

## Effects of Environmental Noises on Fundamental Frequency Contours of Thai Dialects

<sup>1</sup>Suphattharachai Chomphan and <sup>2</sup>Chutarat Chompunth

<sup>1</sup>Department of Electrical Engineering, Faculty of Engineering at Si Racha,  
Kasetsart University, 199 M.6, Tungsookhla, Si Racha, Chonburi, 20230, Thailand

<sup>2</sup>School of Social and Environmental Development, National Institute of Development  
Administration, 118 M.3, Serithai Road, Klong-Chan, Bangkok, 10240, Thailand

---

**Abstract: Problem statement:** Fundamental Frequency (F0) is an important speech feature defining the human speech prosody. It is the resultant of the vibration of human's vocal chords in speech production. In Thai, there are four main dialects spoken by Thai people residing in four core region including central, north, northeast and south regions. Environmental noises are also playing an important role in corrupting the speech quality. The study of effects of noises the F0 contour for Thai dialects reveals the important of noise reduction issue. **Approach:** Four types of environmental noises were recorded with different levels of power. It was subsequently mixed with clean speech. The F0 contours from different dialects, different types of noises and different levels of noises was extracted. The difference in term of Root Mean Square Error (RMSE) between the F0 contour of clean speech and the noise-corrupted speech was calculated. **Results:** In the experiments, each regional dialect includes 10 samples of 10 utterances with male and female speech. Four types of noises include train, factory, car and air conditioner. Moreover, five levels of each type of noise are varied from 0-20 dB. The results show that effects of distinguish types of noises are different. Four kinds of regional dialects also cause the differences in RMSEs. **Conclusion:** The recorded noises deteriorate the F0 contours for all Thai dialects.

**Key words:** Thai dialects, fundamental frequency contour, analysis of fundamental frequency, environmental noise, noise effect

---

### INTRODUCTION

In the recent study on modeling of F0 contour with noisy environment, the simulated noises deteriorate the Fujisaki's model parameters (Fujisaki and Sudo, 1971; Mixdorff and Fujisaki, 1997; Seresangtakul and Takara, 2003). However the study on the direct effect of noises on the fundamental frequency contour has not been conducted (Chomphan, 2010a; 2010b). This study proposes an analysis the differences between the F0 contour of clean speech and the noise-corrupted speech in term of RMSE. An example of F0 contour is shown in logarithmic scale in Fig. 1. The concerned Thai main dialects include the standard Thai or Central dialect, Lanna or North dialect, Lao-style or Northeast dialect and South dialect. Meanwhile the selected four types of noises are air-conditioner, car, factory and train noises.

### MATERIALS AND METHODS

**Experimental design:** The procedure of the experiment is designed and depicted in Fig. 2. The first step is to

construct the speech database of four Thai dialects. Meanwhile, the noise database is also constructed with four different types including air-conditioner, car, factory and train noises. The F0 contours of clean speech are extracted from the speech database in the "calculation of F0 contour" stage. Moreover, the clean speech from speech database is mixed with all four types of noises from the noise database in the "noises mixing noises with clean speech" stage. Subsequently, the F0 contours of noise-corrupted speech are extracted in another stage of "calculation of F0 contour". The differences in terms of RMSE are then calculated in the "RMSE calculation" stage. In the data analysis stage, the RMSE values are analyzed comparatively (Chomphan and Kobayashi, 2008; 2009).

The main four dialects are spoken in all four regions of Thailand. They are chosen to be concerned in this study. They include the standard Thai or Central dialect, Lanna or North dialect, Lao-style or Northeast dialect and South dialect (Chomphan and Kobayashi, 2007a; 2007b).

---

**Corresponding Author:** Suphattharachai Chomphan, Department of Electrical Engineering, Faculty of Engineering at Si Racha, Kasetsart University, 199 M.6, Tungsookhla, Si Racha, Chonburi, 20230, Thailand

**Environmental noises:** Four types of noises include train, factory, car and air conditioner. They are mixed directly with the pre-recorded clean speech in the speech database. Before mixing noises with the clean speech, the noise volume or power are adjusted in several exact levels. As for the level variation of noises, the levels of each type of noise are varied from 0, 5, 10, 15 and 20 dB, respectively.

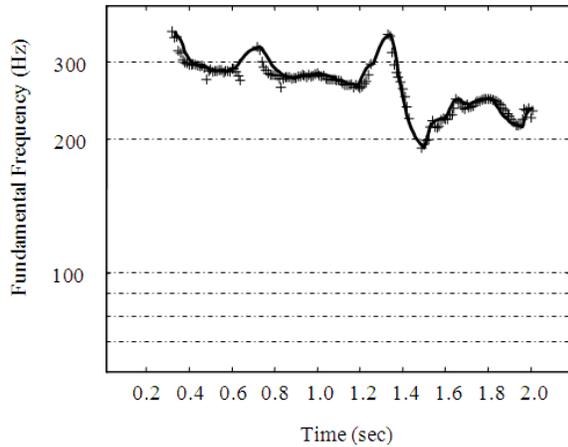


Fig. 1: An example of F0 contour with a trend line in a logarithmic scale

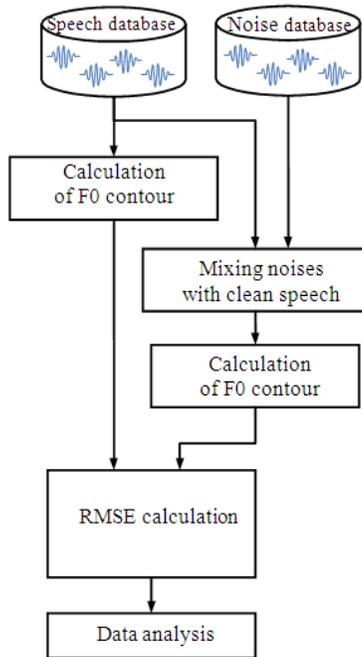


Fig. 2: Experimental flow chart determining the procedures in the study

## RESULTS

In our speech database, we use ten sentences in Thai for male and female genders. The sentences have been recorded in four Thai dialects of standard Thai (Center-dialect), Lanna Thai dialect (North-dialect), Lao-style Thai dialect (Northeast-dialect) and South Thai dialect (South-dialect). Each dialect contains a hundred utterances of samples. Therefore we have four hundred utterances of samples for each gender (Mixdorff and Fujisaki, 1997; Chomphan and Kobayashi, 2007a; 007b).

From the “data analysis” stage in Fig. 2, the following charts are summarized. First, the RMSE values from four different types of noises at the noise level of 0dB with both male and female genders are presented in Fig. 3. Second, the RMSE values from four different types of noises at the noise level of 20 dB with both male and female genders are presented in Fig. 4. Third, the RMSE values from four different Thai dialects and four distinguished types of noises at the noise level of 0dB for male speech are presented in Fig. 5. Fourth, the RMSE values from four different Thai dialects and four distinguished types of noises at the noise level of 0 dB for female speech are presented in Fig. 6.

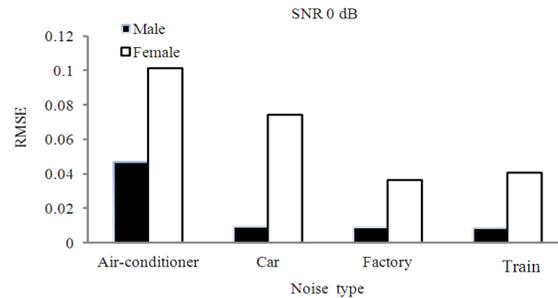


Fig. 3: RMSEs from four different types of noises at the noise level of 0dB with both male and female genders

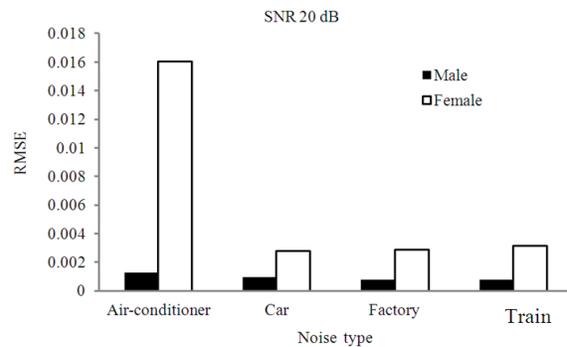


Fig. 4: RMSEs from four different types of noises at the noise level of 20dB with both male and female genders

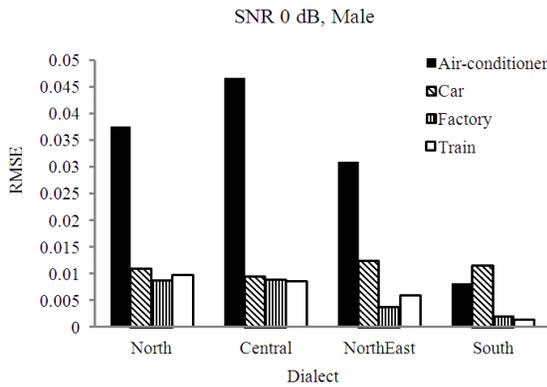


Fig. 5: RMSEs from four different Thai dialects and four distinguished types of noises at the noise level of 0dB for male speech

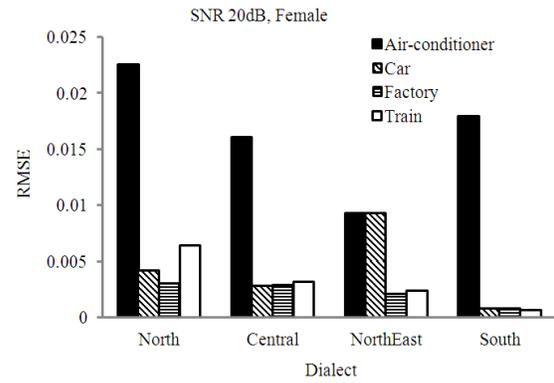


Fig. 8: RMSEs from four different Thai dialects and four distinguished types of noises at the noise level of 20dB for female speech

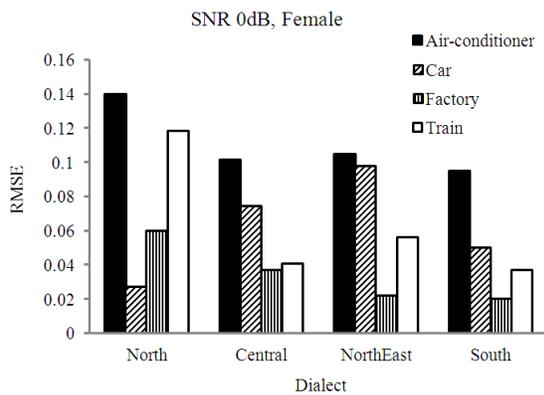


Fig. 6: RMSEs from four different Thai dialects and four distinguished types of noises at the noise level of 0dB for female speech

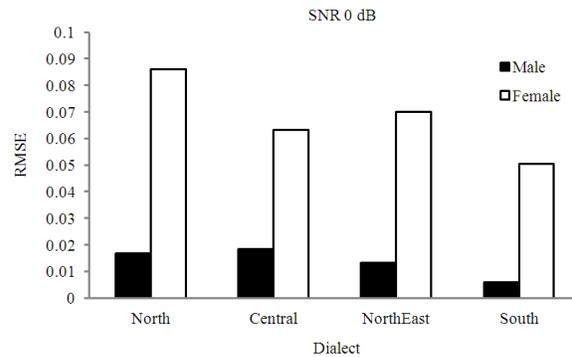


Fig. 9: RMSEs from four different Thai dialects with both male and female speech

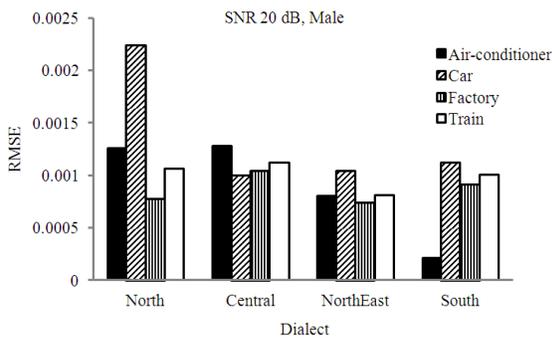


Fig. 7: RMSEs from four different Thai dialects and four distinguished types of noises at the noise level of 20dB for male speech

Fifth, the RMSE values from four different Thai dialects and four distinguished types of noises at the noise level of 20 dB for male speech are presented in Fig. 7.

Sixth, the RMSE values from four different Thai dialects and four distinguished types of noises at the noise level of 20 dB for female speech are presented in Fig. 8. Seventh, the RMSE values from four different Thai dialects with both male and female speech are presented in Fig. 9. It has been noted that the other levels of noises have the corresponding results to both 0 and 20 dB.

## DISCUSSION

From Fig. 3 and 4, it can be obviously seen that female speech gives the higher RMSEs than that of male speech for all four types of noises. Then comparing among types of noises, the air-conditioner gives the highest RMSEs for both levels of noises, meanwhile the other types of noises give similar values of RMSEs. From Fig. 5 (male speech, 0 dB), it can be seen that air-conditioner noise gives the highest values of RMSEs. The second order is of car noise except for the south dialect. From Fig. 6 (female speech, 0 dB), it can be seen that air-conditioner noise also

gives the highest values of RMSEs. The second order is of car noise except for the north dialect. From Fig. 7 (male speech, 20 dB), it can be observed that car noise gives the highest values of RMSEs for nearly all dialects except for north dialect. From Fig. 8 (female speech, 20 dB), it can be noticed that air-conditioner noise gives the highest values of RMSEs for all dialects. From Fig. 9, it can be observed that female speech gives much higher values of RMSE than that of male speech. Moreover the south dialect gives the lowest values of RMSEs for both genders. These results can be implied that female speech is affected by environmental noises much more than that of male speech.

### CONCLUSION

This study presents a study of effects of noises on F0 contour for Thai dialects. Four types of environmental noises are recorded with different levels of power. The differences of F0 contours between the clean speech and the noise-corrupted speech are calculated in terms of RMSEs. All in all, the simulated noises deteriorate the F0 contours differently depending on type of noise, level of noise, gender and dialects.

### ACKNOWLEDGEMENT

The author is grateful to Kasetsart University for the research scholarship through the Center for Advanced Studies in Industrial Technology.

### REFERENCES

Chomphan, S. and T. Kobayashi, 2007a. Design of tree-based context clustering for an HMM-based Thai speech synthesis system. Proceeding of the 6th ISCA Workshop on Speech Synthesis (SSW6), ISCA, Aug. 2007, Bonn, Germany, pp: 160-165.

Chomphan, S. and T. Kobayashi, 2007b. Implementation and evaluation of an HMM-based Thai speech synthesis system. Proceeding of the 8th Annual Conference of the International Speech Communication Association, Aug. 27-31, ISCA, Antwerp, Belgium, pp: 2849-2852.

Chomphan, S. and T. Kobayashi, 2008. Tone correctness improvement in speaker dependent HMM-based Thai speech synthesis. *Speech Commun.*, 50: 392-404. DOI: 10.1016/j.specom.2007.12.002

Chomphan, S. and T. Kobayashi, 2009. Tone correctness improvement in speaker-independent average-voice-based Thai speech synthesis. *Speech Commun.*, 51: 330-343. DOI: 10.1016/j.specom.2008.10.003

Chomphan, S., 2010a. Analytical study on fundamental frequency contours of Thai expressive speech using Fujisaki's model. *J. Comput. Sci.*, 6: 36-42. DOI: 10.3844/jcssp.2010.36.42

Chomphan, S., 2010b. Fujisaki's model of fundamental frequency contours for Thai dialects. *J. Comput. Sci.*, 6: 1263-1271. DOI: 10.3844/jcssp.2010.1263.1271

Fujisaki, H. and H. Sudo, 1971. A model for the generation of fundamental frequency contours of Japanese word accent. *J. Acoust. Soc. Jap.*, 57: 445-452.

Mixdorff, H. and H. Fujisaki, 1997. Automated quantitative analysis of F0 contours of utterances from a German ToBI-labeled speech database. Proceeding of the Eurospeech, September 22-25, ISCA, Rhodes, Greece, pp: 187-190.

Seresangtakul, P. and T. Takara, 2003. A generative model of fundamental frequency contours for polysyllabic words of Thai tones. Proceedings of the International Conference on Acoustics, Speech and Signal Processing, IEEE, April 6-10, 2003, pp: 452-455. DOI: 10.1109/ICASSP.2003.1198815