Synthesis of Hindi Consonants with Naturally Sounding Quality using PC Based Klatt Synthesiser

Rajesh Verma and Shyam S. Agrawal

Central Electronics Engineering Research Institute Centre, CSIR Complex, NPL Campus
Hill Side Road, New Delhi 110 012, INDIA
Fax : 91-11-578 8347, Email: rv@ceerid.ernet.in, sagrawal@ceerid.ernet.in

ABSTRACT: This paper describes experiments conducted to synthesise Hindi consonants with naturally sounding quality using the PC based Klatt synthesiser. All the most frequently used 29 Hindi consonants have been synthesised in the form of CV and VC syllables. For this purpose, a detailed spectral analysis was carried out to determine minute acoustic differences among the various categories of consonants such as unaspirated/aspirated, voiced/unvoiced, dental/retroflex, liquid/trill etc. The synthesis was done using a modified version of Klatt Synthesiser simulated on a PC. Variations in important parameters were made until a satisfactory quality of consonant was perceived and a clear distinction occurred with its cognate sound.

INTRODUCTION

In the past KLSYN88 has been successfully used for voice quality variations of male, female speakers for American English and several other languages [1]. A PC based version of KLSYN88 was developed later at this institute, with the modified aspirated source to suit the synthesis of Hindi sounds as shown in Figure 1. The current experiment was conducted to use this synthesizer to generate high quality synthetic Hindi sounds, with an objective to use this data in development of unlimited vocabulary Text to Speech conversion system for Hindi [2].

![Figure 1. Block diagram of PC-KLATT formant synthesizer](image)

SOUNDS OF HINDI AND THEIR ANALYSIS

There are 10 vowels and 29 most frequently used consonants in Hindi. These consonants possess certain special features, not available in European Languages or American English. Certain stops and affricates which use both voicing and aspiration are the distinct sounds of Hindi. For this reason aspiration source of the Klatt synthesiser has been modified to have energy even at a frequency as low as 300 Hz. The vowel quality features such as breathiness or laryngealization can be effectively controlled by glottal source parameters like Q and TL [3]. All the 29 consonants of Hindi were analysed in CV and VC form in five vowel contexts /a/, /i/, /u/, /e/ and /o/, using standard LPC/FFT techniques, to look for important acoustical landmarks in these consonants i.e. voicebar, burst frequencies, VOT, formant transitions from consonant to vowel and vice-versa.
OBSERVATIONS

During synthesis of syllables, special attention was paid to the parameters such as F2, F3, A2F, A3F, A4F and that play important role in distinction between cognate sounds like /t/ and /t*/. QQ, TL and corresponding formant bandwidths are kept same. A comparison spectrograms of synthetic syllables /ta/, /t*a/, /ti/ and /t*i/ are shown in Figure 2. Also the variations in acoustic realisation of sound /r/ was studied in minute details. It was observed that /r/ behaves like a vowel in word end position as in /sansar/, while, it looks like a trill in word initial position and like fricative if it is followed by a retroflex as in /rasht*ra/. Spectrogram for these variations are shown in Figure 3.

CONCLUSIONS

Feasibility of using Klatt synthesizer for close imitation of natural sounding Hindi consonants has been shown. All the 29 CV and 29 VC syllables have been generated. Consonants had been synthesized with clear distinctions with their cognate sounds.

ACKNOWLEDGEMENTS

The authors are grateful to Prof. RN Biswas, Director, CEERI, Pilani for permitting to publish this work. They are also thankful to Prof K N Stevens, MIT, Cambridge, USA for his useful suggestions.

REFERENCES