CONCLUSION

This thesis focused on the problem of current Tamil speech recognizers which are still constrained to speaker-dependent small vocabulary isolated speech recognition tasks. This research focused on releasing these constraints and provides a basis on developing the speaker-independent large vocabulary Tamil continuous speech recognition system. In short the system can be termed as Intelligent system for Tamil speech recognition.

This thesis focuses on both recognition task of small vocabulary and medium vocabulary based Tamil speech database. In small vocabulary tasks, the Tamil isolated word and Tamil isolated numeral recognition (with vocabulary of 10 Tamil digits 0-9) and Tamil connected word and Tamil connected numeral recognition (Pathin ondru, Iruvathi Ondru) will be considered, where word models are used. In medium vocabulary task, instead of developing a complete medium vocabulary continuous speech recognition system, the thesis is scoped to Tamil phoneme classification and segmentation of medium vocabulary Tamil continuous speech database (with vocabulary of 240 words).

The performance study reveals that the system comes out with impressive performance in clean conditions (Noiseless) for Isolated word and Numeral recognition, regardless of the speaker. The system performance is slightly reduced for the case of Continuous speech, which can be due to the problem in segmentation of the Continuous speech. The system comes up with the same performance as in clean environment for any normal environment (noisy), which in turn shows that the system is unaffected by that of noise added to the signal.
The noise performance studied with PSNR reveals that the system came up with excellent recognition accuracy even in the case of very low value of PSNR, which in turn reveals that the system developed is highly independent of the noise added to the signal.

**Future Work**

There are several ways of expanding this system for example by including speaker identification functionality and adding new solutions for noise robustness. Other ways to continue is to expand to a text-independent system. Since DTW lacks the generalization power, this would require a pattern matching technique such as HMM and GMM could be used. The MFCC and signal processing part of this work would still be viable to use. The conversion of speech to text is another expansion that could be added to this system. But this requires a very large database of words and also requires more amount of training.