A Contrastive Investigation of Diphthongs between Standard Mandarin and Shanghai Accented Mandarin

Jue Yu, Aijun Li**, Xia Wang****

* School of foreign languages, Zhejiang University
**Institute of Linguistics, Chinese Academy of Social Sciences
*** Nokia Research Center, Beijing
*erinyu@sohu.com **liaj@linguistics.cass.net.cn ***Xia.S.wang@nokia.com

Abstract

9 diphthongs [ai, ei, ao, ou, ia, ie, ua, uo, ye] for 10 SC and 20 ASH speakers with medium accent were compared in this paper. In order to find out the differences in diphthongs between SC and ASH, the vowel charts and the formant patterns were plotted for them. Finally, comparative distinctions were obtained and analyzed.

1. Introduction

Spoken Chinese comprises many regional varieties, called dialects. There are 9 dialectal areas in China: *Guan, Jin, Wu, Hui, Xiang, Gan, Kejia, Yue and Min*. People from different dialectal areas might not be able to communicate with each other simply because the differences among the dialects are so significant. Mandarin, or *Putonghua*, (hereafter referred to as SC or Mandarin), which is widely used throughout China in almost every activity from news broadcasting to commercial trades, would be a good choice as a means of communication. [Hou, 2002]

Although lots of people CAN speak Mandarin, they speak it with different accents, depending on how well they grasp the language. The Mandarin they speak is always affected by their native dialects phonetically, lexically and syntactically. For a long time, dialectal differences have been widely investigated for dialect identification, language (L2) learning and pronunciation modeling for Automatic Speech Recognition (ASR) (Tsukada, 2002; Pellegrino et.al., 1999; Goreman, 1999; Peters, 1999; Gronnum, 1993). Especially in Chinese ASR systems, how to deal with and tackle the accent issue is becoming a big challenge due to the variability of the language. We hope the contrastive study from a phonetic point of view on regional accented Mandarin will shed a light on the Chinese ASR framework, and L2 learning and evaluation.

In my early paper (Yu Jue, 2003), I have already made a contrastive study on the acoustic differences of basic vowels between Standard and Shanghai-accented Mandarin (ASH). But in addition to the rich inventory of basic vowels [a, o, ə, i. u, y, ɿ, ʅ, ɚ], SC has 9 diphthongs [ai, ei, ao, ou, ia, ie, ua, uo, ye]. Now, I would like to continue my study to find out the differences in diphthongs between SC and ASH, and also concluded the acoustic and articulatory features of ASH diphthongs.

2. Material

The database used in this study was NOKIA-CASS, a special speech corpus designed and recorded for comparison study for three major regions: Shanghai, Suzhou and Ningbo. Here only the speech corpus for Shanghai part was used. All the recorded utterances were phonetically annotated on orthographic and pronunciation tiers by using Praat (http://www.praat.org). Pronunciation variables or phonemic changes caused by dialects were annotated dedicatedly. Prosodic and stress structure were annotated by C-ToBI (Li, 2002).

Accents of Shanghai Mandarin were first classified into three categories as light, medium and heavy by subjective criteria from dialectologists and objective criteria from statistical results obtained from the phonetic annotation. With the consideration of the generalization of the results, the phonetic level analysis on Shanghai accent was made mainly for the medium accent. 9 diphthongs [ai, ei, ao, ou, ia, ie, ua, uo, ye] for 10 SC and 20 ASH speakers with medium accent were compared. Each diphthong was investigated in 4 zero-initial monosyllabic words bearing 4 tones. For each diphthong, the first 3 formant trajectories were performed. We first manually measured the frequency values and the corresponding time of each formant trajectory at 5 points of start, the first target, the transition, the second target and the end point. Then we did normalization and interpolation in order to make all the diphthongs articulated by different speakers comparable. After that, we plotted the vowel charts and the formant patterns for them to find out the differences in diphthongs between SC and ASH.

3. Results and discussion

3.1. The phonology of SC and SHD

In addition to a rich inventory of vowels [a, o, ə, i. u, y, ɿ, ʅ, ɚ], SC has 9 diphthongs [ai, ei, ao, ou, ia, ie, ua, uo, ye]. In traditional Chinese phonetics, diphthongs are classified into two kinds, “Front Sonority” and “Back Sonority” (Wu Zhongji, 2000). The former refers to diphthongs beginning with the syllabic nucleus and ending with a high vowel, like [ai, ei, ao, ou]; the latter refers to diphthongs beginning with a medial [i, u, y] and ending with the syllabic nucleus, like [ia, ie, ua, uo, ye].

For the Back Sonorities, the medials [i, u, y] play the role of phonemic distinction when followed by the same syllabic nucleus. Medials are short but rather clear, and even consistent in fast speech. For the Front Sonorities, the ending vowels [i, u] are usually undershot and weakened in natural speech.

In Shanghai dialect, altogether there are 8 diphthongs [iA, iɔ, iE, iɤ, uA, uE, uø, yø], less than the diphthongs in SC. In fact, almost all the Front Sonority diphthongs disappear in Shanghai dialect, which is the distinctive feature shared by all Wu dialects. For example, [ai, ao] are monophthongized into [E, ɔ]. Due to this difference, we predict it would have some influence on the acquired SC diphthongs

3.2. Comparison based on the F1-F2 vowel charts

Fig.1 and Fig.2 show the vowel charts of diphthongs for SC and ASH male speakers respectively. The arrows indicate the formant movements from the first target to the second target values in F1- F2 panel. It is generally agreed that diphthongs can be viewed as a movement from an onset vowel to an offset vowel. However, whether the diphthong begins from and ends at the exact targets and to what extent the diphthong vary with the original vowels are comprehensive, because in the production of diphthongs, target undershoots may take place due to the relative short duration allocated for a particular target phone and the coarticulatory influence of the adjacent phones. In Fig.3 and Fig.4, the diphthong elements of SC and ASH male speakers are compared with their corresponding target vowels respectively to examine this phenomenon. We can see that in both SC and ASH, the vowel [a] in the diphthong [ai] [ia] is slightly more fronted than the single vowel [a] and the [a] in [au] is quite back, due to the coarticulatory effect of the adjacent vowels respectively. While under the influence of different low vowels, [i] and [u] in [ia, ai, ie, ei, ua, uo, ou] are much lower than the target vowels in SC and ASH. [ua] and [ye] show some differences between SC and ASH male speakers: in SC male speakers, the [y] in [ye] is more fronted than the target vowel while the [u] in [ua] is more back. But in ASH male speakers, these two vowels could almost reach the positions of their target vowels.

Now we'd like to examine the differences in the frequency values of vowels in each diphthong between SC and ASH male speakers. Fig.9, Fig.10 and Fig. 11 give out the vowel charts of [ai, ei, au], [ia, ie, ou], [ua, ye, uo] for SC and ASH male speakers. Three Front Sonority diphthongs [ai, ei, au] are plotted in Fig.9. From the figure, we can see that the two vowels of [ei] in ASH male speakers are more frontal and much higher than those in SC while the two vowels of [au] are somewhat more back and much lower; moreover, the [a] in [ai] for ASH male speakers is more frontal and higher than SC. In Fig.10, the three diphthongs [ia, ie, ou] in ASH male speakers are all more frontal than in SC. What's more, the two vowels of [ie] in ASH are both higher than in SC while the first vowels in [ia, ou] for ASH male speakers are higher than SC but the second vowels are both lower. Three Back Sonority diphthongs [ia, ie, ou] are plotted in Fig. 11. Among them, the two vowels of [ua, uo] in ASH are lower and more back than in SC, while the two vowels of [ye] in ASH are somewhat higher and more frontal. To sum up, in comparison with the diphthongs of SC male speakers, the Back Sonority diphthongs beginning with [i, y] of ASH male speakers are all more frontal and higher (except [ia]) while those beginning with [u] all more back and lower; regarding the Front Sonority diphthongs, data lacks consistency. The two vowels of [ei, ou] in ASH are more frontal than in SC but the two vowels of [au] in ASH more back. Comparing Fig.1 and Fig. 2, we find out that the whole vocalic space of diphthongs in ASH male speakers keeps more peripheral than that in SC male speakers, which is in accordance with the comparative results of the male monophthongs. It indicates that the production of diphthongs for ASH male speakers is under the influence of the original basic vowels.

We do the same analysis to female speakers. Fig.5 and Fig.6 are the vowel charts of diphthongs for SC and ASH female speakers. Fig.7 and Fig.8 show the comparative results of the vowels in diphthongs and their corresponding target vowels in SC and ASH female speakers respectively. From Fig.7 and Fig.8, we can see the comparative results of female speakers, on the whole, are almost the same with male speakers except [ye] and [ua]. In fact, the [y] in [ye] and the [a] in [ua] for ASH female speakers, not like ASH male speakers, both deviate from the positions of their target vowels. But for SC female speakers, the [y] in [ye] and the [a] in [ua] show the same features as SC male speakers.

Now we'd also like to make a comparative study about the differences in the frequency values of vowels in each diphthong between SC and ASH female speakers. Fig.12, Fig.13 and Fig.14 are the vowel charts of [ai, ei, au]，[ia, ie, ou]，[ua, ye, uo] for SC and ASH female speakers. From the figures, we find out that all diphthongs (except [ua, uo]) in ASH female speakers are more frontal than those in SC female speakers. Moreover comparing Fig.5 and Fig. 6, we can see that, like ASH male speakers, the whole vocalic space of diphthongs in ASH female speakers also keeps more peripheral than that in SC female speakers, which is in accordance with the comparative results of the female basic vowels. It indicates that the production of diphthongs for ASH female speakers is also under the influence of the original basic vowels.

3.3. Comparison based on the formant patterns

By observing the formant patterns of different diphthongs for male speakers in SC and ASH, we find out that except [uo], the formant patterns of all other diphthongs between SC and ASH male speakers show differences. From Fig.15, Fig.22, Fig.23 and Fig.28, we can see the first formant trajectories of these diphthongs are similar for SC and ASH male speakers, but there are differences in the second and the third formant trajectories. Although the whole formant patterns of [ei, ia, ua] for ASH male speakers deviate from those for SC male speakers, the differences are slight (see Fig.16, Fig.21 and Fig.27). To sum up, the formant patterns of [ai, ie, ou, ye, ao] for SC and ASH male speakers show greater differences, and the third formant trajectories of [ai, ou, ao] for ASH male speakers are lower than that for SC male speakers. This indicates that ASH male speakers, generally speaking, produce round-lip sound much rounder than SC male speakers but produce [a] with much smaller open mouth.

As for female speakers, the results are very different from male speakers. Almost the formant pattern of each diphthong shows some differences. The second and the third formant trajectories of [ai, ao, ou, ye] for ASH female speakers are

located at a higher position than for SC female speakers, and show greater range of change (see Fig.18, Fig.20, Fig.26, Fig. 31). From Fig.19, Fig.24, Fig.25, Fig.30 and Fig. 32, we can see: the third formant trajectories of [ia, ua] for ASH female speakers are located at a higher position than for SC female speakers, and the whole formant patterns of [ei, ie] show slighter difference between SC and ASH female speakers but the formant pattern of [uo] for ASH female speakers is totally different from that for SC female speakers. Comparative speaking, the formant patterns of [ai, ao, ou, ye, ia, ua] show greater differences between SC and ASH female speakers and the third formant trajectories of [ai, ao, ou, ye, ia, ua] for ASH female speakers are all located at higher position than for SC female speakers. This indicates that ASH female speakers, different from ASH male speakers, usually produce round-lip sound less round and produce [a] with larger open mouth than SC female speakers. In addition, both in SC and ASH, all diphthongs of female speakers show greater range of change than male speakers. That is maybe the reason why female speakers usually produce sounds much clearer than males.

4. Conclusion

By comparing of vowels, diphthongs and triphthongs between standard and Shanghai-accented mandarin on the vowel charts and formant patterns, we find out that:

1. In the production of diphthongs for both SC and ASH speakers, target undershoots take place due to the relative short duration allocated for a particular target element and the coarticulatory influence of the adjacent element.
2. The whole vocalic space of diphthongs in ASH speakers keeps more peripheral than that in SC speakers, which is in accordance with the comparative results of monophthongs.
3. To sum up, in comparison with the diphthongs of SC male speakers, the back sonority diphthongs beginning with [i, y] of ASH male speakers are all more frontal and higher (except [ia]) while those beginning with [u] all more back and lower; regarding the front sonority diphthongs, data lacks consistency. But all diphthongs (except [ua,uo]) in ASH female speakers are more frontal than those in SC female speakers.
4. The formant patterns of [ai, ie, ou, ye, ao] for SC and ASH male speakers show greater differences, and the third formant trajectories of [ai, ou, ao] for ASH male speakers are lower than that for SC male speakers. But as for female speakers, the formant patterns of [ai, ao, ou, ye, ia, ua] show greater differences between SC and ASH and the third formant trajectories of [ai, ao, ou, ye, ia, ua] for ASH female speakers are all located at higher position than for SC female speakers.

5. References

[1] Goreman, C.G., 1999. Dialect Identification From Prosodic Cues. *ICPhS99*, 1237-1240.

[2] Hou, J. Y., 2002. *The Outline of Modern Chinese Dialects*. Shanghai Education Publishing House.

[3] Pellegrino, F. and Barkat M., 1999. Investigating Dialectal Differences via Vowel System Modeling: Application to Arabic. *ICPhS99,* 145-148.

[4] Peters, J., 1999. The Timing of Nuclear High Accent in German Dialects. *ICPhS99*, 1877-1880.

[5] Tsukada, K., 2002. An Acoustics Comparison between American English and Australian English Vowels. *ICSLP'2002*, 2257-2260.

[6] Wu, Z.J., 2000. From Traditional Chinese Phonology to Modern Speech Processing: Realization of Tone and Intonation in Standard Chinese. *Report of Phonetic Research 2000*, 1-11.

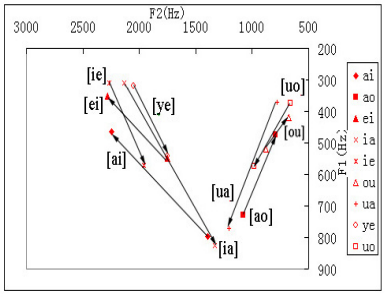

Fig.1: *Vowel chart for SC males*

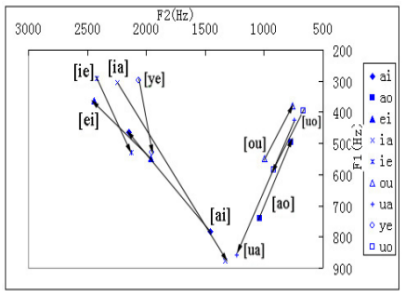

Fig.2: *Vowel chart for ASH males*

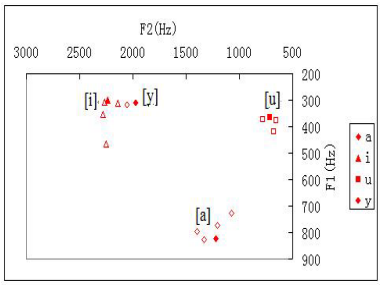

Fig.3: *The vowels in diphthongs of SC males compared with target vowels (empty —in diphthongs, solid —target vowels)*

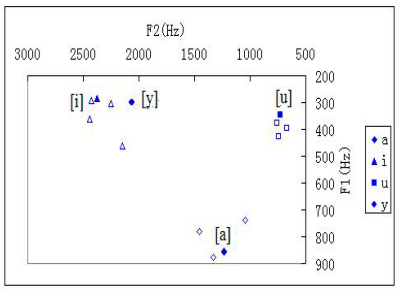

Fig.4: *The vowels in diphthongs of ASC males compared with target vowels (empty —in diphthongs, solid —target vowels)*

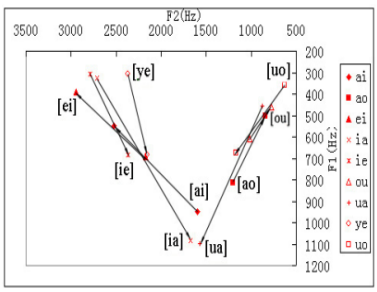

Fig.5: *Vowel chart for SC females*

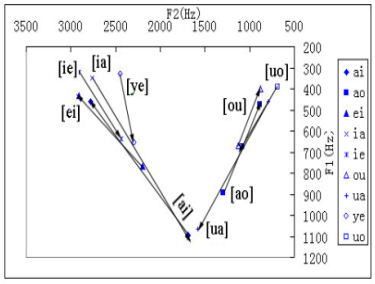

Fig.6: *Vowel chart for ASH females*

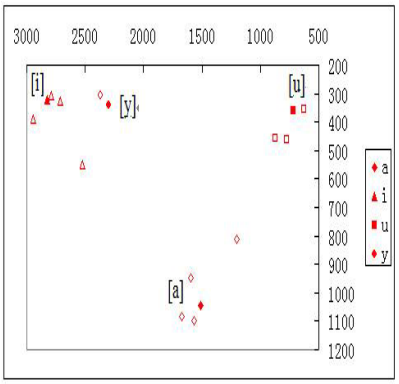

Fig.7: *The vowels in diphthongs of SC females compared with target vowels (empty —in diphthongs, solid —target vowels)*

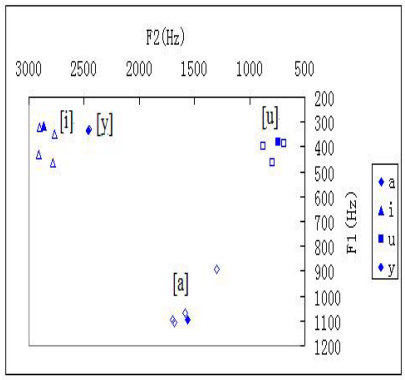

Fig.8: *The vowels in diphthong of ASH females compared with target vowels (empty spot —in diphthongs, solid spot — target vowels)*

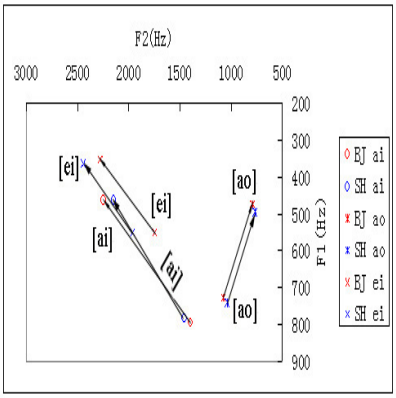

Fig.9: *Vowel chart of [ai,ei,ao] for males (ASH —big-size arrow, SC —small-size arrow)*

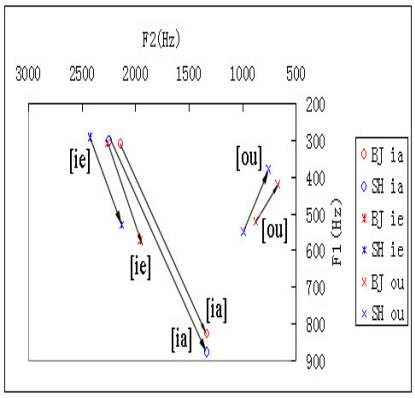

Fig.10: *Vowel chart of [ia,ie,ou] for males (ASH —big-size arrow, SC —small-size arrow)*

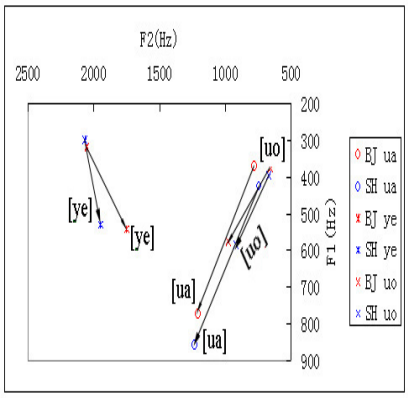

Fig.11: *Vowel chart of [ua,ye,uo] for males (ASH —big-size arrow, SC — small-size arrow)*

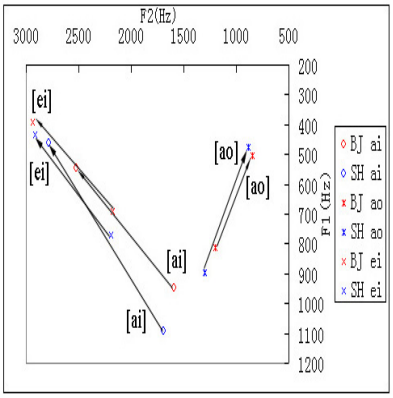

Fig.12: *Vowel chart of [ai,ei,ao] for females (ASH —big-size arrow, SC — small-size arrow)*

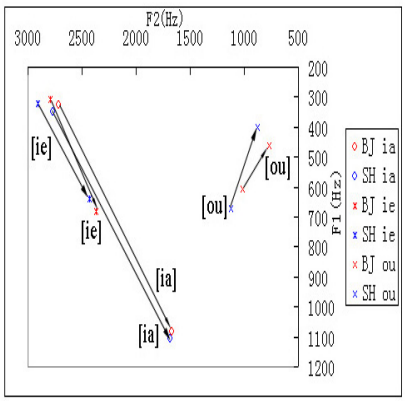

Fig.13: Vowel chart of [ia,ie,ou] for females (ASH —big-size arrow, SC — small-size arrow)

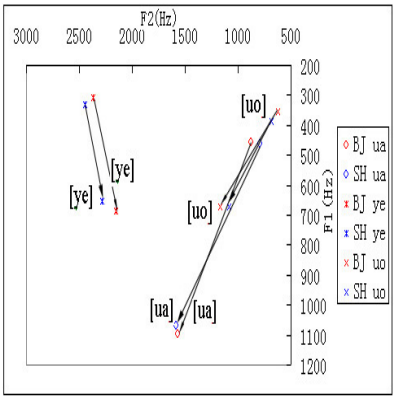

Fig.14: *Vowel chart of [ua,ye,uo] for females (ASH —big-size arrow, SC — small-size arrow)*

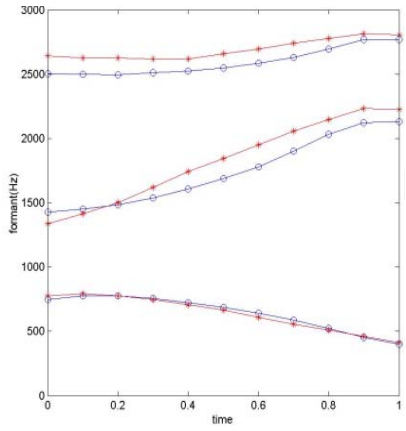

Fig.15: *Formant pattern of [ai] for males (empty spot —ASH, solid spot —SC)*

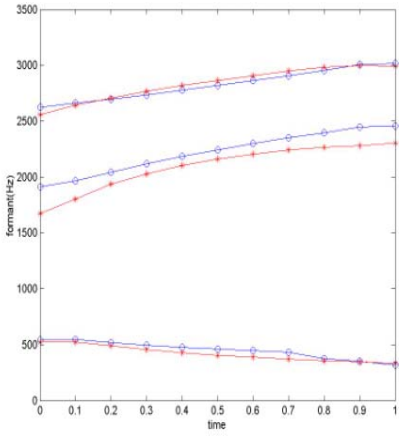

Fig.16: *Formant pattern of [ei] for males (empty spot —ASH, solid spot —SC)*

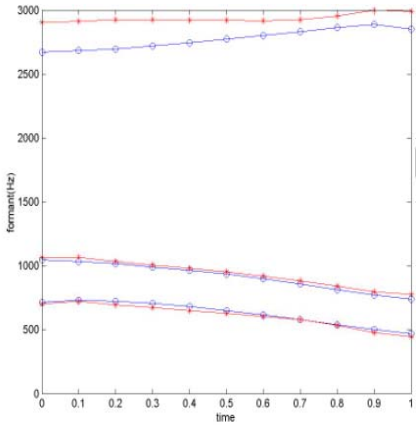

Fig.17: *Formant pattern of [ao] for males(empty spot —ASH, solid spot —SC)*

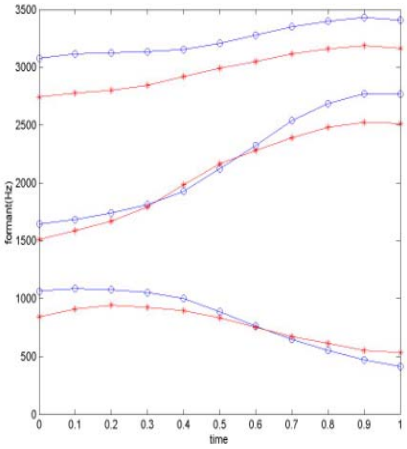

Fig.18: *Formant pattern of [ai] for females (empty —ASH solid —SC)*

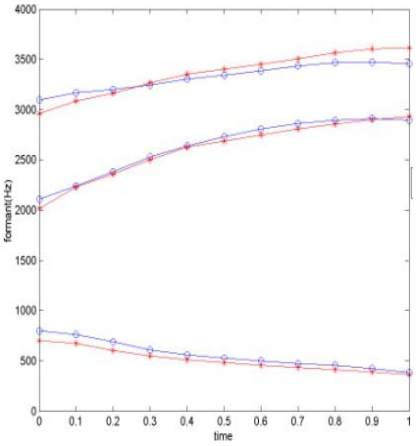

Fig.19: *Formant pattern of [ei] for females (empty —ASH, solid —SC)*

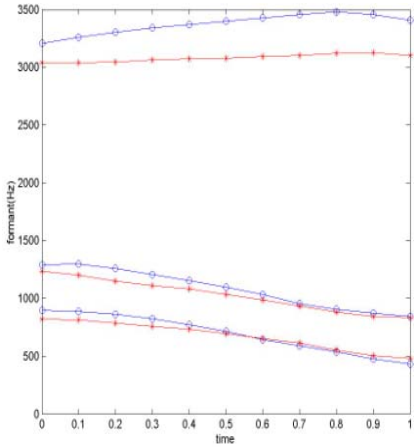

Fig.20: *Formant pattern of [ao] for females (empty —ASH, solid —SC)*

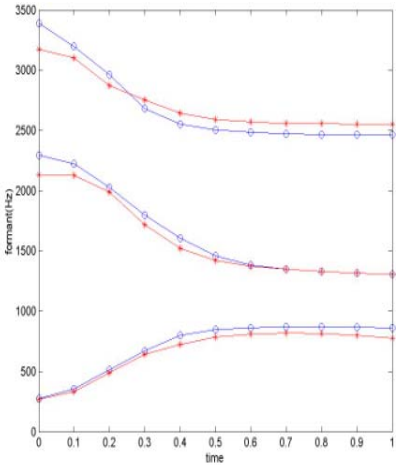

Fig.21: *Formant pattern of [ia] for males (empty spot —ASH, solid spot —SC)*

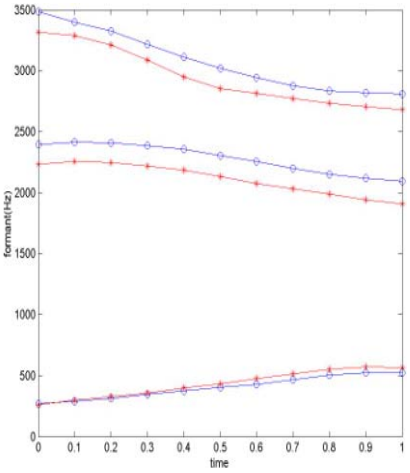

Fig.22: *Formant pattern of [ie] for males (empty spot —ASH, solid spot —SC)*

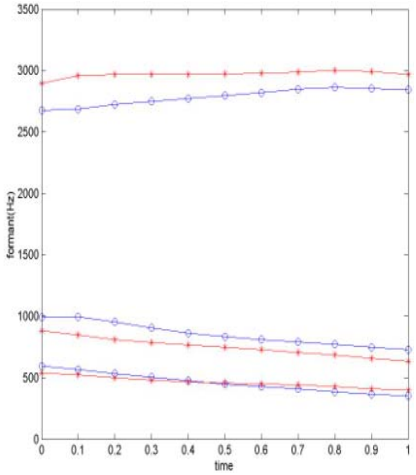

Fig.23: *Formant pattern of [ou] for males (empty —ASH, solid —SC)*

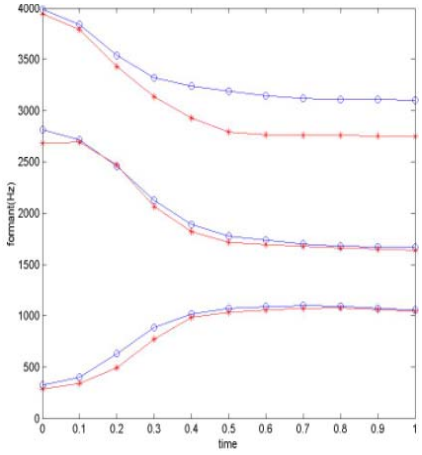

Fig.24: *Formant pattern of [ia] for females（empty spot —ASH, solid spot —SC）*

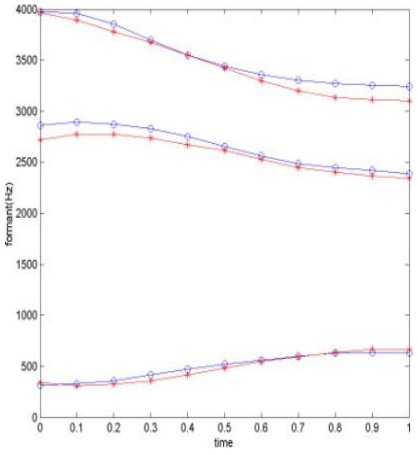

Fig.25: *Formant pattern of [ie] for females (empty —ASH, solid —SC）*

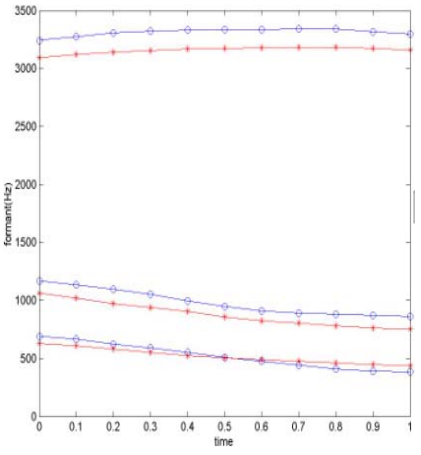

Fig.26: *Formant pattern of [ou] for females (empty —ASH, solid —SC）*

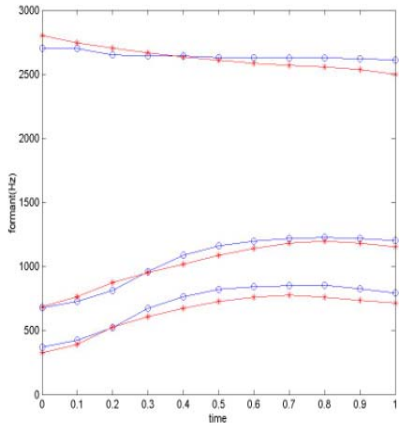

Fig.27: *Formant pattern of [ua] for males（empty spot —ASH, solid spot —SC）*

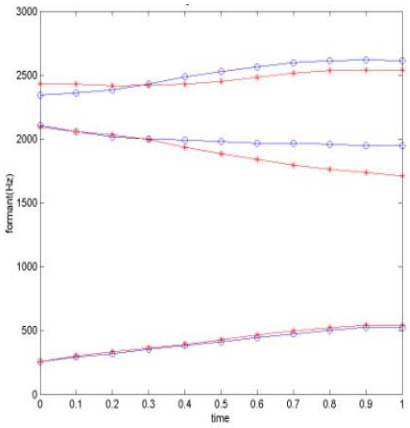

Fig.28: *Formant pattern of [ye] for males（empty spot —ASH, solid spot —SC）*

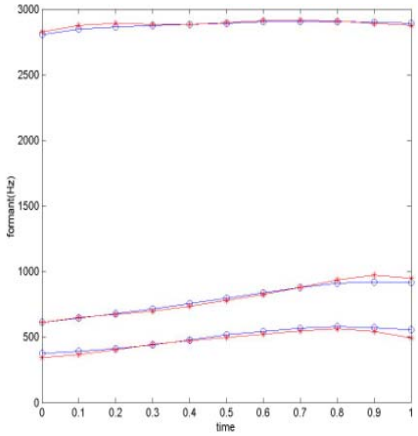

Fig.29: *Formant pattern of [uo] for males (empty —ASH, solid —SC）*

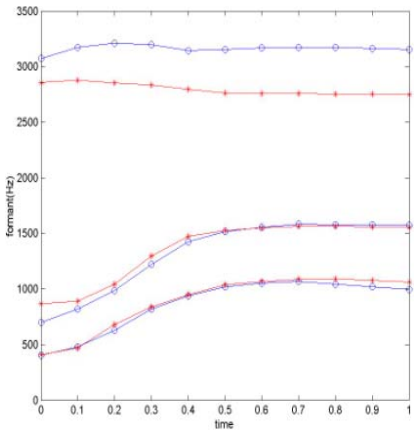

Fig.30: *Formant pattern of [ua] for females (empty —ASH, solid SC）*

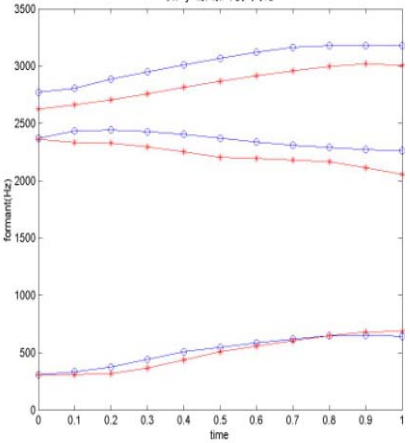

Fig.31: *Formant pattern of [ye] for females（empty —ASH, solid —SC）*

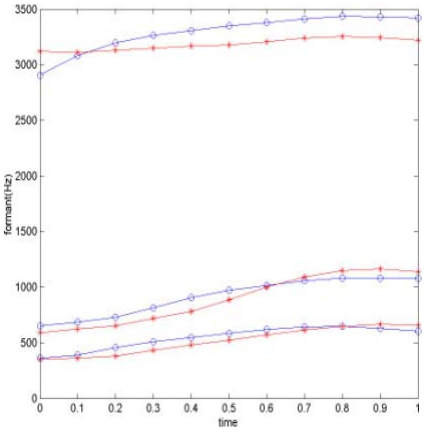

Fig.32: *Formant pattern of [uo] for females（empty —ASH, solid —SC）*