cartilages: the epiglottis, thyroid, cricoids, and arytenoids cartilages. The cricoids and thyroid cartilages are connected by the cricothyroid joint, while the cricoids and arytenoids cartilages are connected by cricoarytenoid joint (Fig. 3). The movement of the cricothyroid joint changes the length of vocal folds. The movements of arytenoids cartilage contribute to the abduction and adduction of the vocal folds (Hirose_1997).

Movement of cricothyroid and cricoarytenoid are controlled by the intrinsic muscles. Elongation and stretching of the vocal folds are achieved by the contraction of CT. CT is the important intrinsic muscle for the operation of the cricothyroid articulation which affects F0. Continuous contraction of CT produces an increase of F0; conversely, relaxation of CT lowers F0 (Gay et al. 1972). CT narrows the angle between the cricoid and thyroid cartilages, increasing tension on the vocal folds. Vocal folds are attached at the back of the arytenoids cartilage and at the front to the thyroid cartilage. Once the CT contracted, thyroid cartilage moves forwards and thus vocal folds are lengthened or stretched by their anterior-posterior plane.

Increased CT activity tenses the vocal folds, which helps to suppress voicing in

voiceless consonants. This behavior in term leads to a higher F0 on the onset of phonations for the adjacent following vowels (Hoole, 2004; Honda 2004; Whalen et al., 1999). Higher CT activity suggests a raising pattern of F0 (Vilkman, Aatonen, Laine & Raimo, 1989). Researches (Whalen et al., 1999; Vilkman, Aatonen, Laine & Raimo, 1989) indicate that CT has a higher correlation with the rising and falling of F0, unlike the findings in Honda’s (1983) research. Honda (1983) found a paradoxical CT activity in the lower F0 region of a speaker’s range (in this case, at the end of sentence); the increases in CT were correlated with decreases of F0.

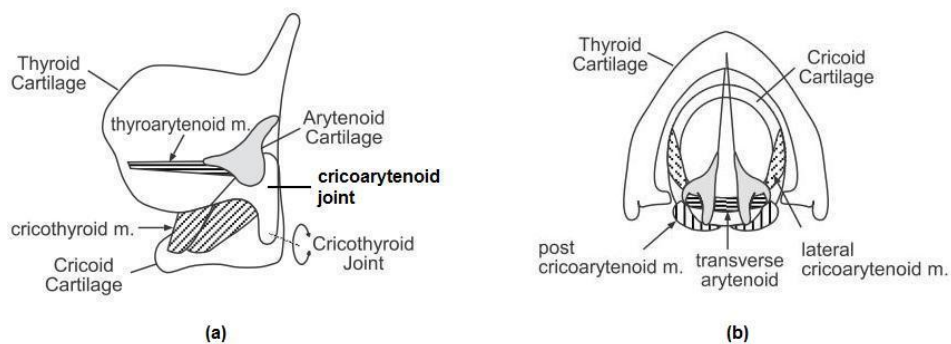


Figure 3. The intrinsic muscles of larynx (adopted from Honda (2004))

Other intrinsic muscles

Vocal folds are also affected by other adductor and abductor muscles. The posterior cricoarytenoid muscle (PLA) is the only abductor muscle, while another three—the interarytenoid muscle (INA), lateral cricoarytenoid (LCA) and the thyroarytenoid (TA) muscles—are adductor muscles (Hirose_1997).

Posterior cricoarytenod muscle (PCA) is the biggest and most powerful muscle of the larynx muscles. It extends from the posterior surface of the cricoid to the each side of

the arytenoid. Contraction of LCA stimulates the translation movement which parts the vocal process. Therefore, the lower vocal folds separated (see Fig. 3b) (Marchal, 2009).

Lateral cricoarytenoid muscle (LCA) originates at the upper edge of the cricoids arch and it insets in the lateral part of the arytenoids cartilage (Fig. 3b). LCA is the smallest of the intrinsic muscles. Its contraction produced the self-pivoting of the arytenoids. As a result, the vocal processes close and the length of the vibration part of the vocal folds is reduced (Handa, 1983).

Thyroarytenoid muscle (TA) originates on the inner surface of the thyroid cartilage and it is slender at the top and thick at the bottom (Fig. 3b). TA is a very fast muscle. It opposes the cricoarytenoid muscle and its main function is to draw forward the arytenoids cartilages, thus shortening the vocal folds and decreasing their tension (Marchal, 2009). The cricothyroid (CT) muscle and thyroarytenoid (TA) muscle work together and affect F0. The CT muscles raise F0 by elongating the vocal folds, whereas the TA muscles raise F0 by increasing the stiffness of vibration when F0 is low.

Extra-laryngeal Mechanisms

Laryngeal Height

It has been shown by previous researches that the larynx moves up and down as F0 rises and falls (Honda, 1999; Steven, 1977). Vilkman (1996) proposed that the vertical movement of larynx played a role in determining F0. Magnetic resonance image (MRI) recoding of the head and neck region were obtained for three male subjects to tracings of the jaw, hyoid bone, laryngeal cartilage, and cervical spine were compared

in high and low F0 range. In the high F0 range, the hyoid bones moved horizontally while the larynx height remained relatively constant. In the low F0, the results indicate that vertical movement of the larynx is a crucial component of the F0 control mechanism.

In addition to fixing the larynx to neighboring organs, extrinsic muscles are responsible for the vertical movement of larynx. They also induce changes in the degree of vocal folds. Bothorel (1980) established the incidence of vertical movement by the hyoid bone during speech. He found that the hyoid bone is systematically higher for voiceless consonants than for voiced consonant. Furthermore, He noticed a correlation between the elevation of the hyoid bone and rises in F0. In high F0 range, hyoid bones are responsible for higher F0.

Extrinsic muscles

How does the extrinsic muscle affect the larynx height? Marchel (2009) pointed out that the direction action muscle on the vertical movement of larynx is stylopharyngeus muscle. Stylopharyngeus muscle originates at the base of the styloid and is attached by pharynx, epiglottis, the upper bone of thyroid cartilage, and the upper edge of the cricoids cartilage. Marchal (2009) indicated that stylopharygeus muscle raises the pharynx and the larynx. Once the stylopharyngeus muscle raises the larynx height, F0 increases. Apart from the elevators muscle of the larynx, the sternothyroid is the depressor of larynx (Fig. 4). This muscle runs from the sternum to the thyroid cartilage. In contracting, it fixes the attachment point of the thyroid and lowers the larynx. Consequently a lowering of F0 is observed (Swashima and Hirose, 1983).

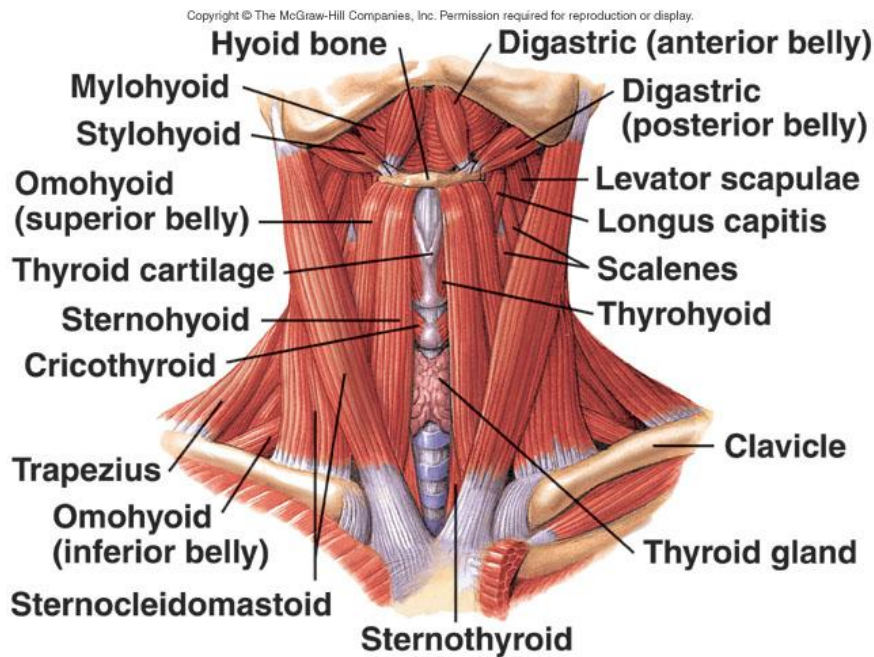


Figure 4. The extrinsic muscles of larynx (adopted from

<http://www.rci.rutgers.edu/~uzwiak/AnatPhys/APFallLect14.html>)

2.2.2 Aerodynamic factors

Air Pressure

Substantial vocal fold vibration occurs when a speaker achieves an appropriate balance among trans-glottal pressure, vocal folds thickness and tension, and degree of abduction of many laryngeal muscles. The laryngeal muscles have been reviewed above. In this section; the aerodynamic factors—air pressure and airflow—will be discussed.

The variation in subglottal pressure plays a central role in speech production. This pressure has to be sufficiently strong enough to overcome the resistance to airflow presented by the glottis and upper airways (Marchal, 2009). Recordings of subglottal pressure during speech reveal it to be posited: by being correlated with F0 (Ohala, 1975). In addition, Lieberman and Atkinson suggest that in certain circumstances the

F0 variations are caused by the subglottal pressure variations which in turn are caused by variation in the pulmonic expiratory force. Miller et al. (1987) established that laryngeal frequency is a function of the length and tension of the vocal fold and of subglottal pressure. He indicated that if tense is kept constant, an increase in subglottal pressure leads to an increase in frequency. Furthermore, Chomsky and Halle (1968) suggested that aspirated stops are produced with heightened subglottal pressure in contrast to unaspirated stops which would have normal subglottal pressure. Taken together, it can be assumed that aspirated stops have a higher F0 than unaspirated stops.

Moreover, trans-glottal air pressure is one of the primary variables controlling the vibratory behavior of the larynx (Stevens, 1977). During vowel production, the trans-glottal pressure is equal to the pressure measured just below the larynx, as there is no pressure decrease in the un-obstruent vocal tract for open vowel articulation. The phonation threshold pressure represents the minimum trans-glottal pressure required for phonation and F0 is influenced by trans-glottal pressure (Owaki_2010). Owaki (2010) investigated the relationship between the change in F0 and per unit change of trans-glottal pressure by a rubber model. The results showed that on the lower side of the modal register ($F_0 < 200\text{-}250\text{ H}$), trans-glottal pressure decreased with increasing F0.

Airflow

Aside from air pressure, airflow is another aspect researchers take into consideration in the aerodynamic factors. Different types of phonation could be determined by the rate of airflow (Reeze and Jongman, 2008). Glottal airflow is known to be the source of sound in voice phonation. Temporal and spectral features of the glottal flow pulses

are of primary importance in both speech analysis and synthesis. According to the Bernoulli Effect, a higher airflow rate decreases the pressure that passes through the vocal folds. The decreased pressure forces the vocal folds to adduct and vibrate more rapidly. Consequently, F0 will be raised.

Dart (1987) measured intraoral air pressure and airflow of Korean fortis and lenis stops in word initial position. He reported that lenis stops shows greater airflow, but less air pressure. Conversely, fortis stops were reported as having less airflow, but greater air pressure. Dart (1987) indicated that the airflow differences during the production of Korean stops are due to the differences in glottal aperture. Dart also suggested the unbalanced pattern of airflow and air pressure for fortis stops is because of the adducted vocal folds before release.

Cho et al. (2002) examined the aerodynamic features of Korean stops in two dialects, namely Seoul standard Korean and Cheju dialect Korean. In the experiment, they included all three stops categories and measured the intraoral pressure and intraoral airflow of Korean bilabials. The results showed that the fortis has less airflow rate than the lenis stops and aspirated stops. The maximum intraoral pressure is smallest for lenis stops, then intermediate for fortis stops and largest for aspirated stops. The patterns of airflow and air pressure reported by Cho et al. (2002) are shown in Figure 5.

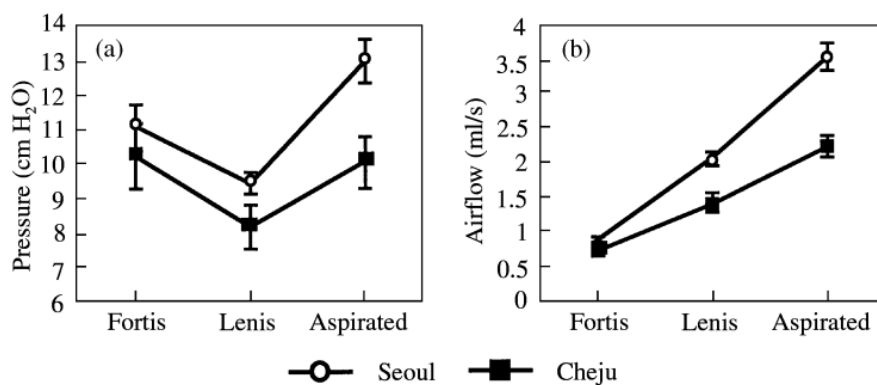


Figure 5. Oral airflow and air pressure from Seoul and Cheju speakers (adopted from Cho et al., 2002, p210))

2.2.3 Speaking rate

Speaking rate is another factor discussed in this paper. It varies naturally within and between speakers (Taso and Weismer, 1997). Rate changes arise from a modified specification of stiffness in the articulators. The biomechanical properties of the moving structures give rise to the individual movement characteristics (Gracco, 1994). It could be conjectured that as more muscular effort is expended to increase stiffness in the system, greater adductory effort is applied to the vocal folds, thus leading to an increase in the sound pressure level and F₀. One of the issues of the present study investigates is whether speaking rate has an influence on increasing F₀.

Dromey and Raming (1998) investigated the effect of speaking rate on F₀. Ten subjects (5 male and 5 female) repeated the sentence “I sell a sapapple again.” under five rate conditions. The result indicated that the F₀ for males increased with rate. However, there was no significant result for females. Furthermore, F₀ variability increased across the range of slow to fast rates.

Prompted by the result of Dromey and Raming (1998), it seems that there is no

consistent agreement on the aspiration perturbation because speaking rates were not well-controlled in the previous researches. In the present study, speaking rate will be one of the factors to examine the relation between aspiration and F0.

2.3 Introduction of Mandarin

The main purpose of the present study is to investigate the relation between aspiration and Mandarin tones. Since the relation between voicing/aspiration and F0 are mentioned above, in the following part, Mandarin syllables and tones and the F0 contour of mandarin tones will be illustrated in the following section. The Mandarin syllable will be presented in the structure of initials, the finals, and the tones.

2.3.1 Initials

The initial represents the consonantal beginning of a syllable. Since Mandarin does not have consonant cluster, the consonantal beginning of a syllable can only be a single consonants (Li, 1989). Mandarin has 21 initials. The initials of Mandarin are provided in table 1 in terms of the International Phonetic Alphabet (IPA).

Table 1. Mandarin initial consonants

PQA Manner	bilabial	Labiodental	alveolar	retroflex	Palate-alveolar	Velar
Plosive	p, p ^h		t, t ^h			k, k ^h
Nasal	m		n			
Fricative		f	s	ʂ	ç	ɣ
Affricate			ts, ts ^h	tʂ, tʂ ^h	tç, tç ^h	
Lateral			l	ɭ		

2.3.2 Finals

The final is the part of the syllable excluding the initial (Li, 1989). There are 35 finals in Mandarin and they are listed in table 2 in IPA symbols. The velar nasal [ŋ] occurs only as a part of a final, never as an initial.

Table 2. Mandarin finals

6 simple finals	a, e, i, o, u, y
13 compound finals	ai, au, ei, ia, iau, iɛ, iou, ou, ua, uai, yɛ, uei, uo
16 nasal finals	an, ɤn, iɛn, in, uan, üɛn, uen, yn aŋ, eŋ, iaŋ, iŋ, iuŋ, uŋ, uaŋ, uɤŋ

2.3.3 Mandarin tones

Lexical tones are pitch patterns that provide contrast in word meaning, especially in tonal languages such as Mandarin (Chao, 1948; Xu, 1997), Taiwanese (Peng, 1997), Korean (Kim, 2002), and Thai (Abramson, 1979, Gandour et al., 1994). Pitch or tone is a function of the rate of vocal folds vibration (Ohala, 1979). Changes in F0 are made by manipulating tension in the vocal folds and the tension is increased or decreased by the laryngeal muscles (Gay, 1972; Lofqvust, 1984; Vilkmán, 1996). Mandarin phonetically distinguishes four tones, with tone 1 having high-level pitch, tone 2 high-rising pitch, tone 3 low-dipping pitch, and tone 4 high-falling pitch (Chao, 1948) (Table 3).

Table 3. Mandarin character differ in tones

Chinese character	English gloss	Tone number	Tone description	Tone range
媽	Mother	First	High-level	55

麻	Hemp	Second	Rising	35
馬	Horse	Third	Dipping	214
罵	Scold	Fourth	Falling	51
嘛	Question marker	Neutral	Neutral	variable

Many phonetic studies have investigated the F0 contour of Mandarin (Moore & Jngman, 1997; Xu, 1997; Liu, 2004). These studies indicate that F0 height and F0 contour are the primary acoustic parameters to characterized Mandarin tones. Xu (1997) examined acoustic variations of tones in Mandarin. Eight male native speaker of Mandarin were asked to produce two kinds of reading lists (monosyllabic and disyllabic reading list) in different carrier sentence as well as in isolation. F0 curves (maximum F0, and minimum F0 of each segment) and F0 value at five points (beginning, one quarter, midpoint, three quarters, and end of the segment) were measured. The result showed that the F0 patterns of tones produced in isolation reflect relatively directly the canonical forms of the tones. In multisyllable utterances, the canonical forms will be distorted by various factors, including the adjacent onset and offset value of the neighboring tones.

An illustration of the mean F0 contours of Mandarin's four tones of the monosyllabic /ma/ in isolation (Fig. 6). Time is normalized, and all tones are plotted with their average duration proportional to the average of tone 3. Tone 1 is relatively high compared to the other tones. Tone 2 exhibits a rising pattern, and its onset occurs in the middle region of the F0 range. The contour of tone 3 occupies the lowest region of the F0 overall range area and it is close in frequency to that of tone 2. Lastly, tone 4 begins at highest region and falls to the bottom if the F0 range.

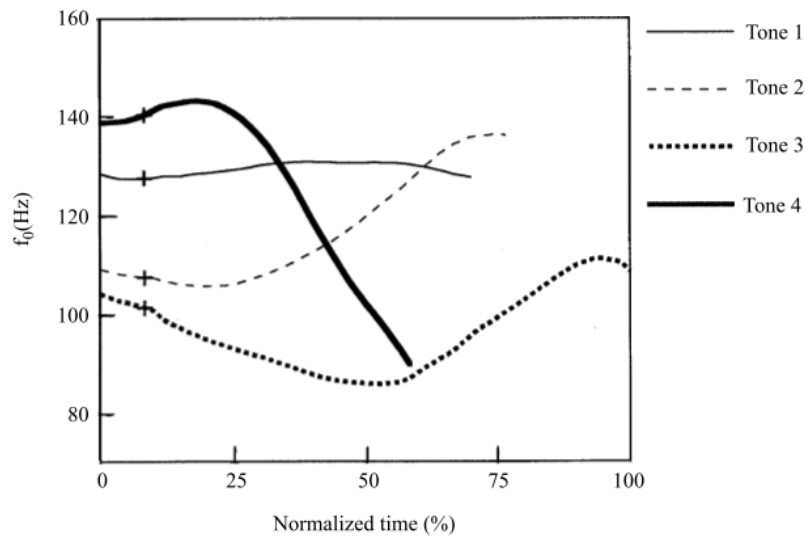


Figure 6. Mean F0 contour of four Mandarin tones in the monosyllable /ma/ (adopted from Xu, 1997, p67).

In order to investigate this issue further, the present study is designed on a rigorous experimental paradigm and aims to uncover possible modulating factors for the perturbation effect of aspiration. More consistent measurements of this will be provided and help to clarify the effect of consonant aspiration on the F0 of following vowels. Stimuli adopted in the experiment are composed of three places of articulation, four vowels, and four tones. According to the significant rising effect only found in female (Lai, 2004) and the opposite result which F0 is lower after aspirated stops in Mandarin (Xu and Xu, 2003), Gender is one of the factors in the present study. Apart from these, Rate which inspired by the result of Dromey and Raming (1998) is included in order to test their effect on the perturbation phenomenon.

Chapter III Methodology

The main goal in the present study is to test if aspiration has a rising effect on F0 of the following vowels in Mandarin. There are 17 voiceless consonants in Mandarin, and 12 out of them, 6 stops and 6 affricates, are paired by the aspirated and unaspirated distinction. The six pairs are: /p/, /p^h/, /t/, /t^h/, /k/, /k^h/, /ts/, /ts^h/, /tʂ/, /tʂ^h/, /tɕ/, and /tɕ^h/.

In this section, subject information, stimuli construction, recording instruments, experiment procedure, data measurement, and analysis are described. The method which is familiar with the delayed shadowing paradigm (Carroll, 2008) was adopted in the experiment. In the delayed shadowing task, subjects have to repeat immediately what they hear. However, in the present study, subjects repeated the stimuli in a carrier sentence after a controlled time. The interval time between the end of sound file and the hint on the screen is 500 ms seconds.

3.1 Subject

In the present experiment, two stages of recording were needed to perform the paradigm of delayed shadowing. Recording from the first stage served as the token for the task in the second stage. There are 21 subjects who participated in the present study. A female speaker was recorded in the first stage and 20 speakers (10 males, 10 females) in the second stage.

The first stage participant

In the first recording, a female subject (age 23; student of National Chiao Tung

University) was recorded. The subject has no reported history of speech or hearing disorders.

The second stage participant

There are 20 native speakers (10 males and 10 females) of Mandarin Chinese participated in the second recording. They are non-smokers with no history of speech or hearing disorder. These participants (age range from 20 to 30) are students from National Chiao Tung University, National Tsing Hua University, and National Hsinchu University of Education.

3.2 Stimuli

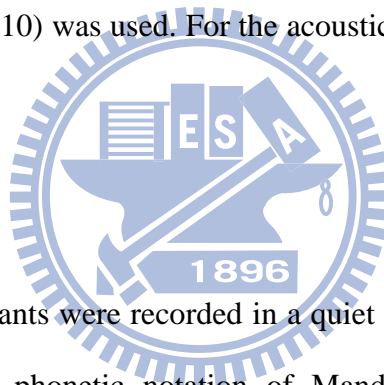
Six stop consonants of 3 places of articulation (bilabial, alveolar, velar) are included: /p/, /t/, /k/, /p^h/, /t^h/, and /k^h/. Each stop consonant was placed in the initial position of each target word. Every stop consonant is followed by a vowel /i/, /a/, and /u/. Because there are too many lexical gaps in the combination of velar stop (/k/ and /k^h/) and the high front vowel (/i/), central vowel /ə/ was added in the velar stop and vowel combination. All the stimuli are: /pi/, /pa/, /pu/, /p^hi/, /p^ha/, /p^hu/, /ti/, /ta/, /tu/, /t^hi/, /t^ha/, /t^hu/, /ki/, /kə/, /ka/, /ku/, /k^hi/, /k^hə/, /k^ha/, /k^hu/. Each token is matched with the four tones in Mandarin. A complete wordlist can be found in Appendix 1 (total 80 target words).

Because the speaking rate (fast vs. slow) is one of the factors in the present study, in order to test the perturbation effect of speaking rates, the isolated words recording was not adopted. All the stimuli were embedded in a carrier sentence: *Wo3 nian4 ___ zhe4 ge4 zi4* ('I read this word ___').

3.3 Instrument

All utterances were recorded using a MR-1000 recorder and a SHURE SM 57 dynamic instrument microphone. The microphone was adjusted to maintain a constant distance (15 cm) from the participants. The participants (2nd stage) heard stimuli through a Creative HQ-1500 headphone.

In the first stage recording, an Acer computer and Microsoft Office PowerPoint 2007 were used. Stimuli were presented by the PowerPoint for controlling the speaking rate in the first stage. In order to present participants stimuli and visual hints at the same time in the second recoding, the software Paradigm was used. Praat (by Boersma, Paul & Weenink, David, 2010) was used. For the acoustic analysis, statistical analysis was done in SPSS 15.0.



3.4 Procedure

In both recordings, participants were recorded in a quiet room. All stimuli in the first stage were transcribed in phonetic notation of Mandarin (bo-po-mo-fo) for the convenience of pronunciation and to denote lexical gaps mentioned earlier. The stimuli (total 80 target words) were shown at random on a computer screen. Speaking rate (slow vs. fast) was regulated via the slide transition of Microsoft Office PowerPoint. The slow speaking rate is close to normal speaking speed whereas fast speed is about two times faster than norm speed.

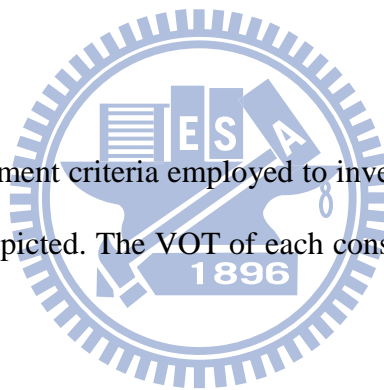
Each token in the carrier sentence was read once in the first stage. A total of 160 sentences were recorded in the first stage (80 * 2 rates = 160). Before the recording, participant in the first stage had the practice section for ten minutes. In the recording, if the participant felt tired, participant can take a break whenever she wanted. After

finished the first stage recording, each carrier sentence will be separated by Praat, and then saved each of them as a sound file for the second stage recording.

The participants heard the sounds recorded in the first stage and then were asked to repeat what they heard with the same speaking rate in the second stages. These sound files were presented using Paradigm in which participants could hear the sounds and see the visual cue which reminded participants to repeat the sentence. As mention above, the interval time is 500 ms between the sound file and visual hint on the screen. Each token was repeated three times. A total of 480 sentences were recorded by every subject. There were 9600 tokens ($480 * 20 = 9600$) to measure.

3.5 Measurements

In this section, the measurement criteria employed to investigate the acoustic property of stops and vowels was depicted. The VOT of each consonant and F0 of each vowel were measured.



Voice Onset Time (VOT)

The stimuli were voiceless aspirated and unaspirated stops, so first measurement is VOT. VOT is the duration of the period of time. The measurement was between the release of a plosive and the beginning of vocal fold vibration (Fig. 7). This period is usually measured in milliseconds (ms).

Fundamental Frequency (F0)

F0 of each vowel was measured from the onset of F1 to the offset of F2 as shown in Fig. 7. There are three F0 measurements in the present study. One is normalized F0, another is first 100 ms, and the other is onset F0. The reason why these three ways of

F0 measurements adopted in the present research will be explained in the discussion section.

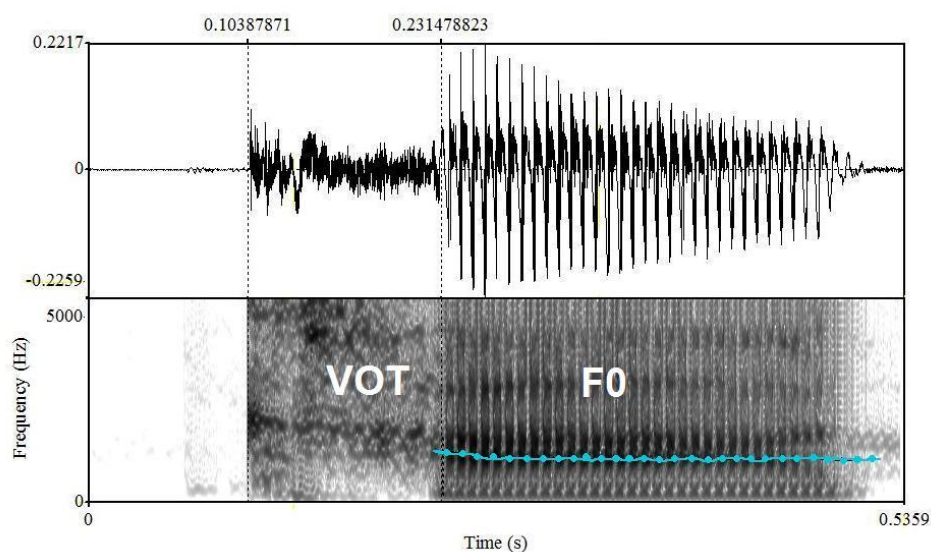


Figure 7. The sample *kha* was pronounced by a male speaker in Tone 1 shown in Praat screen. The upper panel displays the waveform while the lower panel shows the spectrogram of VOT (128 ms) the tonal contour (289 ms). The x-axis is time in seconds and the y-axis is frequency in Hz.

The first one is that F0 was measured by using a normalization script developed by Yi Xu. Normalization is completed by 10 to get the sampling time steps (Fig. 8). For example, a 500 ms F0 contour would be normalized using 50 ms as the sampling steps. The advantage of using the normalized script is that all vowels will equally have ten values no matter the duration of the vowel was.

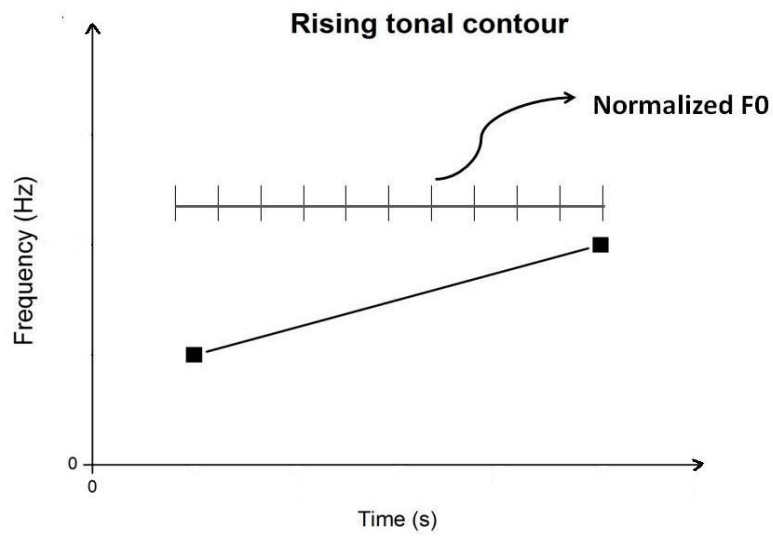


Figure 8. This picture shows the normalized F0 in the rising tonal contour.

Besides, the second measurement of F0 is that the F0 contour of every 10 ms was measured as well. Only 10 points was extracted (including the onset F0) of each F0 contour as shown in Fig. 9. If the onset is at 20 ms, then the next point will be 30 ms and so on. The reason why every 10 ms measurement is added in the study is that normalized F0 measurement may dissipate the rising or lowering effect of the aspiration on whole F0.

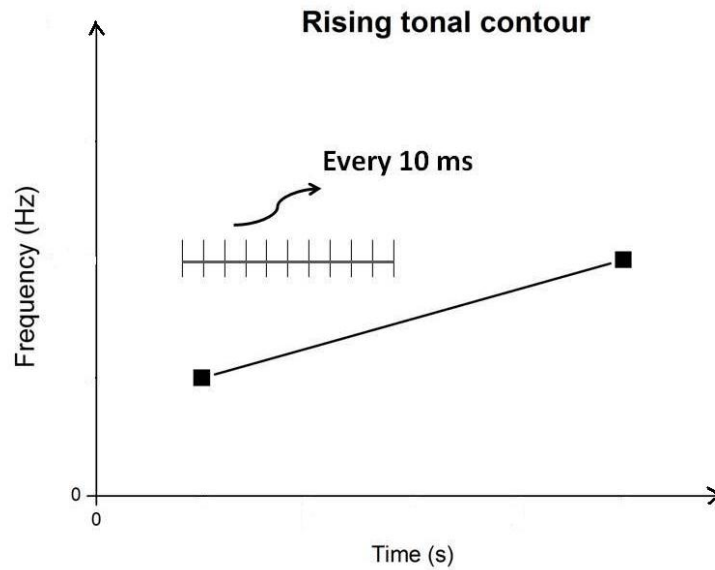


Figure 9. This picture shows the every 10 ms F0 in the rising tonal contour.

3.6 Statistical analysis

Both the measurements of VOT and F0 value will be analyzed by the statistical way with SPSS. The statistic results of VOT and F0 were showed in the following section. Since there were three factors (Aspiration, gender, and speaking rate) involved in the experiment, the three-way ANOVA tests were performed in SPSS for statistical analysis. The within-subjects factors are Aspiration (aspirated and unaspirated) and Speaking rate (fast and slow). The between-subject factor is Gender (male and female).

Chapter IV Results

A series of three-way repeated measures analysis of variance (ANOVA) were conducted to evaluate the effect of Gender, Aspiration and Rate on VOT and F0. The within subjects factors are Aspiration (aspirated and unaspirated) and Speaking rate (fast and slow). The between subject factor is Gender (male and female). A series of repeated measurement is applied on VOT and F0.

4.1 Voice Onset Time (VOT)

The VOT will be illustrated in the following part.

Table 4. Result of a 3-way ANOVA test on VOT

Source of effect	F	P	
Aspiration	(1, 78)= 7578.1	$p < .001$	*
Rate	(1, 78)= 872.7	$p < .001$	*
Gender	(1, 78)= .226	$p = .636$	n.s
Aspiration x Gender	(1, 78)= .82	$p = .367$	n.s
Gender x Rate	(1, 78)= 9.13	$p = .003$	*
Aspiration x Rate	(1, 78)= 619.98	$p < .001$	*
Aspiration x Rate x Gender	(1, 78)=12.17	$p = .001$	*

The results of ANOVA of VOT are shown in Table 4. The main effect of Aspiration is significant [$F(1, 78) = 7578.1, p < .001$]. Duration after aspirated stops (102.8 ms) is significantly longer than after unaspirated stops (24.8 ms) (Fig. 10). The main effect of Rate is significant [$F(1, 78) = 872.7, p < .001$]. Duration in slow speaking rate (71.1 ms) is longer than that in fast speaking rate (52.5 ms) (Fig. 11). The main effect of Gender was not significant [$F(1, 78) = 226, p < .636$].

The interaction between Gender and Aspiration [$F(1, 78) = .82, p = .367$] was not significant. The interaction between Gender and Rate [$F(1, 78) = 9.13, p = .003$] and between Aspiration and Rate [$F(1, 78) = 619.98, p < .001$] were significant (Fig. 13 and Fig. 14).

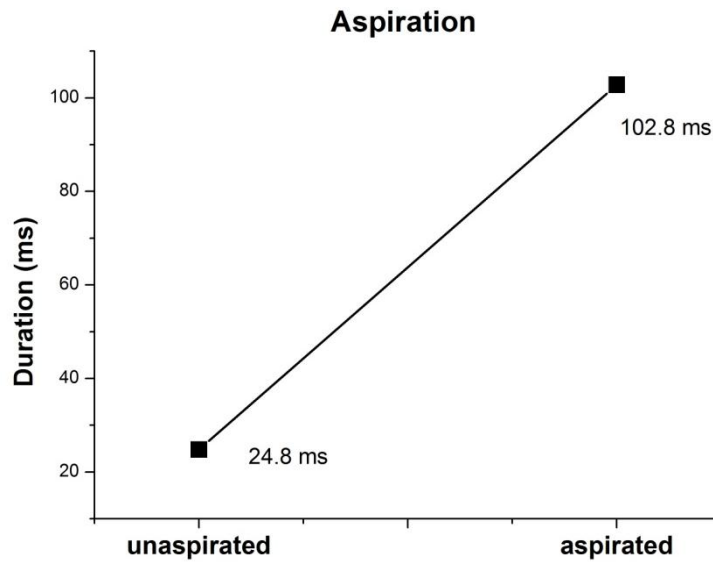


Figure 10. The main effect of Aspiration on VOT

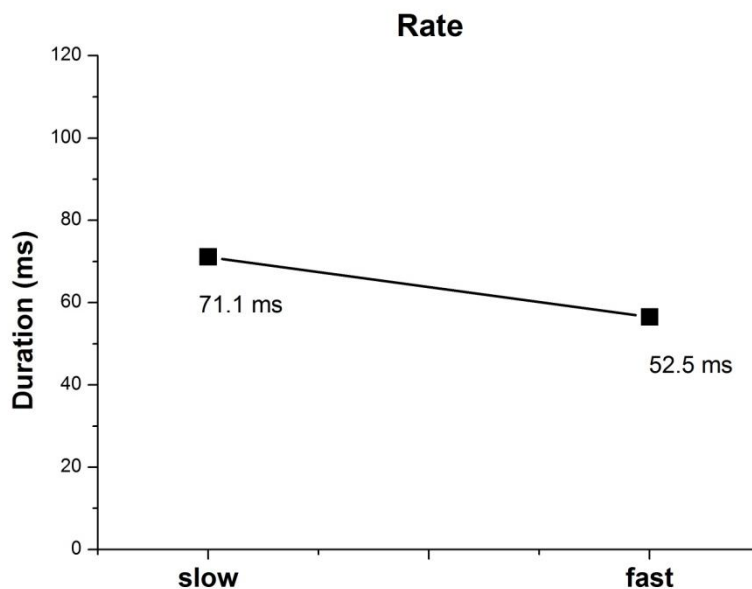


Figure 11. The main effect of Rate on VOT

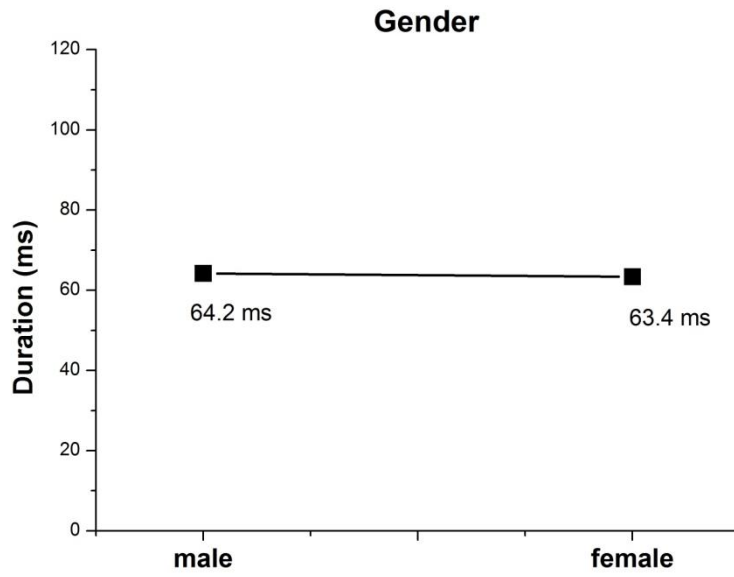


Figure 12. The main effect of Gender on VOT

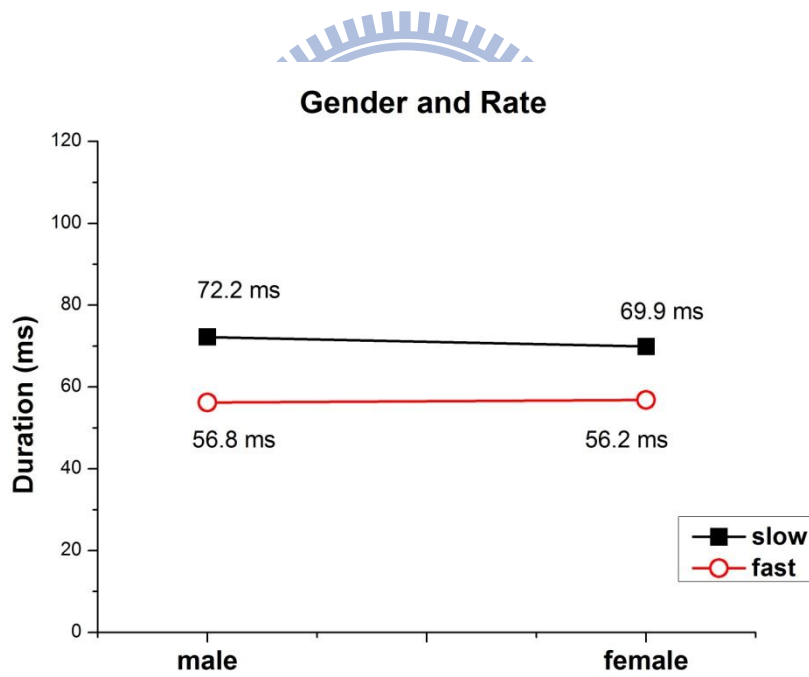


Figure 13. The interaction between Gender and Rate on VOT

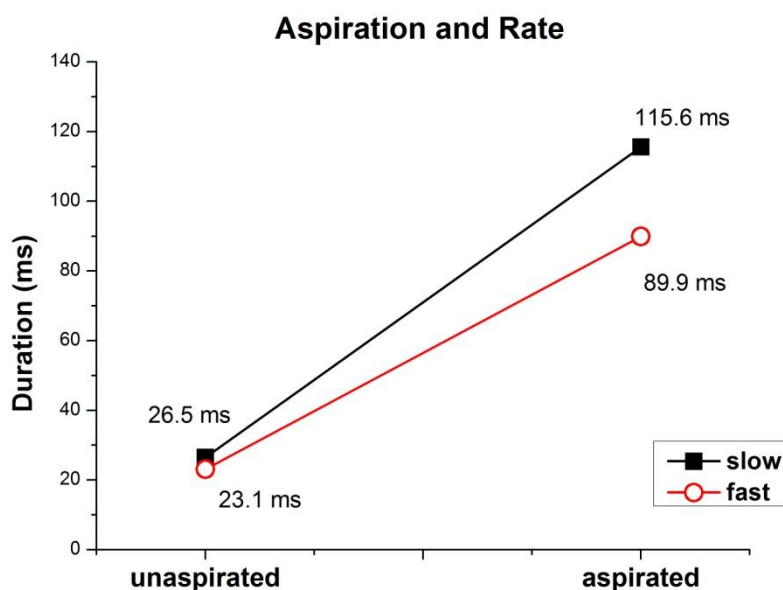


Figure 14. The interaction between Aspiration and Rate on VOT

There are several findings in the analyses above. First, the VOT is longer in aspirated consonants than that in the unaspirated consonants, and this is true across two speaking rates. Also both females and males show the tendency (i.e. both of them produce the aspirated consonants longer than the unaspirated ones.) moreover, there is an interaction between aspiration and rate. Specifically, the faster the speech is, the VOT is shorter; the slower the speech is, the VOT is longer.

Besides, VOT was also analyzed separately by four tones. The results of different tones were shown in Appendix. The results across four tones are as the same as the results of overall VOT (tones are not analyzed separately). VOT is longer in aspirated consonants than that in unaspirated consonants. As the speaking rate is slower, the VOT is longer. However, there exists no significant effect in Gender. There is an interaction between Aspiration and Rate (i.e. the faster the speaking rate is, the VOT is shorter and vice versa). The F0 analysis will be depicted in the following section.

4.2 Fundamental Frequency (F0) Measurement

There are three types of measurement of F0. One is normalized F0, another is onset F0, and the other is the first 100 ms F0. In the following section, the results will focus on the first 100 ms F0. The other two results (normalized and onset F0) are shown in Appendix.

Tone 1

Table 5. Result of an ANOVA test on first 100 ms F0 (Tone 1)

Source of effect	F	<i>P</i>	
Aspiration	(1, 198)= 91.6	$p < .001$	*
Rate	(1, 198)= 58.39	$p < .001$	*
Gender	(1, 198)= 4468.62	$p < .001$	*
Aspiration x Gender	(1, 198)= 5.72	$p < .001$	*
Gender x Rate	(1, 198)= 12.78	$p < .001$	*
Aspiration x Rate	(1, 198)= 29.96	$p = .001$	*
Aspiration x Rate x Gender	(1, 198)= 15.26	$p = .568$	n.s

The results of ANOVA on Tone 1 are shown in Table 5. The main effect of Aspiration is significant [$F(1, 198) = 91.6, p < .001$]. F0 after aspirated stops (162.7 Hz) is significantly higher than after unaspirated stops (158.2 Hz) (Fig. 15). The main effect of Gender is also significant [$F(1, 198) = 4468.62, p < .001$]. Female's F0 (205.6 Hz) is higher than male's (129.3 Hz) (Fig. 16). The main effect of Rate is significant [$F(1, 198) = 58.39, p < .001$]. F0 in fast speaking rate (162.7 Hz) is higher than that in slow speaking rate (159.9 Hz) (Fig. 17).

The interaction between Aspiration and Gender is significant [$F(1, 198) = 5.72, p < .001$] (Fig. 18). Furthermore, the interaction between Gender and Rate [$F(1, 198)$

= 12.7, $p < .001$] and between Aspiration and Rate [$F = (1, 198) = 29.96, p = .001$] are also significant (Fig. 19 and Fig. 20).

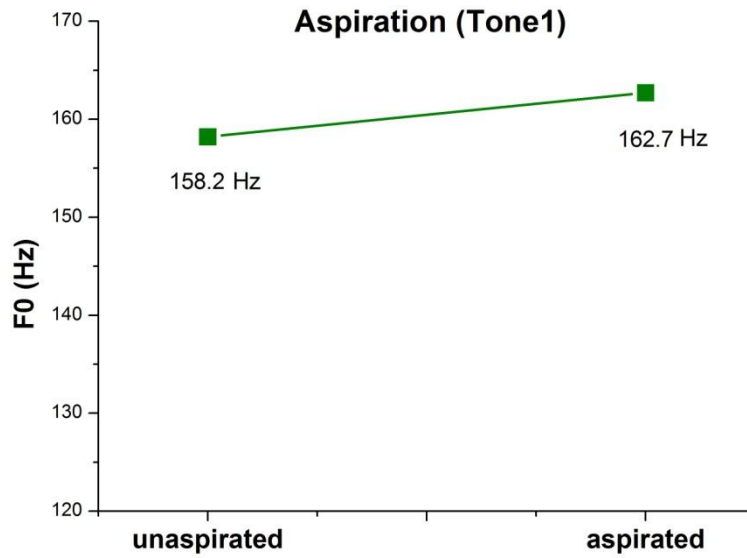


Figure 15. The main effect of Aspiration on 100 ms F0 in Tone 1

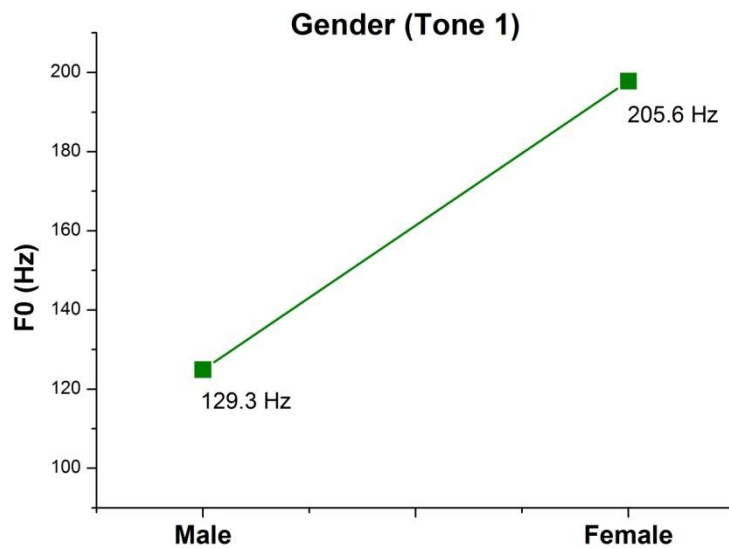


Figure 16. The main effect of Gender on 100ms F0 in Tone 1

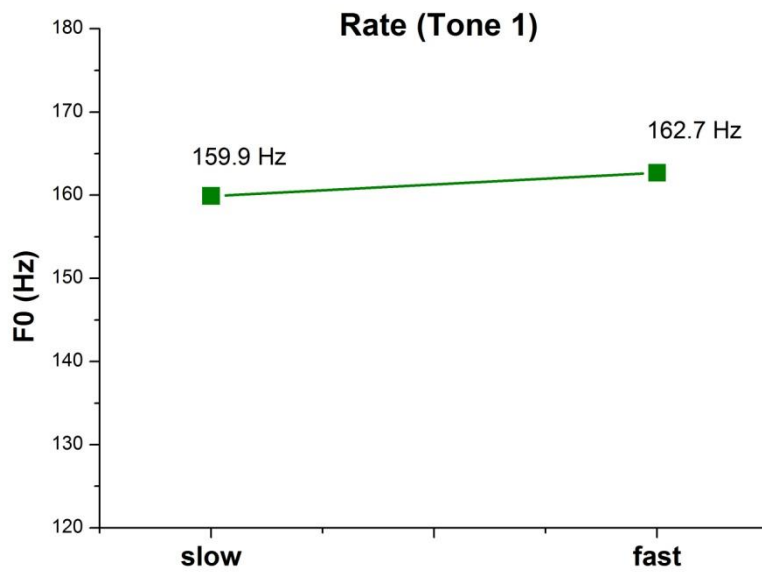


Figure 17. The main effect of Rate on 100 ms F0 in Tone 1

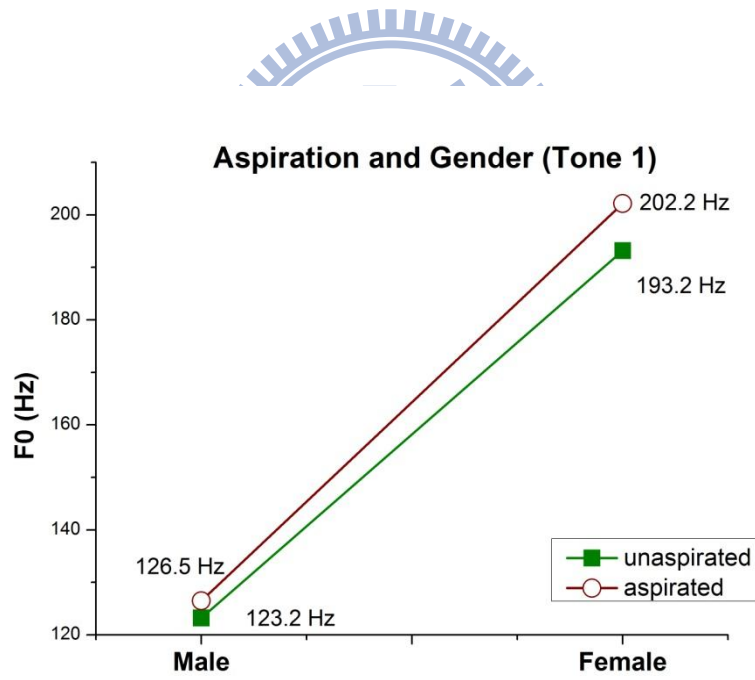


Figure 18. The interaction between Aspiration and Gender on 100 ms F0 in Tone 1

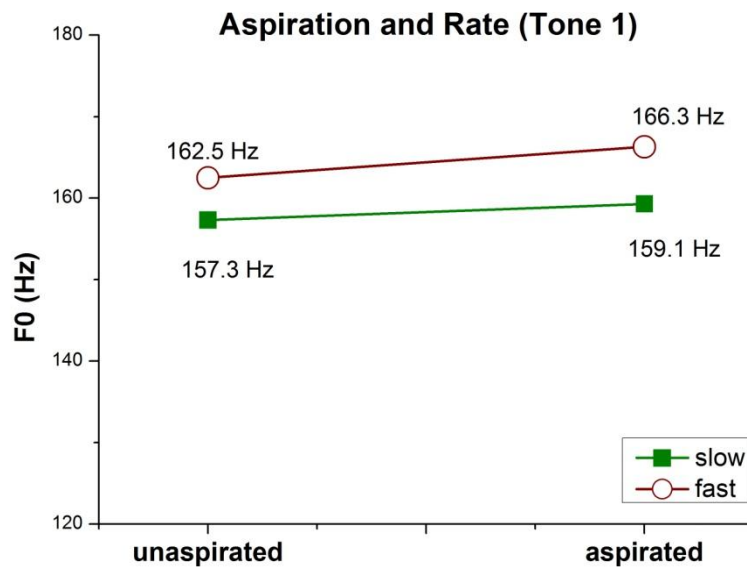


Figure 19. The interaction between Aspiration and Rate on 100 ms F0 in Tone 1

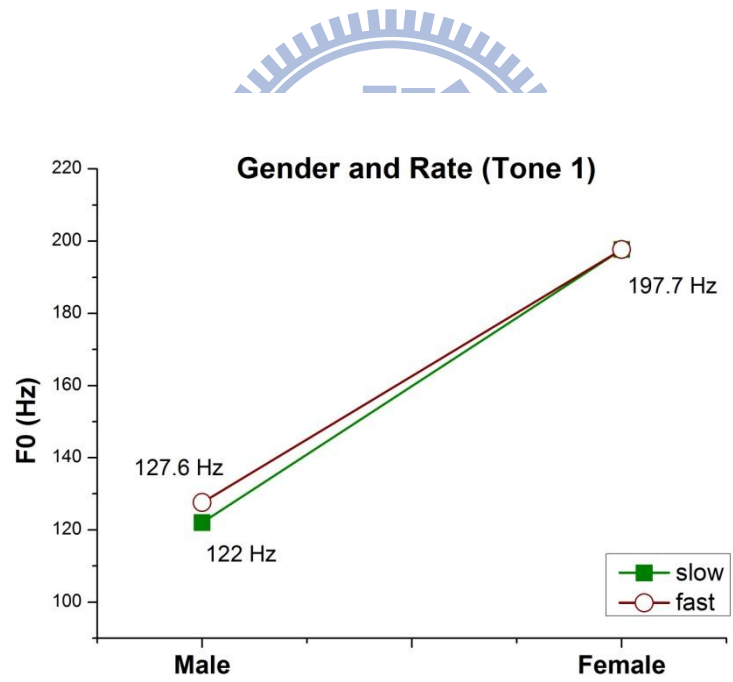


Figure 20. The interaction between Gender and Rate on 100 ms F0 in Tone 1

Tone 2

Table 6. Result of an ANOVA test on first 100 ms F0 (Tone 2)

Source of effect	F	P	
Aspiration	(1, 198)= 203.64	$p < .001$	*
Rate	(1, 198)= 80.04	$p < .001$	*
Gender	(1, 198)= 7202.369	$p < .001$	*
Aspiration x Gender	(1, 198)= 42.95	$p = .018$	*
Gender x Rate	(1, 198)= 80.67	$p < .001$	*
Aspiration x Rate	(1, 198)= 11.64	$p < .001$	*
Aspiration x Rate x Gender	(1, 198)= .327	$p < .001$	*

The results of ANOVA on Tone 1 are shown in Table 6. The main effect of Aspiration was significant [$F(1, 198) = 203.64, p < .001$]. F0 after aspirated stops (134.2 Hz) is significantly higher than after unaspirated stops (128.9 Hz) (Fig. 21). The main effect of Gender was significant [$F(1, 198) = 7202.369, p < .001$]. F0 of female (159.5 Hz) is higher than that of male (103.6 Hz) (Fig. 22). The results showed a significant effect of Rate [$F(1, 198) = 80.04, p < .001$]. F0 in fast speaking rate (133.6 Hz) is higher than that in slow speaking rate (129.5 Hz) (Fig. 23).

The interactions between Aspiration and Gender [$F(1, 198) = 42.95, p = .018$], between Gender and Rate [$F(1, 198) = 80.67, p < .001$], and between Aspiration and Rate [$F(1, 198) = 11.64, p < .001$] are all significant (Fig. 24, 25 and 26).

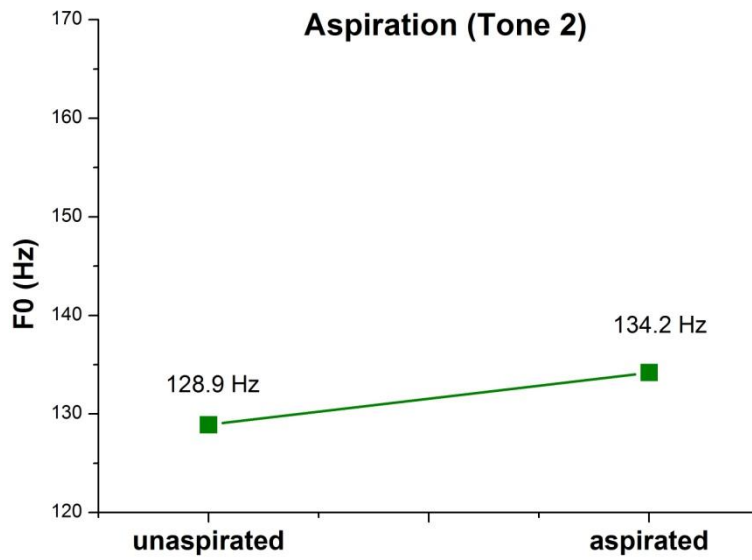


Figure 21. The main effect of Aspiration on 100 ms F0 in Tone 2

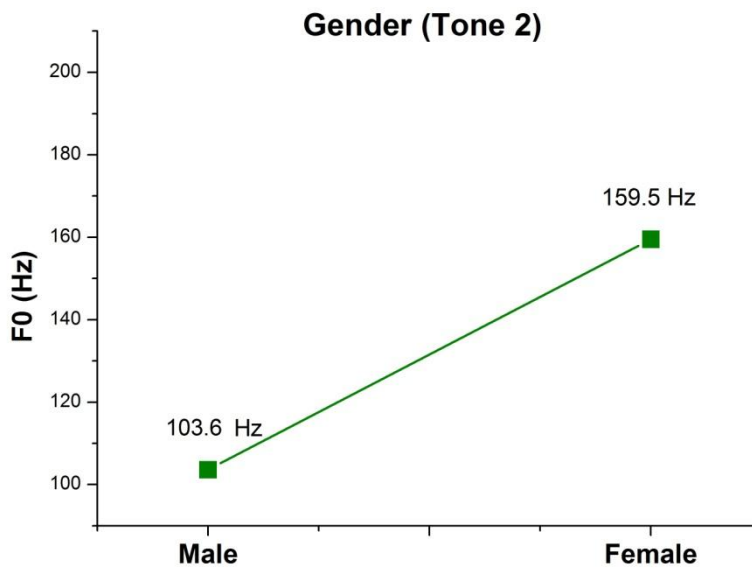


Figure 22. The main effect of Gender on 100 ms F0 in Tone 2

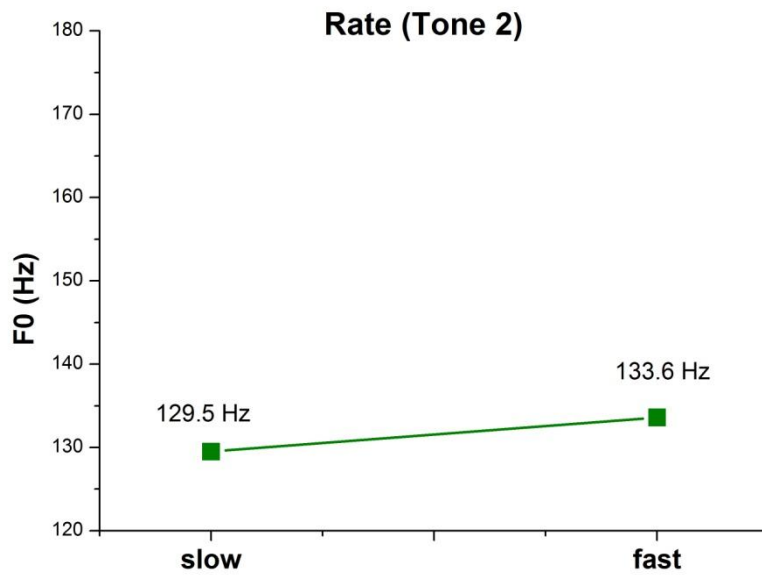


Figure 23. The main effect of Rate on 100 ms F0 in Tone 2

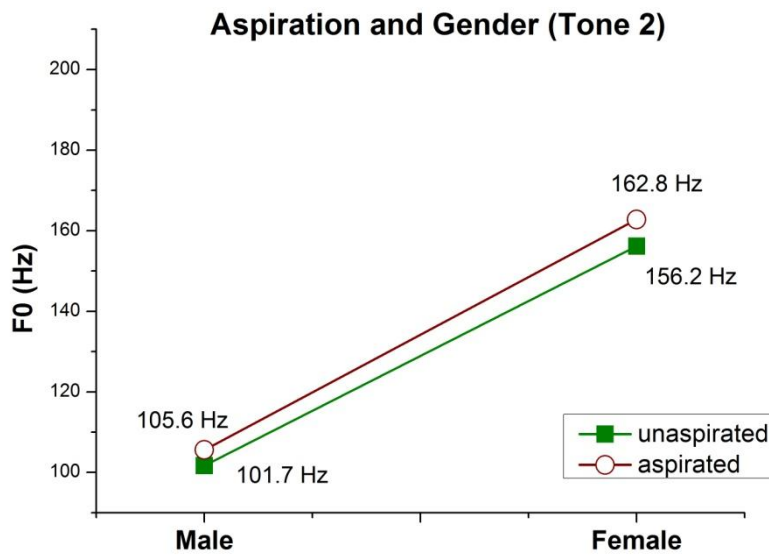
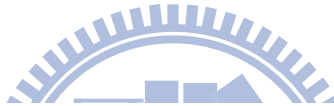


Figure 24. The interaction between Aspiration and Gender on 100 ms F0 in Tone 2

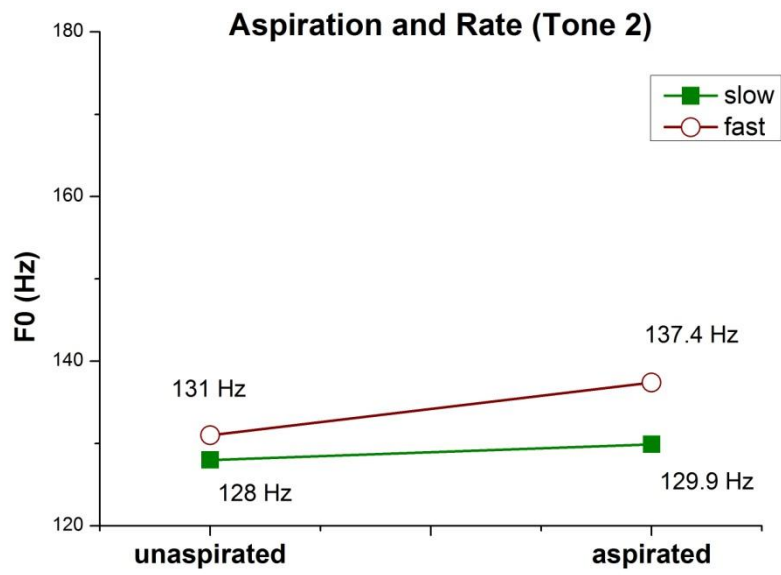


Figure 25. The interaction between Aspiration and Rate on 100 ms F0 in Tone 2

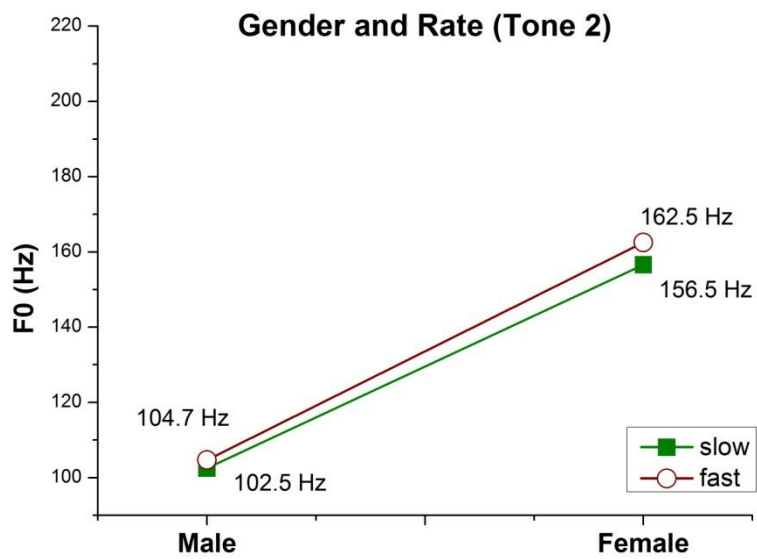
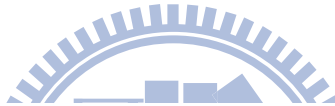


Figure 26. The interaction between Gender and Rate on 100 ms F0 in Tone 2

Tone 3

Table 7. Result of an ANOVA test on first 100 ms F0 (Tone 3)

Source of effect	F	<i>P</i>	
Aspiration	(1, 198)= 109.48	$p < .001$	*
Rate	(1, 198)= 21.25	$p < .001$	*
Gender	(1, 198)= 1326.31	$p < .001$	*
Aspiration x Gender	(1, 198)= 133.71	$p < .001$	*
Gender x Rate	(1, 198)= 40.2	$p < .001$	*
Aspiration x Rate	(1, 198)= 1.75	$p = .187$	n.s
Aspiration x Rate x Gender	(1, 198)= 17.45	$p < .001$	*

The results of ANOVA on Tone 3 are shown in Table 7. The main effect of Aspiration is significant [$F(1, 198) = 109.48, p < .001$]. F0 is higher after aspirated stops (134 Hz) than that after unaspirated stops (125.7 Hz) (Fig. 27). The main effect of Gender is significant [$F(1, 198) = 1326.31, p < .001$]. Female's F0 (155.6 Hz) was found higher than male's (104.1 Hz) (Fig. 28). The main effect of Rate is also significant [$F(1, 198) = 21.25, p < .001$]. F0 in slow speaking rate (131.4 Hz) is higher than that in fast speaking rate (128.3 Hz) (Fig. 29).

The interactions between Gender and Aspiration (Fig. 30) [$F(1, 198) = 133.71, p < .001$], and between Gender and Rate (Fig. 31) [$F(1, 198) = 40.2, p < .001$] are significant. However, the interaction between Aspiration and Rate (Fig. 32) [$F(1, 198) = 1.75, p = .187$] is not significant.

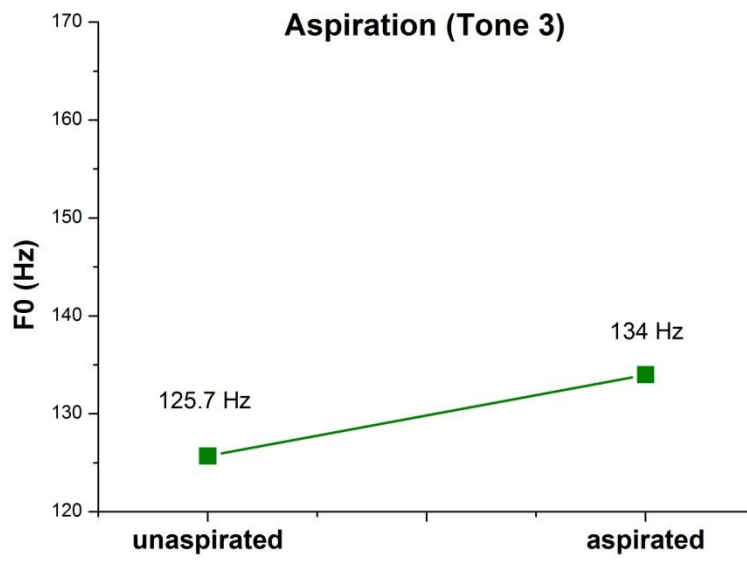


Figure 27. The main effect of Aspiration on 100 ms F0 in Tone 3

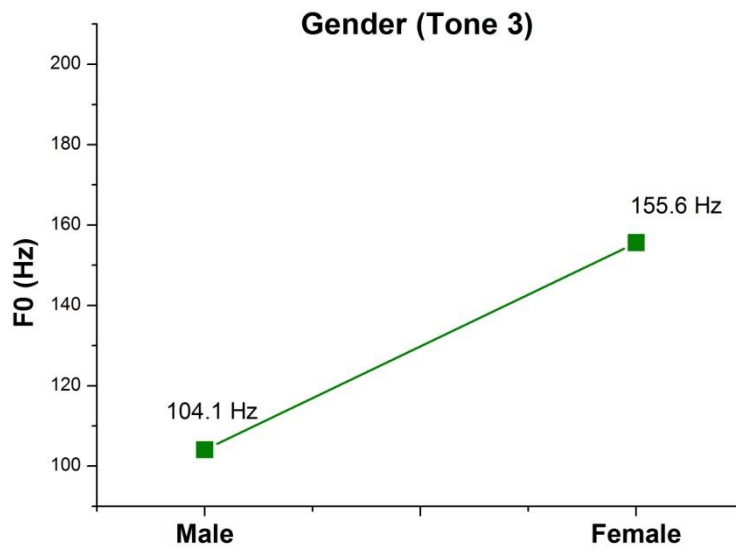


Figure 28. The main effect of Gender on 100 ms F0 in Tone 3

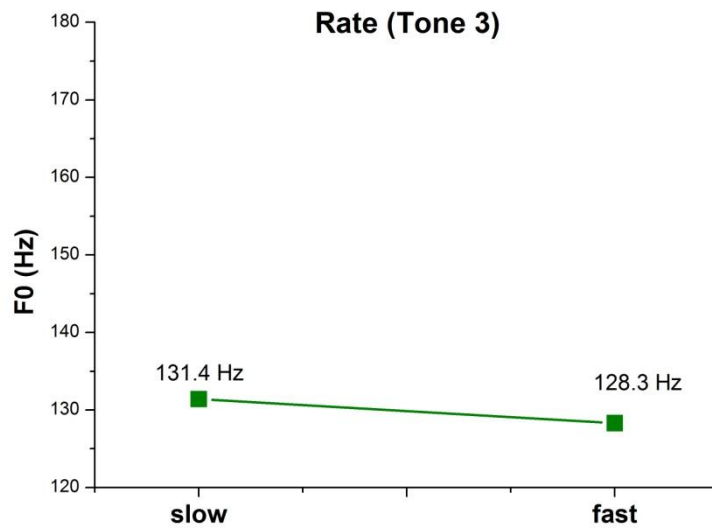


Figure 29. The main effect of Rate on 100 ms F0 in Tone 3

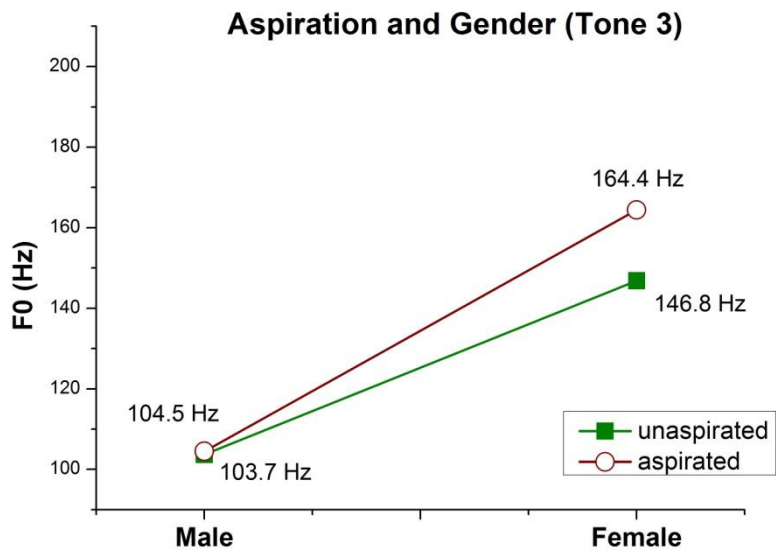


Figure 30. The interaction between Aspiration and Gender on 100 ms F0 in Tone 3

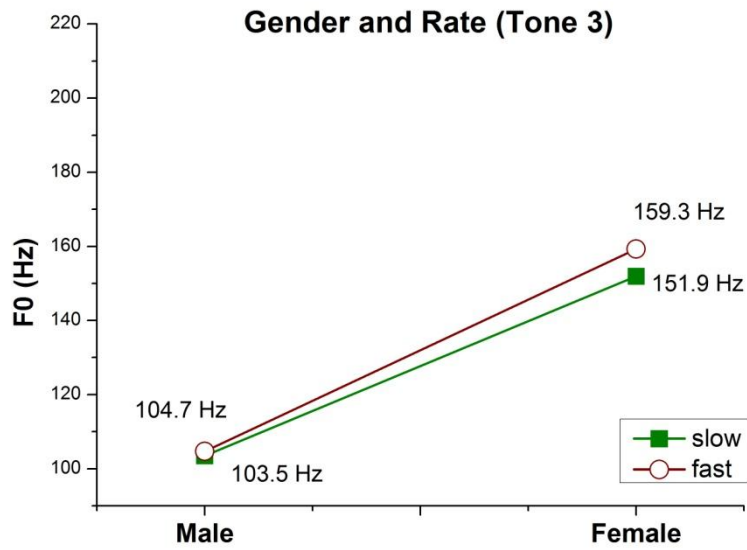


Figure 31. The interaction between Gender and Rate on 100 ms F0 in Tone 3

Tone 4

Table 8. Result of an ANOVA test on first 100 ms F0 (Tone 4)

Source of effect	F	P	
Aspiration	(1, 198)= 93.84	$p < .001$	*
Rate	(1, 198)= 213.61	$p < .001$	*
Gender	(1, 198)= 1945.84	$p < .001$	*
Aspiration x Gender	(1, 198)= 24.81	$p < .001$	*
Gender x Rate	(1, 198)= 183.68	$p < .001$	*
Aspiration x Rate	(1, 198)= 59.81	$p = .037$	*
Aspiration x Rate x Gender	(1, 198)= 79.53	$p = .016$	*

The results of ANOVA on Tone 4 are shown in Table 8. The main effect of Aspiration is significant [$F(1, 198) = 93.84, p < .001$]. F0 after aspirated consonants (169.6 Hz) is higher than that after unaspirated consonants (165.4 Hz) (Fig. 32). The main effect of Gender is significant [$F(1, 198) = 1945.84, p < .001$]. Female's F0 (205.6 Hz) is higher than male's (129.3 Hz) (Fig. 33). The main effect of Rate [$F(1, 198) = 213.61,$

$p < .001$] is significant. F0 is higher at slow speaking rate (170 Hz) than at fast speaking rate (164.9 Hz) (Fig. 34).

The interaction between Gender and Aspiration is significant [$F(1, 198) = 24.81, p < .001$] (Fig. 35). The interactions between Gender and Rate [$F(1, 198) = 183.68, p < .001$], and between Aspiration and Rate [$F(1, 198) = 59.81, p = .037$] are significant as well (Fig. 36 and Fig. 37).

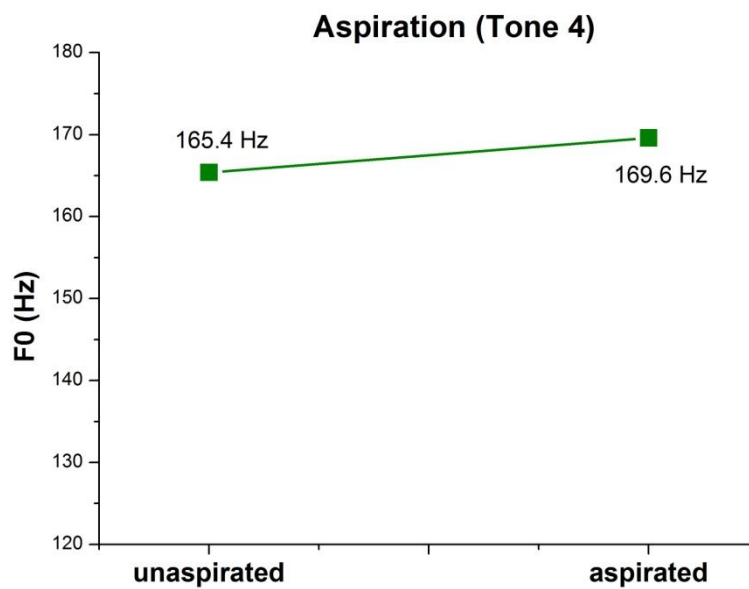


Figure 32. The main effect of Aspiration on 100 ms F0 in Tone 4

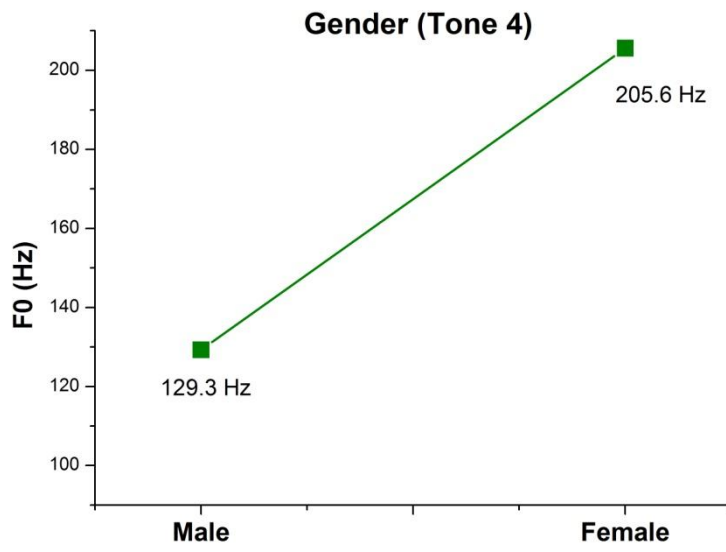


Figure 33. The main effect of Gender on 100 ms F0 in Tone 4

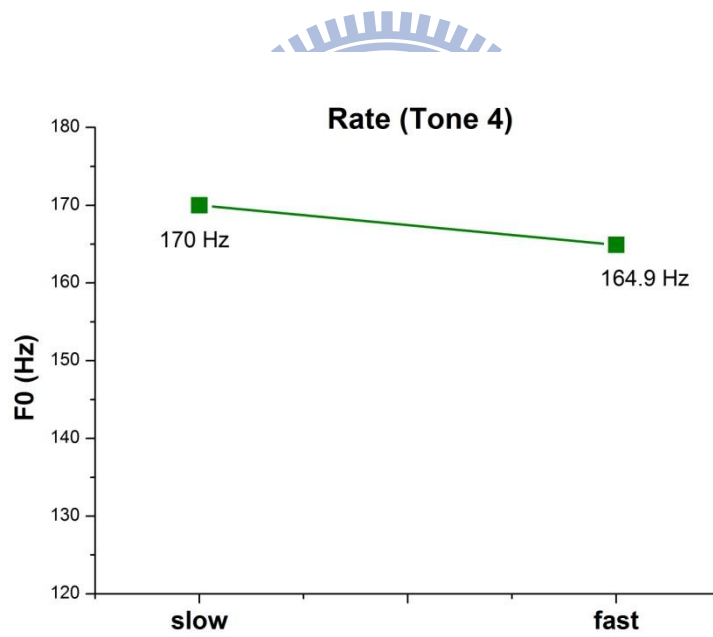


Figure 34. The main effect of Rate on 100 ms F0 in Tone 4

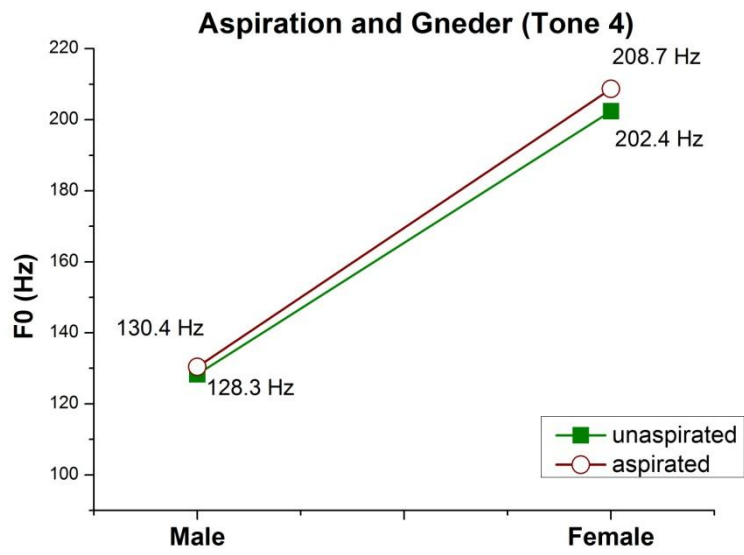


Figure 35. The interaction between Aspiration and Gender on 100 ms F0 in Tone 4

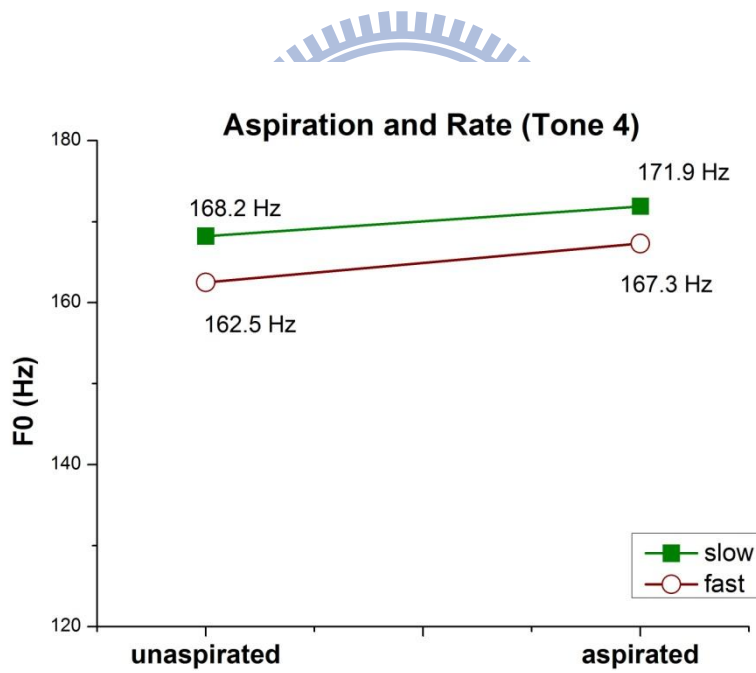


Figure 36. The interaction between Aspiration and Rate on 100 ms F0 in Tone 4

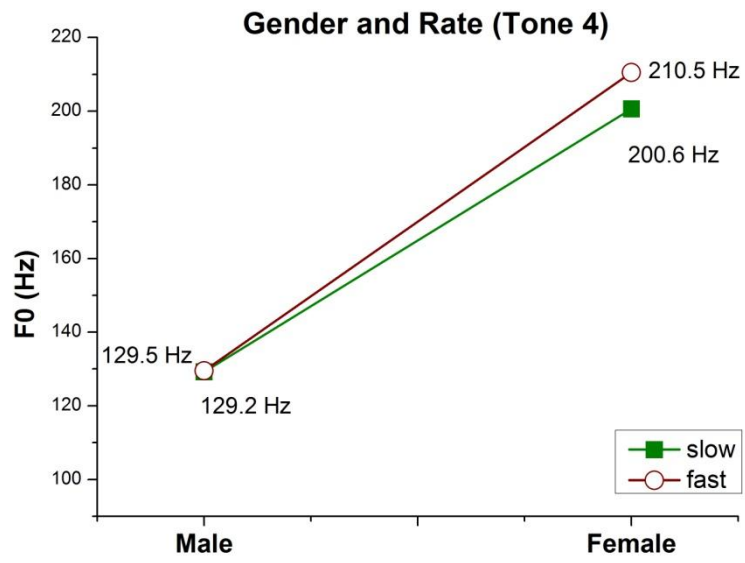


Figure 37. The interaction between Gender and Rate on 100 ms F0 in Tone 4



Chapter V Discussion

5.1 Summary of results

The present study tested the perturbation effect of prevocalic aspirated stops on F0 of following tones in Mandarin. Twenty participants were recorded producing 40 minimal pairs of CV syllables with the consonants contrasting in aspiration. These minimal pairs are embedded in the carrier sentence: *Wo3 nian4 __ zhe4 ge4 zi4* ('I read this word __'). The measurements include VOT of consonants and F0 of each vowel (normalized F0 and the onset F0).

Results showed that VOT is longer in aspirated stops than in unaspirated stops which agree with the results of Lisker and Hombert (1964). Unsurprisingly, the VOT duration is longer in slow speaking rate than that in fast speaking rate. The result of VOT duration is not significant between genders, unlike Lai's (2004) result. The discussion will be illustrated in the subsequent section.

Results from normalized F0 indicated that F0 is higher after aspirated prevocalic stops. The results in the present study echo the results in Cantonese (Zee, 1980), and in Taiwanese (Lai, 2004). In addition, F0 was found to be higher in fast speaking rate as well as in the female speakers. The interaction between Aspiration and Gender is significant but only in tone 1 and tone 3. Moreover, it is significant in the interaction between Aspiration and Rate. However, results from the onset F0 did not show a clear pattern as the results from normalized F0. It might be caused by the measuring difficulty at the onset of voicing in vowels.

In order to explain the aspiration perturbation more clearly, the following areas are

addressed as: VOT and F0. Some parameters controlling F0, such as aerodynamic and laryngeal features, will be illustrated as well.

5.2 VOT

It was found that VOT is longer in aspirated stops than in unaspirated stops for the different phonetic characteristics, as well in the slow speaking rate. The present data confirms and extends earlier findings (Diehl et al. 1980; Summerfield, 1981; Miller, 1986). This finding accords with the results of Miller (1986). He depicted that VOT between voiced and voiceless stops also changed with rate, increasing with increasing syllable duration. As speaking rate became slower, overall syllable duration became longer, VOT, especially the voiceless consonants, became longer. As speaking rate declines and overall syllable duration increases, there is a concomitant increase in the VOT value of the consonants. The conclusions of the present study show that there are no differences between genders. In comparison with the result of Lai (2004), she indicated that the VOT difference between aspirated and unaspirated stops is significantly greater for male than for female speakers. Yet, the present results are not in complete agreement with Lai's (2004) results. No significant differences between aspirated and unaspirated stops were found for males.

As for interaction, there is no significant difference between Aspiration and Gender but between Aspiration and Rate. The duration differences between fast and slow speaking rate are larger in aspirated stops than in unaspirated stops. This finding is supported by Pind (1999). Speaking rate has been shown to have much larger effect on VOT of aspirated stops than unaspirated stops in Icelandic (Pind, 1999). In the case of unaspirated stops of the present study, it seems to be stable across the effect of contextual factors such as speaking rate. One possible cause is that the size of the

phonetic category of unaspirated stops is small, ranging from 1 ms to 30 ms VOT. As a consequence, there is little acoustic space in which exemplars of the category may vary. This small range contrasts with the aspirated stops which range to approximately 130 ms (Kessinger & Blumstein, 1998).

Lai (2004) compared the correlation between aspirated F0 and VOT. The result showed that there is a significant correlation between VOT differenced and onset frequency differences. As the difference in VOT between aspirated and unaspirated onset increased, the difference in onset frequency decreased. Longer VOT means that the perturbation effect of the aspiration decreased as tone increases. In other words, long aspirated stops might have minimal effects on the following F0 onset. Unfortunately, the significant correlation is not found in the present study ($p = .054$). Tse (2005) investigated the interaction between VOT and lexical tones in Cantonese. The results showed that there are a correlation of tonal onset to VOT, but could not conclude whether or not this is due to the tonal onset or to the F0 onset. Therefore, no relation between VOT and F0 was found in the present study. Further research is needed to investigate whether VOT has any effect on the F0 of following vowel.

5.3 F0

In the methodology, there are three types of measurements (normalized F0, onset F0, and first 100 ms F0) were adopted. Tones have different durations. Tone 3 has the longest duration and the Tone 4 is the shortest one (Xu, 1997). The normalized measurement can equalize the different tone durations. F0 was extracted by every 10%. The second one measurement is onset F0. If aspiration has the rising or lowering effect, onset F0 will be affected most. The third one is first 100 ms F0. F0 was extracted every 10 ms and there were 10 points recorded. There are three types of

measurement adopted in the present study; however, only the first 100 ms was discussed. Although normalized F0 reduces the duration differences among tones, the perturbation effect of aspiration decrease by the normalization. On the other side, the onset F0 has the most obvious perturbation effect but how to define the onset of F0 raise other issue. Hombert et al. (1979) indicated the voiceless stops may still affect F0 at least 100 ms after vowel onset. That is why first 100 ms was adopted in this paper. The discussion was illustrated as below.

Aspiration

There are various parameters controlling F0 which are affected by aspirated and unaspirated distinction. These parameters could explain why aspiration distinctions have different effects on the F0 of the following vowels. The fact is that the glottis is widely open upon the release of an aspirated stop, as opposite to a more closed position for an unaspirated stop (Hombert, 1978). The more the glottis opened, the more airflow passed. According to the flow rate calculation, it is often estimated by determining the velocity at which flow is through a given cross-section area. The flow equation is: $Q = VA$ (where Q = volumetric flow, V =flow velocity, and A = cross-sectional area).

On the other hand, Müller (1987) indicated that an increase in subglottal pressure leads to an increase in frequency. Chomsky and Halle (1968) also suggested that aspirated stops are produced with heightened subglottal pressure in contrast to unaspirated stops which would have normal subglottal pressure. Previous studies also illustrated that sound required a high rate of air flow; fricative and aspirated stops tend to be produced with a separated opening gestures of glottis (Löfqvist et al., 1984; Löfqvist and Yoshioka, 1980; Yoshioka, 1981, 1982). Subtelny et al. (1969) indicated

that higher airflow rate after the aspirated versus the unaspirated series. According to the Bernoulli Effect, when the speed of flow through a fluid increases, the pressure decreases. Higher airflow rate passes through the glottis, the vocal folds vibration increases due to the decrease of pressure. Then higher airflow raises F0. Based on Bernoulli Effect, it might be suspected that F0 is higher after aspirated stops.

In addition to glottis opening, the stiffness of vocal folds is distinguished in aspiration. Increasing the stiffness of the vocal folds makes the coupling between upper and lower edges of the vocal fold larger. Stiffening of the vocal folds effects glottal vibration, regardless of the size of the glottal aperture. Halle and Stevens (1971) indicated that the voiceless aspirated stop has spread glottis and delayed VOT in initial prestressed position, and tend to increase the F0 of an adjacent vowel, as the feature [+stiff] would suggest.

The average F0 is the lowest for vowels following lax stops, higher following tense stops, and highest following aspirated stops (Kim et al., 2002). Further investigation indicated that Korean aspirated and tense stops correspond to the aspirated and unaspirated stops in Mandarin (Reetz and Jongman, 2009). What is more, Hirose (1974) examined the actions of intrinsic laryngeal muscles in production of Korean stop consonants by electromyography (EMG). The results indicated that the aspirated stop appeared to be characterized by the suppression of all the adductor muscles (vocalis, cricothyroid muscle, interarytenoid muscle, lateral cricoarytenoid, etc.) of the larynx immediately preceding the articulatory release. This suppression was always followed by a steep increase in activity which seemed to correspond to the rapid closure of the glottis after stops release. In the tense stops, the pattern of vocalis and lateral cricoarytenoid activity were most characterized. Both muscle, vocalis

particularly, showed a marked increase in activity before the stop release, which presumably results in an increase of inner tension of the vocal folds and constriction of the glottis during or immediately after the articulatory closure.

Results in the first 100 ms F0 data indicated that the aspiration at the initial position has the rising effect on F0 and the results conform to our hypothesis that F0 is higher after aspirated stops.

Aspiration and Rate

The main effect of speaking rate was significant. F0 was higher at fast speaking rate in Tone 1 and Tone 2 which is supported by Dromey & Ramig (1998). However, the main effect of speaking rate was also significantly higher at slow speaking rate in Tone 3 and Tone 4. It is quite strange to see the contradictory results. If there exists contradictory conditions, it should be that Tone 1 and Tone 4 have the same result and Tone 2 and Tone 3 for their onset tonal contour range. From the present results, it still shows that different rates indeed have influence upon F0 which might help to clarify the conflicting findings of previous researches. The different effect of aspiration of previous studies may come from the non-well controlled speaking rate.

In addition, according to flow equation, the fast speaking rate has a higher airflow. Under the same cross-area, the higher airflow increases higher airflow rate. Thus the fast airflow rate is assumed to trigger a higher F0 (Ladfoged, 1967). Moreover, Verneuil (1996) found that the increasing airflow corresponds with the increasing of F0. This finding supports the results that F0 is higher at fast speaking rate. Besides, the result could be supposed that as more muscular effort while producing fast speaking rate is expended to increase the stiffness of vocal folds. As mentioned above,

stiffening of the vocal folds affects glottal vibration. Halle and Stevens (1971) indicated that when the vocal folds are slackened, there is a decrease in the glottal vibration frequency.

How does Aspiration interaction with Rate? The interaction between Aspiration and Rate is significant in Tone 1, Tone 2, and Tone 4. Sounds or words could be misheard easier at the fast speaking rate. It was found that in the present study that the differences between aspirated stops and unaspirated stops are larger in fast speaking rate than in slow speaking rate. F0 is getting higher after aspirated stops and lower after unaspirated stops. The reason why perturbation effect is stronger at fast speaking rate might be that subjects pronounce sentence more clearly with consciousness. The conscious aspiration emphasis caused by speaker is to increase the discrimination among words in the fast speaking rate.

Generally, the results showed that the different speaking rates have different effect on F0, although the results were not consistent. F0 is significantly higher at fast speaking rate in Tone 1 and Tone 2, and higher at slow speaking rate in Tone 3 and Tone 4. Does the coarticulation make F0 different distribution at different tones? This raises another issue to investigation in future.

Aspiration and Gender

The main effect of Gender is significant. Unsurprisingly, F0 of females is higher than that of males. Some important morphological differences between the male and the female larynx have been described by many researchers (Kahane, 19198; Hirano, 1983; Tieze, 1989). Because of the vocal fold size and mass as well as vocal tract length, physiological differences between genders have different influence on F0. The

anterior-posterior dimension of the male cartilage is approximately 20% larger than that of females (Kahane, 19198). Titze (1989) reported the different larynx size between genders. The thickness of the male vocal fold is about 20% to 30% greater than that of the female as shown in Fig. 59 (a). The greater thickness creates a difficulty in elongation of the membranous portion of the vocal folds. Apart from the thickness of vocal folds, Fant (1976) also showed that the ratio of the total length of the female vocal tract to that of a male is about 0.87 or the membranous vocal fold lengths differ by 60% (Fig. 59 (b)) (Kahane, 1987). The longer vocal fold length triggers the lower F0. These differences in physiological parameters lead to different performance in acoustical performances between genders.

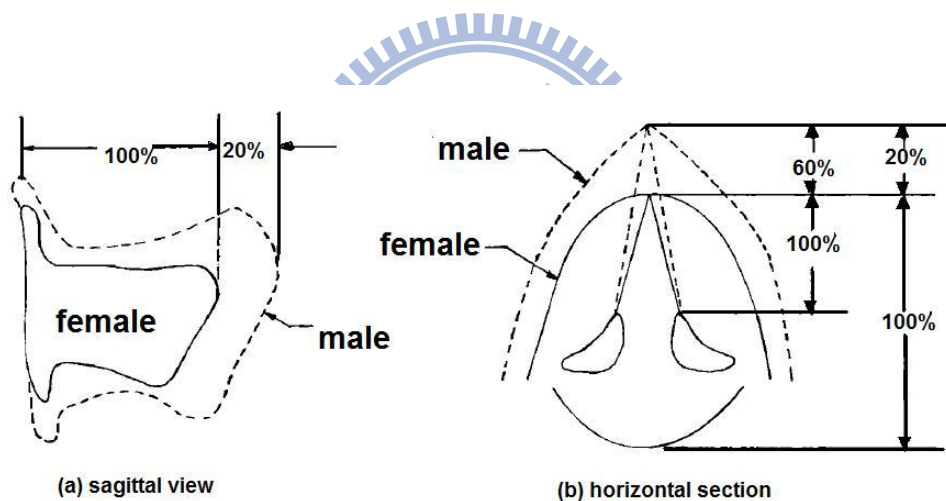


Figure 38. Male-female comparisons of dimensions of the larynx (a) Sagittal view of the thyroid cartilage and (b) horizontal section showing difference in membranous length (Adopted by Kahane, 19198)

Ajamani (1990) pointed out that the thyroid angle in the female larynx is significantly larger than that in males. It varies from 106° to 60° in males and 132° to 88° in females. Because the front parts of the vocal folds are attached at the thyroid cartilage, it could be hypothesized that wider the thyroid angle is, the more the vocal folds are strengthened. To put that differently, the female vocal folds can be strengthened more

widely than male vocal folds. As mentioned above, the wider glottis opened, the more airflow pass. Thus, the high airflow triggers the higher F0 (Ladefoged, 1967).

In addition the different physiological structure of vocal folds, the CT muscle may work differently between genders (Ximenes Filho, 2005). In mammals, the CT articulation has acquired an important function in voice modulation, permitting elongation of the vocal folds as a result of contraction of the CT muscle. The cricothyroid, which by drawing the front edges of the thyroid and cricoids cartilages closer together, lengthens and tenses the vocal cords. Ximenes Filho (2005) investigated the angulations of cricothyroid articulation in the cricoid ring, and distance between cricothyroid and cricoarytenoid articulations from 16 cadavers (9 men, 7 women). He indicated that the CT articulation angle was narrower in men than in women. The major diameter of the thyroid cartilage was wider in men. These different intrinsic physiological structures might explain why female's F0 is higher than male's.

The interaction between Aspiration and Gender was significant across four tones. F0 difference between aspirated and unaspirated stops is significantly greater in female when compared to male participants. The raising effect seems greater in female speakers as well. The physiological difference between females and males could explain the different perturbation effect of aspiration. A good illustration of this is the thickness of vocal folds. The thicker the vocal folds are, the more effort will be exhausted to make the vibration. Simply stated, based on the same subglottal pressure velocity (supposed the velocity of subglottal pressure is the same when males and females pronounce aspirated stops), the thicker the vocal folds are, the harder they vibrate. Furthermore, the range of F0 value between different genders is different. F0

range of females is larger than that of males (Iseli et al., 2006). Thus the F0 differences of aspirated and unaspirated stops could be expanded more because of the larger female F0 range.

Compared with Xu and Xu (2003), the result, opposite to the result of Lai's (2004) and the results of present study, indicated that F0 was lower after the aspirated stops. There are only 7 female participants in Xu and Xu's experiment. Based on the significant interaction between Aspiration and Gender in the present study, it can be speculated that there exists a possible gender influence in the results if researchers did balance the number of each gender. Interestingly, the female results between Xu and Xu and the present study are different either. F0 is higher after the unaspirated stop (Xu and Xu, 2003) which is opposite to that of present results. F0 is about 20-30 Hz higher after unaspirated stops in Xu and Xu (2003). Somehow the F0 differences between aspirated and unaspirated stop are too large. Furthermore, only alveolar stops were used as stimuli in Xu and Xu's experiment. There are 3 places of articulation and four vowels were adopted in the present study. It might cause the different results with these dissimilarities.

5.4 Further research

The first 100 ms F0 of aspirated and unaspirated is also compared. F0 separated by 10 points (first 100 ms) of aspirated and unaspirated stops were drawn pair-wise (Fig. 39 and Fig. 40).

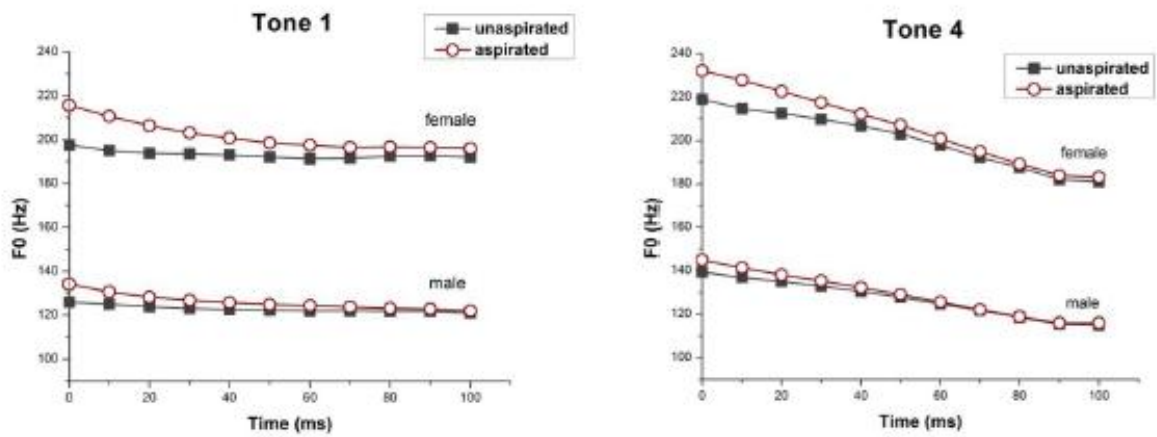


Figure 39. First 100 ms F0 of aspirated and unaspirated stops between male and female (high tones)

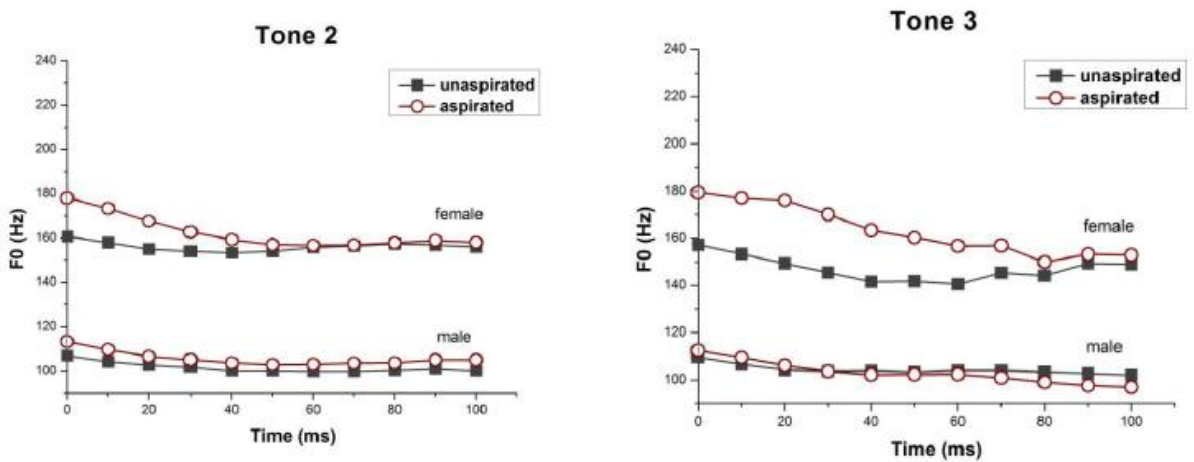


Figure 40. First 100 ms F0 of aspirated and unaspirated stops between male and female (low tones)

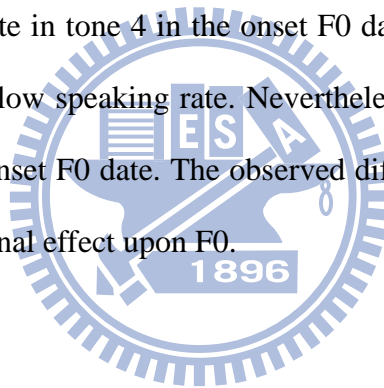
Generally speaking, F0 of aspirated stops is higher than that of unaspirated stops although somehow both of aspirated and unaspirated stop F0 overlap or even F0 of unaspirated stops is higher than that of aspirated stops (the male one in Tone 3).

Form Fig. 39 and 40, the differences between aspirated and unaspirated stops are larger in low tones especially in female. The rising effect of aspiration seems more obvious in low tones. It might be explained by tone inventory. Every tone has itself tone range size. It could conjecture that if the rising effect of aspiration occurs in high

tones, it might cause the change of tone range size. It might not cause the tone range size changing problem in the low tones, yet.

5.5 Conclusion

Consonant aspiration in Mandarin has a significant effect on F0 of the following vowels. The aspirated stop is associated with higher F0 and its unaspirated cognate with lower F0. This effect is attributable to aerodynamic and physiological factors together. Besides, the other two modulating factors—gender difference and speech rate—help to provide more explanation why the results of previous researches are not in agreement. In addition, there are some different distributions in the interaction between Aspiration and Rate in tone 4 in the onset F0 data. Mostly, F0 is higher at a fast speaking rate than at slow speaking rate. Nevertheless, F0 of tone 4 is higher at slow speaking rate in the onset F0 data. The observed differences between tones lead us to further research the tonal effect upon F0.



References

- Abramson, A. S. (1979). The coarticulation of tones: An acoustic study of Thai. *Studies in Tai and Mon-Khmer phonetics and phonology in honour of Eugenie JA Henderson*, 1-9.
- Ajmani, M. L. (1990). A metrical study of the laryngeal skeleton in adult Nigerians. *Journal of anatomy*, 171, 187.
- Bauer, R. S., & Benedict, P. K. (1997). *Modern Cantonese Phonology*: Mouton de Gruyter.
- Chao, Y. R. (1967). Contrastive aspects of the Wu dialects. *Language*, 43(1), 92-101.
- Chao, Y. R., & Pian, R. C. (1948). *Mandarin primer* (Vol. 26): Harvard University Press Cambridge, MA.
- Cho, T., Jun, S. A., & Ladefoged, P. (2002). Acoustic and aerodynamic correlates of Korean stops and fricatives. *Journal of Phonetics*, 30(2), 193-228.
- Clark, J. E., Yallop, C., & Fletcher, J. (2007). *An introduction to phonetics and phonology*: Wiley-Blackwell.
- Clumeck, H. (1976). Patterns of soft palate movements in six languages. *Journal of Phonetics*, 4(4), 337-351.
- Dart, S. N. (1987). An aerodynamic study of Korean stop consonants: Measurements and modeling. *The Journal of the Acoustical Society of America*, 81, 138.
- Dromey, C., & Ramig, L. O. (1998). Intentional changes in sound pressure level and rate: Their impact on measures of respiration, phonation, and articulation. *Journal of Speech, Language, and Hearing Research*, 41(5), 1003.
- Ewan, W. G. (1979). Laryngeal behavior in speech. *Report of Phonology laboratory*, 3, 1-93.
- Fant, G. (1976). Vocal tract energy functions and non-uniform scaling. *J. Acoust. Soc. Japan*, Dec, 1-11.
- Francis, A. L., Ciocca, V., Wong, V. K. M., & Chan, J. K. L. (2006). Is fundamental frequency a cue to aspiration in initial stops? *The Journal of the Acoustical Society of America*, 120, 2884.
- Gandour, J. (1974). Consonant types and tone in Siamese. *Journal of Phonetics*, 2, 337-350.
- Gandour, J., Potisuk, S., & Dechongkit, S. (1994). Tonal coarticulation in Thai. *Journal of Phonetics*, 22(4), 477-492.
- Gandour, J., & Weinberg, B. (1980). On the relationship between vowel height and fundamental frequency: evidence from esophageal speech. *Phonetica*, 37(5-6), 344-354.
- Gay, T., Hirose, H., Strome, M., & Sawashima, M. (1972). Electromyography of the

- intrinsic laryngeal muscles during phonation. *Ann Otol Rhinol Laryngol*, 81(3), 401-409.
- Gay, T., Ushijima, T., Hirose, H., & Cooper, F. S. (1974). Effect of speaking rate on labial consonant-vowel articulation. *Journal of Phonetics*, 2(47-63).
- Gracco, V. L. (1994). Some organizational characteristics of speech movement control. *Journal of speech and hearing research*, 37(1), 4.
- Hanson, H. M., & Stevens, K. N. (2002). A quasiarticulatory approach to controlling acoustic source parameters in a Klatt-type formant synthesizer using HLsyn. *The Journal of the Acoustical Society of America*, 112, 1158.
- Hardcastle, W. J., & Marchal, A. (1990). *Speech production and speech modelling*: Springer.
- Hirano, M. (1981). Structure of the vocal fold in normal and disease states: anatomical and physical studies. *ASHA Rep*, 11, 11-30.
- Hirano, M., Kurita, S., & Nakashima, T. (1983). Growth, development and aging of human vocal folds. *Vocal fold physiology: Contemporary research and clinical issues*, 22-43.
- Hirano, M., Kurita, S., & Sakaguchi, S. (1989). Ageing of the vibratory tissue of human vocal folds. *Acta oto-laryngologica*, 107(5), 428-433.
- Hirano, M., & Ohala, J. (1969). Use of hooked-wire electrodes for electromyography of the intrinsic laryngeal muscles. *J Speech Hear Res*, 12(2), 362-373.
- Hirose, H. (1997). Investigating the physiology of laryngeal structures. In J. L. William J. Hardcastle, Fiona E. Gibbon (Ed.), *The Handbook of Phonetic Sciences, 2nd Edition*: Wiley-Blackwell.
- Hirose, H., Lee, C. Y., & Ushijima, T. (1974). Laryngeal control in Korean stop production. *Journal of Phonetics*, 2(2), 145–152.
- Hombert, J. M. (1975). *Towards a Theory of Tonogenesis: An Empirical, Physiologically and Perceptually Based Account of the Development of Tonal Contrasts in Languages*: University of California, Berkeley.
- Hombert, J. M. (1978). Consonant types, vowel quality, and tone. *Tone: a linguistic survey*, 77-111.
- Hombert, J. M., & Ladefoged, P. (1977). The effect of aspiration on the fundamental frequency of the following vowel. *WORKING PAPERS IN PHONETICS* 36, 33.
- Hombert, J. M., Ohala, J. J., & Ewan, W. G. (1979). Phonetic explanations for the development of tones. *Language*, 55(1), 37-58.
- Honda, K. (1983). Relationship between pitch control and vowel articulation. *Vocal fold physiology*, 286-297.
- Honda, K. (2004). *Physiological factors causing tonal characteristics of speech: from*

global to local prosody.

- Honda, K., Hirai, H., Masaki, S., & Shimada, Y. (1999). Role of vertical larynx movement and cervical lordosis in F0 control. *Language and Speech*, 42(4), 401.
- Hoole, P., Honda, K., Murano, E., Fuchs, S., & Pape, D. (2004). *Cricothyroid activity in consonant voicing and vowel intrinsic pitch.*
- House, A. S., & Fairbanks, G. (1953). The influence of consonant environment upon the secondary acoustical characteristics of vowels. *Journal of the Acoustical Society of America*.
- Iseli, M., Shue, Y. L., & Alwan, A. (2006). *Age-and gender-dependent analysis of voice source characteristics.*
- Jeel, V. (1975). An investigation of the fundamental frequency of vowels after various Danish consonants, in particular stop consonants. *Annual Report of the Institute of Phonetics, University of Copenhagen*, 9, 191–211.
- Kagaya, R. (1974). A fiberscopic and acoustic study of the Korean stops, affricates and fricatives. *Journal of Phonetics*, 2, 161-180.
- Kahane, J. C. (19198). A morphological study of the human prepubertal and pubertal larynx. *American Journal of Anatomy*, 151(1), 11-19.
- Kanade, T., Tian, Y., & Cohn, J. F. (2000). Comprehensive database for facial expression analysis. *fg*, 46.
- Keating, P. A. (1984). Physiological effects on stop consonant voicing. *ucla working papers*, 29.
- Kelso, J. A. S., Vatikiotis Bateson, E., Saltzman, E. L., & Kay, B. (1985). A qualitative dynamic analysis of reiterant speech production: Phase portraits, kinematics, and dynamic modeling. *The Journal of the Acoustical Society of America*, 77, 266.
- Kenstowicz, M., & Park, C. (2006). Laryngeal features and tone in Kyungsang Korean: A phonetic study. *Studies in Phonetics, Phonology, and Morphology*, 12, 247–264.
- Kim, C. W. (1968). The vowel system of Korean. *Language*, 44(3), 516-527.
- Kim, M. R., Beddor, P. S., & Horrocks, J. (2002). The contribution of consonantal and vocalic information to the perception of Korean initial stops. *Journal of Phonetics*, 30(1), 77-100.
- Kessinger, R. H., & Blumstein, S. E. (1998). Effects of speaking rate on voice-onset time and vowel production: Some implications for perception studies. *Journal of Phonetics*, 26(2), 117-128.
- Koenig, L. L. (2000). Laryngeal factors in voiceless consonant production in men, women, and 5-year-olds. *Journal of Speech, Language, and Hearing Research*,

- 43(5), 1211.
- Krakow, R. A. (1999). Physiological organization of syllables: a review. *Journal of Phonetics*, 27(1), 23-54.
- Ladefoged, P. (1967). *Three areas of experimental phonetics*: Oxford University Press.
- Lai Y. (2004). The effect of aspiration on the fundamental frequency of the following tone in Taiwanese. MA thesis. University Kansas.
- Lehiste, I., & Peterson, G. E. (1961). Some basic considerations in the analysis of intonation. *The Journal of the Acoustical Society of America*, 33, 419.
- Li, C. N., & Thompson, S. A. (1989). *Mandarin Chinese: A functional reference grammar*: Univ of California Pr.
- Li, P. (2006). *The Handbook of East Asian Psycholinguistics: Chinese*: Cambridge Univ Pr.
- Lieberman, P. (1967). *Intonation, perception, and language*: MIT Press Cambridge, MA.
- Liu, S., & Samuel, A. G. (2004). Perception of Mandarin lexical tones when F0 information is neutralized. *Language and Speech*, 47(2), 109.
- Lofqvist, A. (1975). Intrinsic and extrinsic F0 variations in Swedish tonal accents. *Phonetica* 31, 228-247.
- Lofqvist, A., McGarr, N. S., & Honda, K. (1984). Laryngeal muscles and articulatory control. *The Journal of the Acoustical Society of America*, 76, 951.
- Lofqvist, A., & Yoshioka, H. (1980). Laryngeal activity in Swedish obstruent clusters. *Journal of the Acoustical Society of America*, 68(3), 792-799.
- Massaro, D. W., Cohen, M. M., & Tseng, C. Y. (1985). The evaluation and integration of pitch height and pitch contour in lexical tone perception in Mandarin Chinese. *Journal of Chinese Linguistics*, 13(2), 267-290.
- Mohr, B. (1971). Intrinsic variations in the speech signal *Phonetica*, 23(2), 65-93.
- Moore, C. B., & Jongman, A. (1997). Speaker normalization in the perception of Mandarin Chinese tones. *The Journal of the Acoustical Society of America*, 102, 1864.
- Newman, S. R., Butler, J., Hammond, E. H., & Gray, S. D. (2000). Preliminary report on hormone receptors in the human vocal fold*. *Journal of Voice*, 14(1), 72-81.
- Ohala, J., & Ewan, W. G. (1979). Phonetic explanations for the development of tones. *Language*, 55(1).
- Ohala, J. J. (1990a). *The phonetics and phonology of aspects of assimilation*.
- Ohala, J. J. (1990b). Respiratory activity in speech. *Speech production and speech modelling*, 23-53.

- Ohala, M., & Ohala, J. (1972). The problem of aspiration in Hindi phonetics. *Annual Bulletin, Research Institute of Logopedics and Phoniatics, University of Tokyo*, 6, 39–46.
- Ohde, R. N. (1984). Fundamental frequency as an acoustic correlate of stop consonant voicing. *The Journal of the Acoustical Society of America*, 75, 224.
- Ostry, D. J., & Munhall, K. G. (1985). Control of rate and duration of speech movements. *Journal of the Acoustical Society of America*, 77(2), 640-648.
- Peng, S. (1997). Production and perception of Taiwanese tones in different tonal and prosodic contexts. *Journal of Phonetics*, 25(3), 371-400.
- Reetz, H., & Jongman, A. (2008). *Phonetics: transcription, production, acoustics, and perception*: Wiley-Blackwell.
- Rosenblum, L. D., & Fowler, C. A. (1991). Audiovisual investigation of the loudness-effort effect for speech and nonspeech events. *Journal of Experimental Psychology: Human Perception and Performance*, 17(4), 976-985.
- Sawashima, M., & Hirose, H. (1980). Ann. Bull. RILP, No. 14, 29-51 (1980) LARYNGEAL GESTURES IN SPEECH PRODUCTION. *Annual bulletin*, 29.
- Sawashima, M., & Hirose, H. (1983). Laryngeal gestures in speech production. *The production of speech*, 11-38.
- Schiavetti, N., & Metz, D. E. (1997). Stuttering and the measurement of speech naturalness. *Nature and treatment of stuttering: New directions*, 398–412.
- Shi, F. (1998). The Influence of Aspiration on Tones. *Journal of Chinese Linguistics*, 26, 126-145.
- Stevens, K. N. (1977). Physics of laryngeal behavior and larynx modes. *Phonetica*, 34(4), 264.
- Strik, H., & Boves, L. (1992). Control of fundamental frequency, intensity and voice quality in speech. *Journal of Phonetics*, 20(1), 15-25.
- Subtelny, J. D., Worth, J. H., & Sakuda, M. (1966). Intraoral pressure and rate of flow during speech. *Journal of speech and hearing research*, 9(4), 498.
- Titze, I. R. (1989). Physiologic and acoustic differences between male and female voices. *The Journal of the Acoustical Society of America*, 85, 1699.
- Titze, I. R. (1992). Phonation threshold pressure: A missing link in glottal aerodynamics. *J Acoust Soc Am*, 91(5), 2926-2935.
- Tsao, Y. C., & Weismer, G. (1997). Interspeaker variation in habitual speaking rate: evidence for a neuromuscular component. *Journal of Speech, Language, and Hearing Research*, 40(4), 858.
- Tse, H. B. (2005). *The Phonetics of VOT and Tone Interaction in Cantonese*.

University of Chicago, Chicago.

- Umeda, N. (1981). Influence of segmental factors on fundamental frequency in fluent speech. *The Journal of the Acoustical Society of America*, 70, 350.
- Vilkman, E., Aaltonen, O., Raimo, I., Arajarvi, P., & Oksanen, H. (1989). Articulatory hyoid-laryngeal changes vs. cricothyroid muscle activity in the control of intrinsic F0 of vowels. *Journal of Phonetics*, 17, 193-203.
- Vilkman, E., Sonninen, A., Hurme, P., & Korkko, P. (1996). External laryngeal frame function in voice production revisited: a review. *Journal of Voice*, 10(1), 198-92.
- Verneuil, A., Kreiman, J., Kevorkian, K., Ye, M., Gerratt, B. R., & Berke, G. S. (1996). Effects of driving pressure and recurrent laryngeal nerve stimulation on glottic vibration in a constant pressure model. *Otolaryngology--Head and Neck Surgery*, 115(1), 15.
- Whalen, D. H., Gick, B., Kumada, M., & Honda, K. (1999). Cricothyroid activity in high and low vowels: Exploring the automaticity of intrinsic F0. *Journal of Phonetics*, 27(2), 125-142.
- Ximenes Filho, J. A., Cardoso Bohadana, S., Hiroshi Tsuji, D., Perazzio, A. F., & Ubirajara Sennes, L. (2005). Anatomy of the cricothyroid articulation: differences between men and women. *The Annals of otology, rhinology & laryngology*, 114(3), 250-252.
- Xu, C. X., & Xu, Y. (2003). Effects of consonant aspiration on Mandarin tones. *Journal of the International Phonetic Association*, 33(02), 165-181.
- Xu, Y. (1997). Contextual tonal variations in Mandarin. *Journal of Phonetics*, 25(1), 61-83.
- Xu, Y. (1999). Effects of tone and focus on the formation and alignment of f0 contours. *Journal of Phonetics*, 27(1), 55-105.
- Yoshioka, H., Lofqvist, A., & Collier, R. (1982). Laryngeal adjustments in Dutch voiceless obstruent production. *Ann. Bull. RILP*, 16, 27-35.
- Yoshioka, H., Lofqvist, A., & Hirose, H. (1981). Laryngeal adjustments in the production of consonant clusters and geminates in American English. *Journal of the Acoustical Society of America*, 70(6), 1615-1623.
- 石汝杰. (2007). 北部語的送氣分調現象. *中國語學*, 107-123.

Appendix

1. Stimuli in the methodology

POA	Vowel	Tone	Unaspirated	Gloss	Aspirated	Gloss
bilabial	i	1	pi	force	p ^h i	a batch
bilabial	i	2	pi	nose	p ^h i	the skin
bilabial	i	3	pi	comparison	p ^h i	the ruffian
bilabial	i	4	pi	have to	p ^h i	the fart
bilabial	a	1	pa	eight	p ^h a	to lie prone
bilabial	a	2	pa	pull out	p ^h a	to crawl
bilabial	a	3	pa	to hold	p ^h a	nonsense word
bilabial	a	4	pa	father	p ^h a	the handkerchief
bilabial	u	1	pu	to eat	p ^h u	to fall forward
bilabial	u	2	pu	not	p ^h u	simple or plain
bilabial	u	3	pu	to mend	p ^h u	general
bilabial	u	4	pu	clothing	p ^h u	a shop
alveolar	i	1	ti	low	t ^h i	to kick
alveolar	i	2	ti	an enemy	t ^h i	to lift
alveolar	i	3	ti	bottom	t ^h i	the body
alveolar	i	4	ti	ground	t ^h i	to shave
alveolar	a	1	ta	to build	t ^h a	to collapse
alveolar	a	2	ta	to answer	t ^h a	nonsense word
alveolar	a	3	ta	to heat	t ^h a	the tower
alveolar	a	4	ta	big or large	t ^h a	to step on
alveolar	u	1	tu	to pout	t ^h u	bald
alveolar	u	2	tu	to read	t ^h u	the picture
alveolar	u	3	tu	to gamble	t ^h u	the dust
alveolar	u	4	tu	the belly	t ^h u	the rabbit
velar	i	1	ki	nonsense word	k ^h i	nonsense word
velar	i	2	ki	nonsense word	k ^h i	nonsense word
velar	i	3	ki	nonsense word	k ^h i	nonsense word
velar	i	4	ki	nonsense word	k ^h i	nonsense word
velar	a	1	ka	the sound of squeaking	k ^h a	californium
velar	a	2	ka	gadolinium	k ^h a	nonsense word
velar	a	3	ka	nonsense	k ^h a	to get stuck

				word		
velar	a	4	ka	embarrassed	k ^h a	noise made in coughing
velar	u	1	ku	aunt	k ^h u	to cry
velar	u	2	ku	a kind of bird	k ^h u	nonsense word
velar	u	3	ku	ancient	k ^h u	bitter
velar	u	4	ku	to be solid	k ^h u	a warehouse
velar	e	1	ke	songs	k ^h e	to carve
velar	e	2	ke	to separate	k ^h e	a shell
velar	e	3	ke	a Chinese family name	k ^h e	thirst
velar	e	4	ke	each	k ^h e	a course



2. The overall results of measurements of F0

Table 9. Overall results of normalized F0 measurement by factors (aspiration, rate, gender) and their interaction

	Tone 1	Tone 2	Tone 3	Tone 4
Aspiration	* asp > un	* asp > un	* asp > un	* asp > un
Rate	* fast > slow	* fast > slow	* fast > slow	n.s
Gender	* f > m	* f > m	* f > m	* f > m
Aspiration x Gender	*	n.s	*	n.s
Gender x Rate	*	n.s	*	*
Aspiration x Rate	n.s	*	*	*
Aspiration x Gender x Rate	*	n.s	n.s	*

Table 10. Overall results of first 100 ms F0 measurement by factors (aspiration, rate, gender) and their interaction

	Tone 1	Tone 2	Tone 3	Tone 4
Aspiration	* asp > un	* asp > un	* asp > un	* asp > un
Rate	* fast > slow	* fast > slow	* slow > fast	* slow > fast
Gender	* f > m	* f > m	* f > m	* f > m
Aspiration x Gender	*	*	*	*
Gender x Rate	*	*	*	*
Aspiration x Rate	*	*	n.s	*
Aspiration x Gender x Rate	*	*	*	*

Table 11. Overall results of Onset F0 measurement by factors (aspiration, rate, gender) and their interaction

	Tone 1	Tone 2	Tone 3	Tone 4
Aspiration	* asp > un	* asp > un	* asp > un	* asp > un
Rate	* fast > slow	* fast > slow	* slow > fast	* slow > fast
Gender	* f > m	* f > m	* f > m	* f > m
Aspiration x Gender	*	*	*	*
Gender x Rate	*	*	*	*
Aspiration x Rate	*	*	n.s	*
Aspiration x Gender x Rate	*	*	*	*

3. Results of a 3-way ANOVA test on VOT

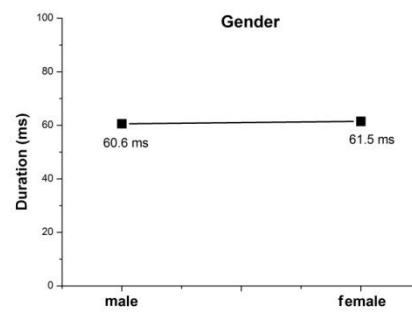
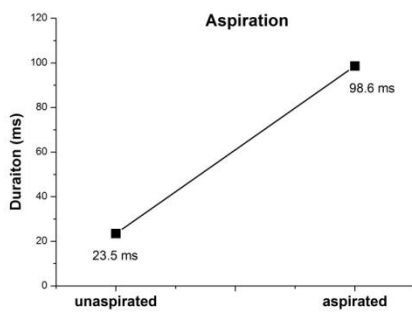
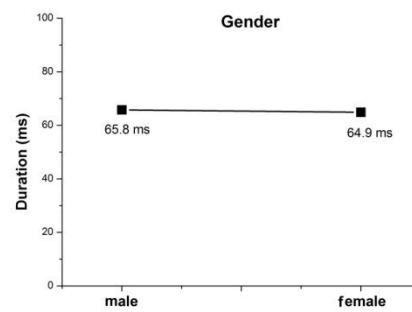
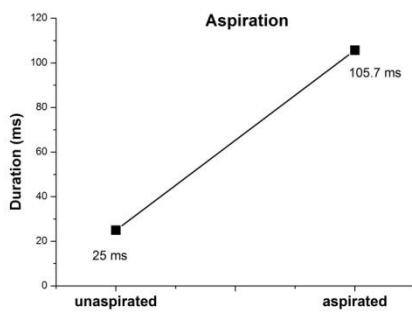
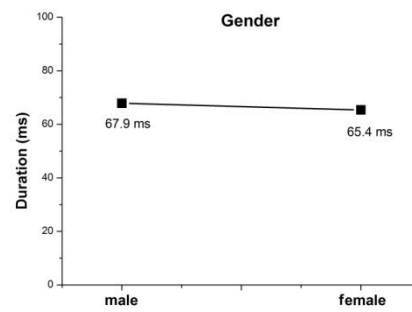
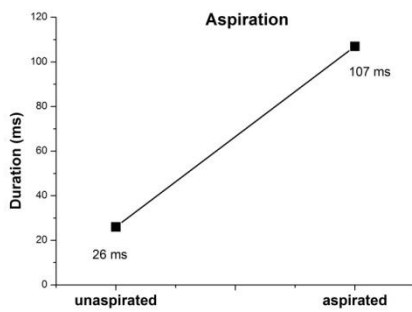
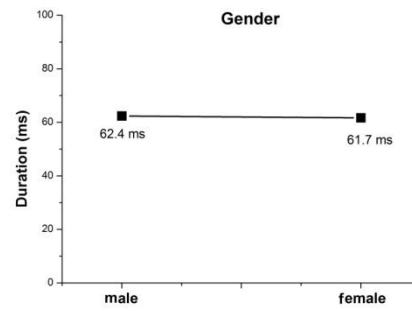
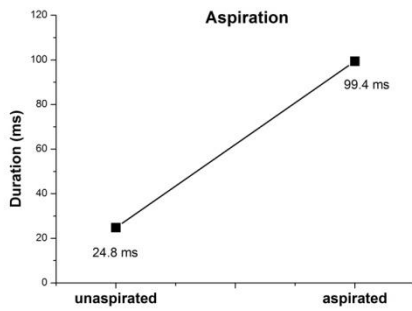


Figure 41. The main effect of Aspiration on VOT across four tones

Figure 42. The main effect of Gender on VOT across four tones

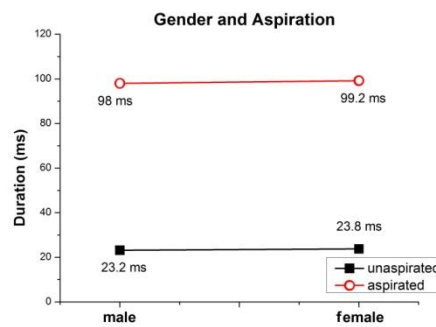
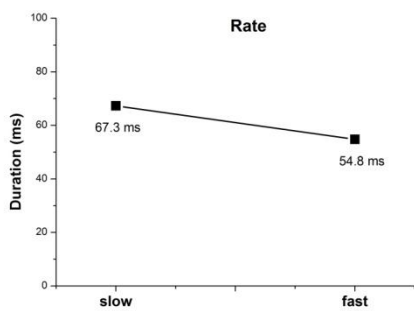
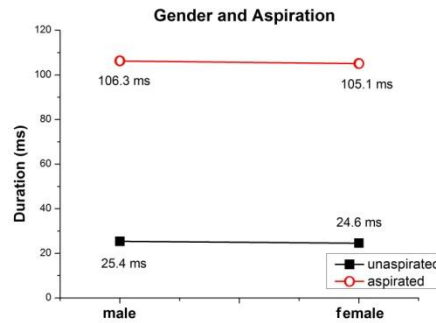
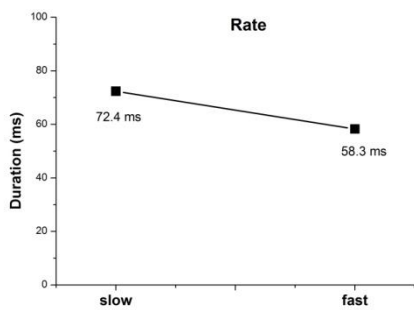
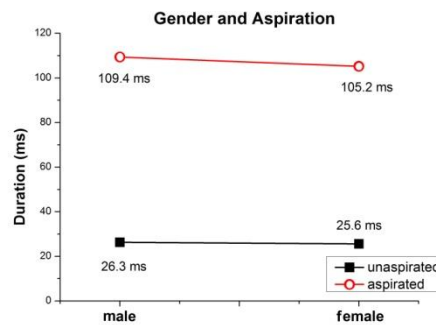
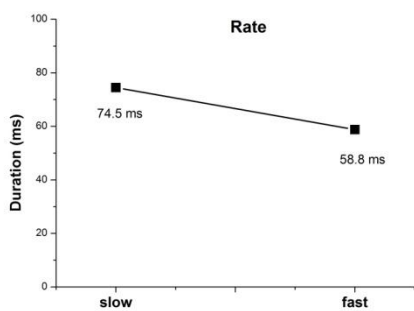
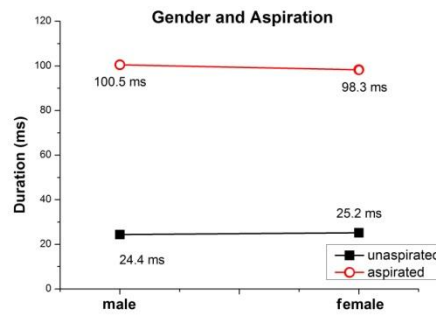
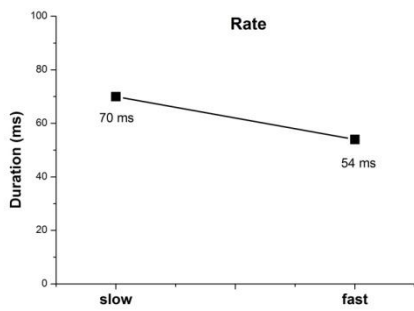


Figure 43. The main effect of Rate on VOT across four tones

Figure 44. The interaction between Aspiration and Gender on VOT across four tones

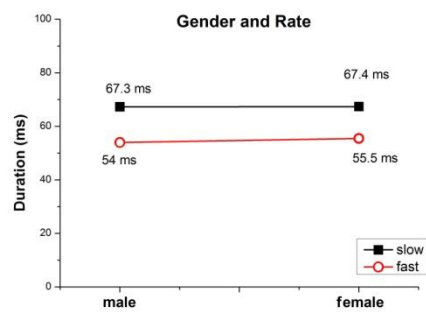
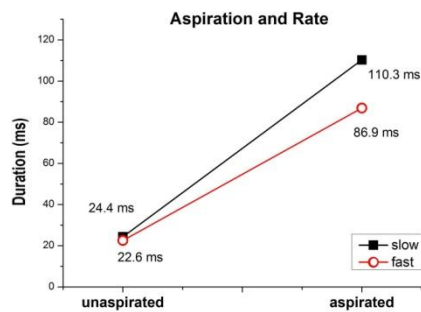
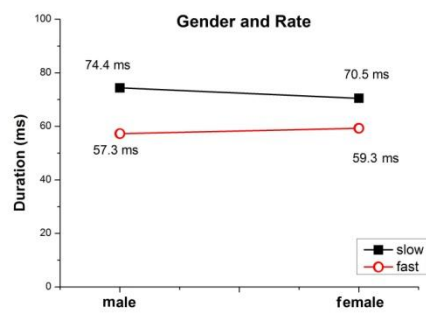
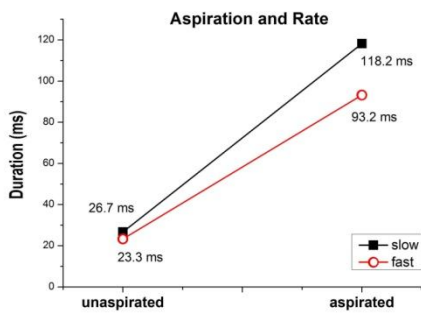
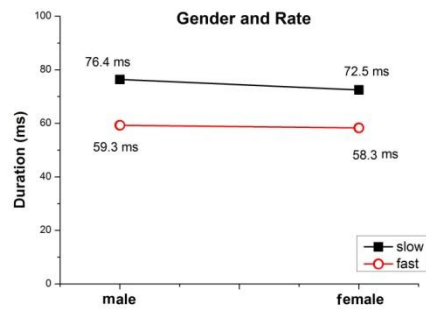
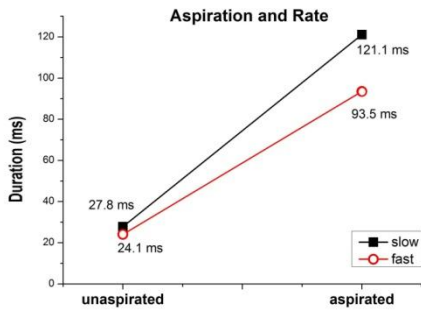
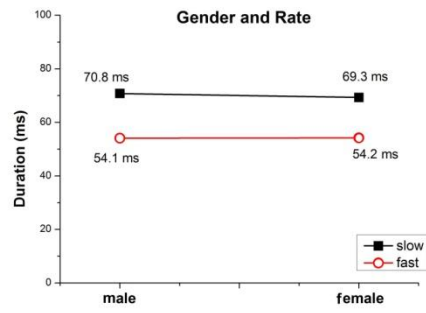
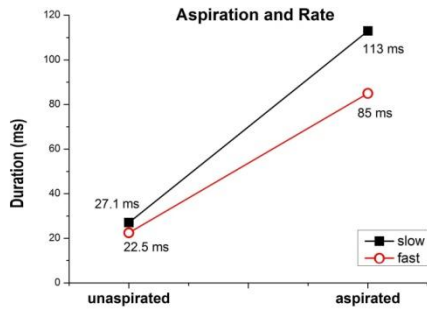


Figure 45. The interaction between Aspiration and Rate on VOT across four tones

Figure 46. The interaction between Gender and Rate on VOT across four tones

4. Results of a 3-way ANOVA test on Normalized F0

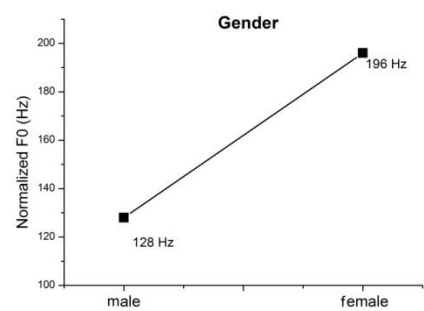
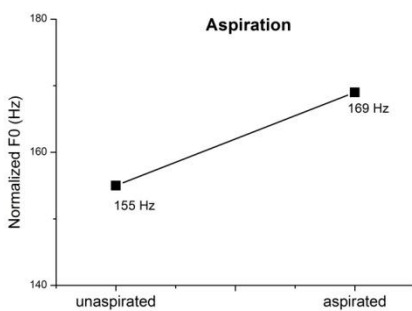
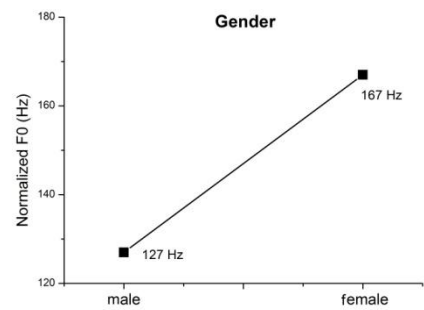
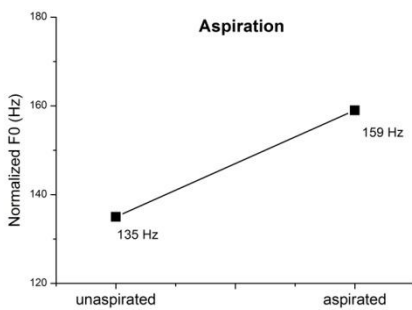
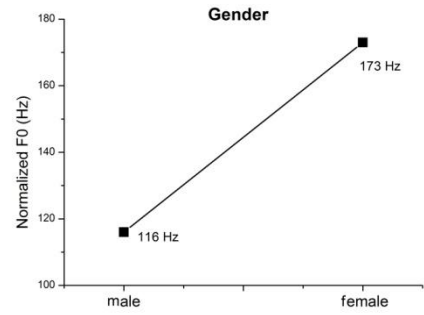
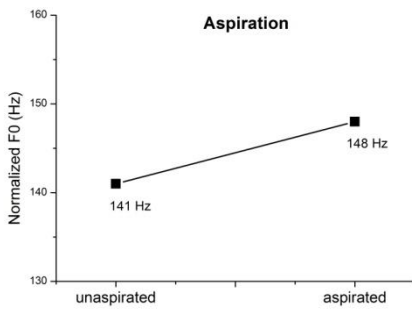
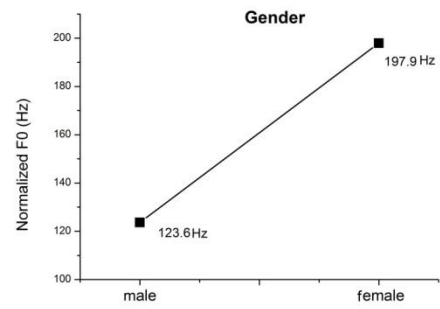
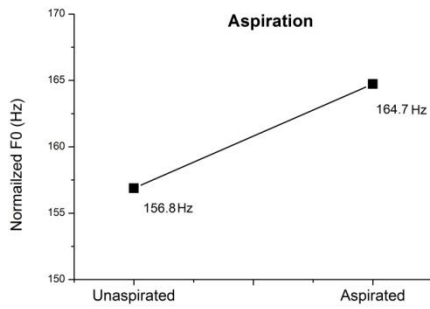


Figure 47. The main effect of Aspiration on normalized F0 across four tones

Figure 48. The main effect of Gender on normalized F0 across four tones

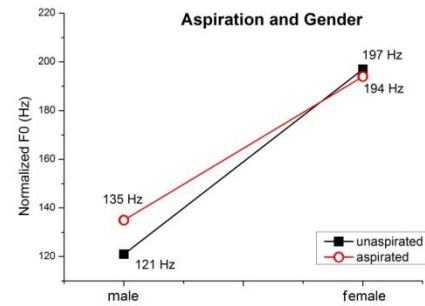
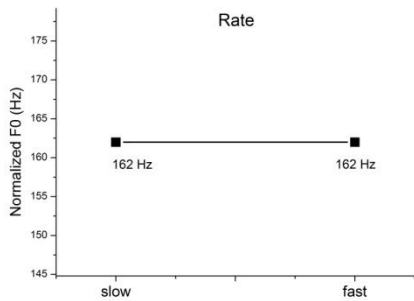
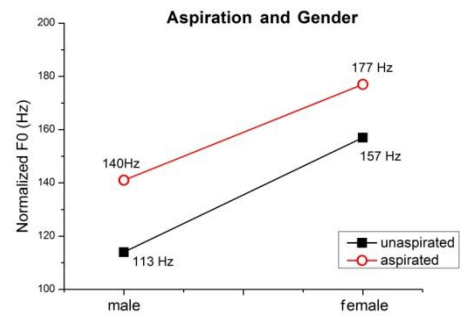
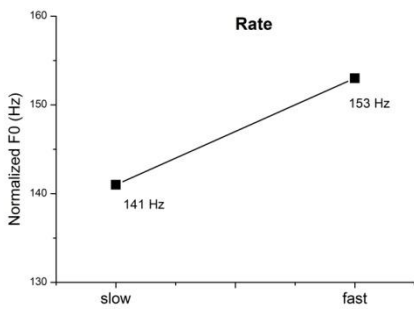
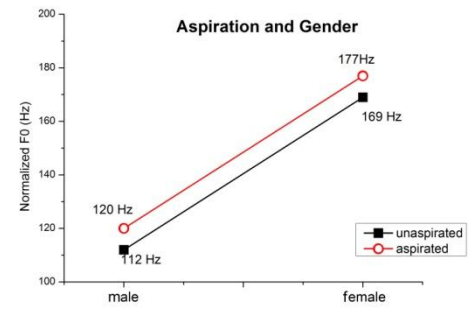
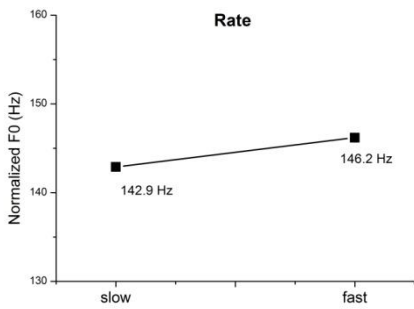
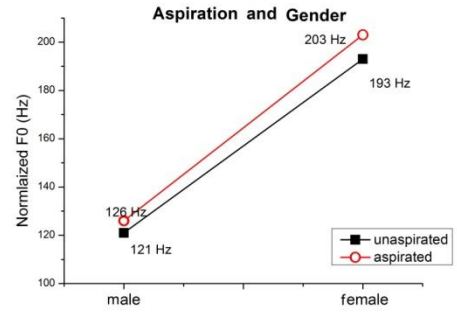
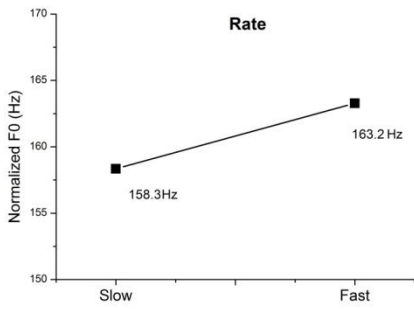


Figure 49. The main effect of Rate on normalized F0 across four tones

Figure 50. The interaction between Aspiration and Gender on normalized F0 across four tones

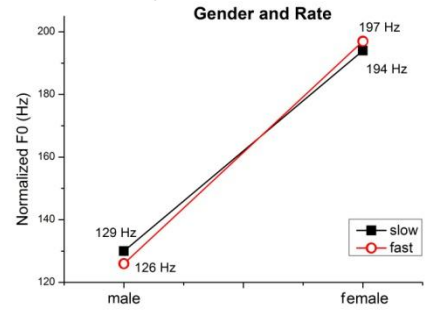
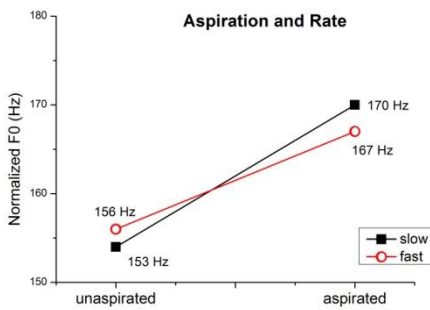
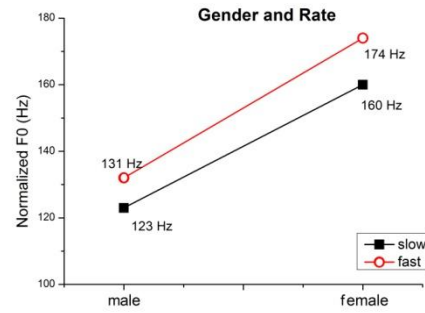
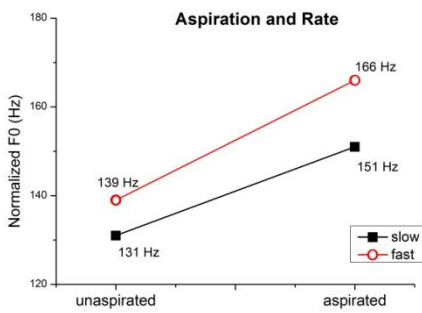
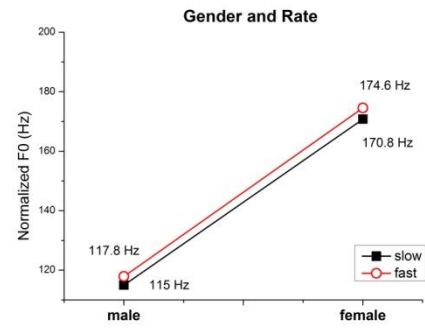
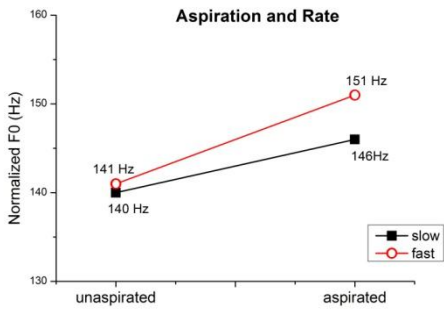
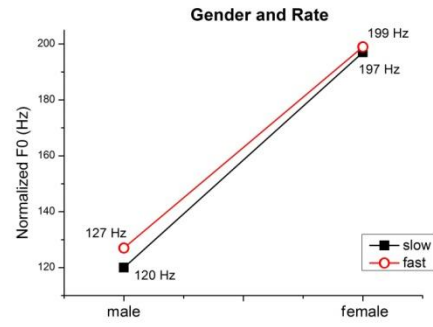
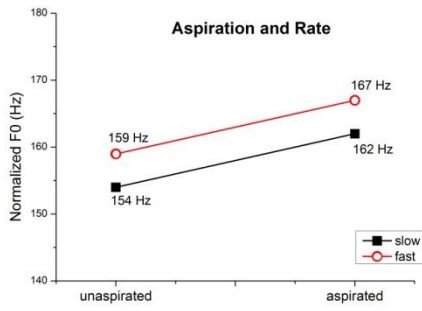


Figure 51. The interaction between Aspiration and Rate on normalized F0 across four tones

Figure 52. The interaction between Gender and Rate on normalized F0 across four tones

5. Results of a 3-way ANOVA test on Onset F0

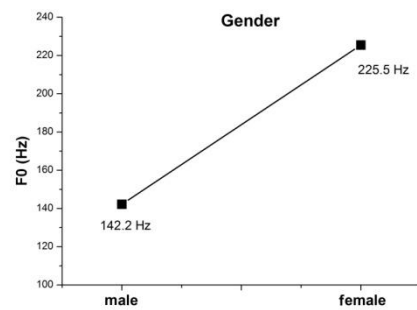
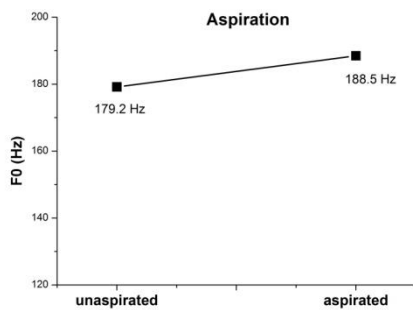
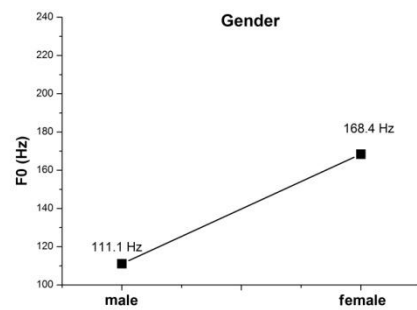
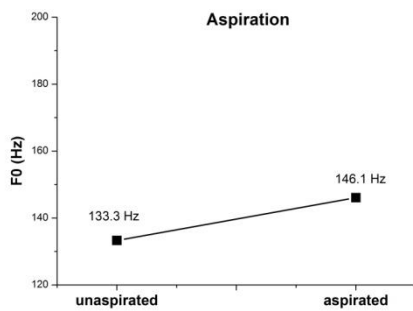
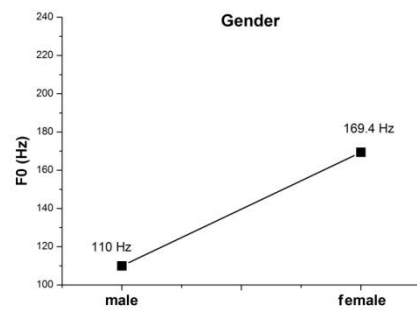
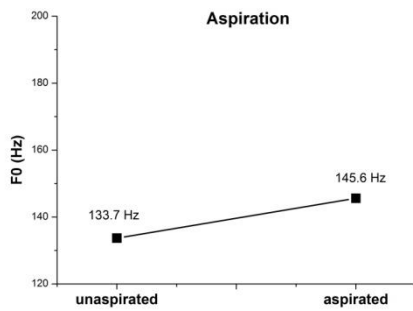
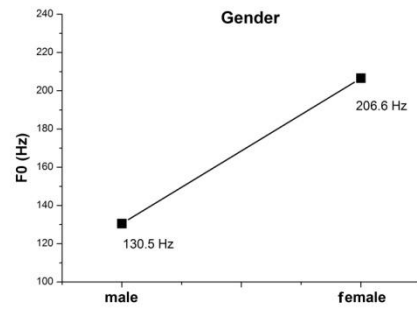
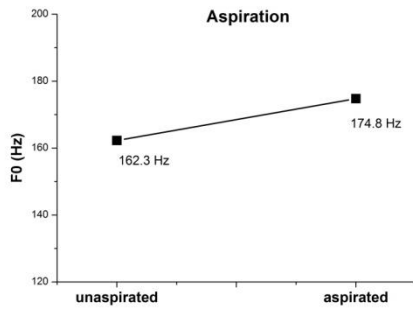


Figure 53. The main effect of Aspiration on onset F0 across four tones

Figure 54. The main effect of Gender on onset F0 across four tones

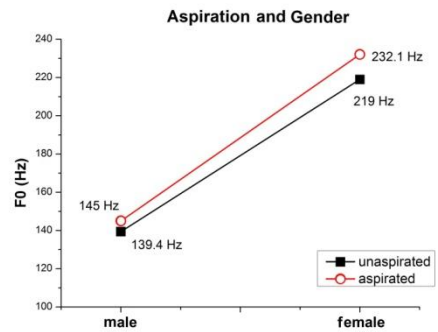
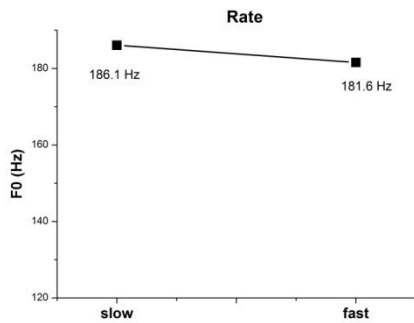
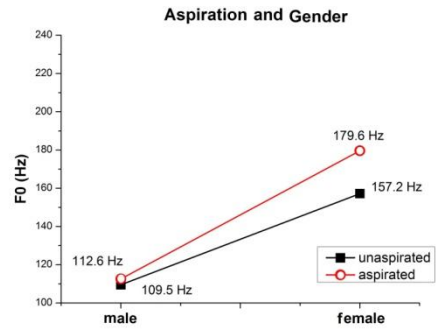
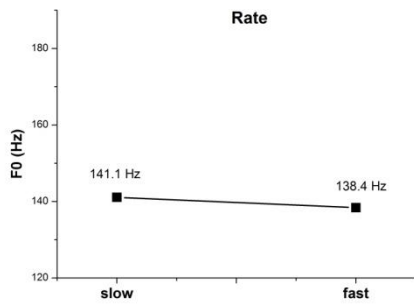
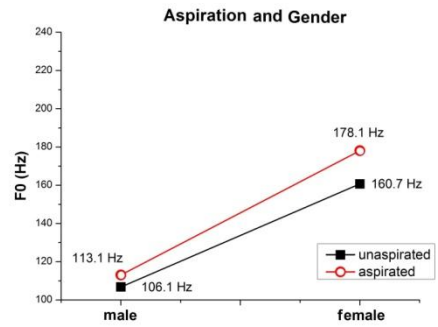
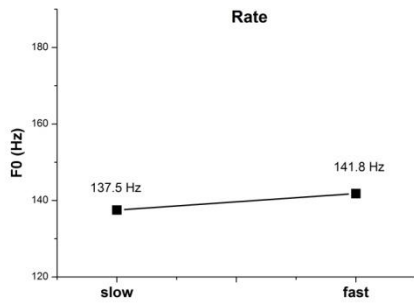
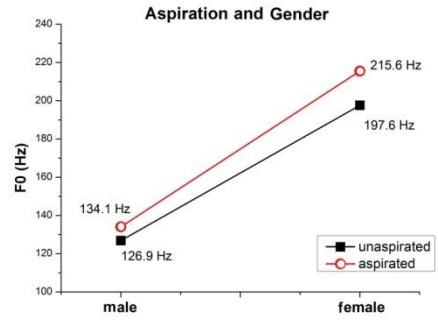
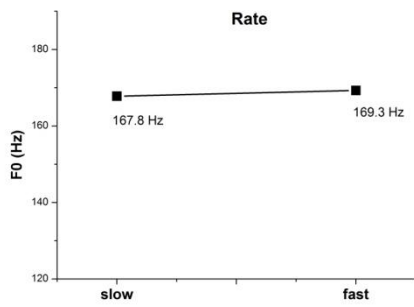


Figure 55. The main effect of Rate on onset F0 across four tones

Figure 56. The interaction between Aspiration and Gender on onset F0 across four tones

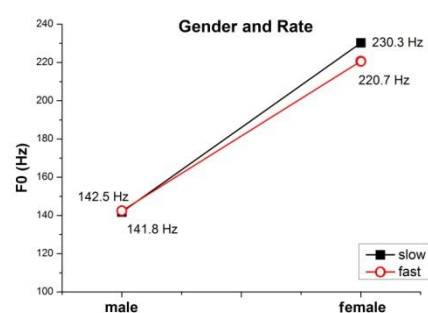
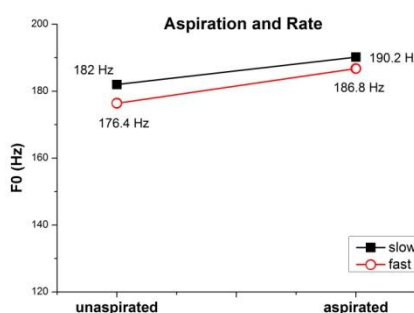
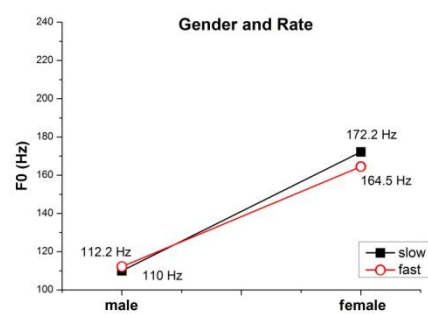
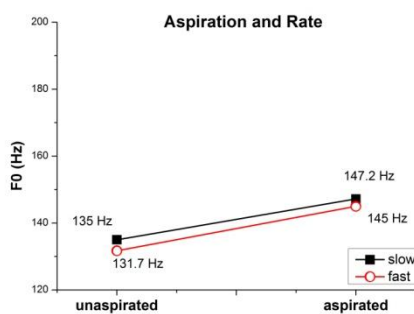
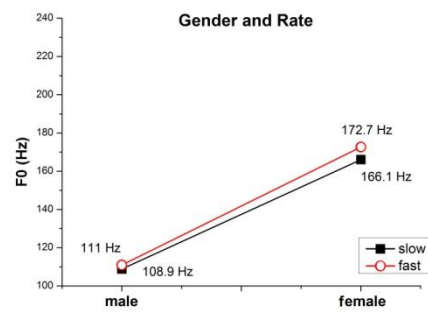
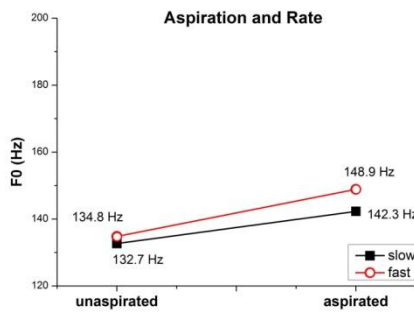
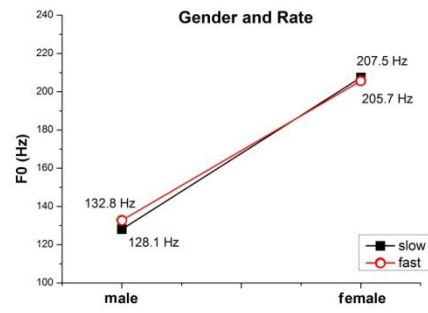
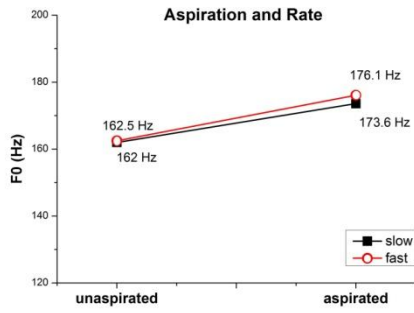


Figure 57. The interaction between Aspiration and Rate on onset F0 across four tones

Figure 58. The interaction between Gender and Rate on onset F0 across four tones