

# Internet Speech Analysis System Using E-mail and Web Technology

Jouji Miwa and Masaru Kumagai

Department of Computer and Information Science, Iwate University  
4-3-5 Ueda Morioka, Iwate, 020-8851, Japan

## ABSTRACT

This paper describes a system for speech analysis on the Internet using E-mail and Web technology called as INSAS/M.

A user's speech waveform attached in E-mail is analyzed and converted to an image file with the system. Then, the user can easily see the analyzed result in a web page with a browser.

The analysis functions are formant locus and section displayed by LPC analysis, fundamental frequencies and so on.

The system can be used at any time, in anywhere and for anyone so that it is very suitable for a phonetician and a busy language learner in the world.

## 1. INTRODUCTION

A tool of speech analysis is important for a phonetician, a language learner and so on. But there are problems that the tool is expensive and necessary for installation.

We have developed a system for speech analysis on the Internet using E-mail and Web technology called as INSAS/M [1].

A user's speech waveform attached in E-mail is analyzed and converted to an image file with the system. Then, the user can easily see the analyzed result in a web page with a browser.

The system has merits such as no charge, no installation and on the Internet. And the system can be used at any time, in anywhere and for anyone so that it is very suitable for a phonetician and a busy language learner in the world.

## 2. SYSTEM OVERVIEW

### 2.1 SYSTEM CONFIGURATION

Fig. 1 shows a block diagram of an Internet speech analysis system. Speech is recorded in a file with an application program such as "sound recorder". A user's speech waveform attached in E-mail is analyzed and converted to an image file in a speech analysis server. Then, the user can easily see the analyzed result in a web page with a browser.

### 2.2 FLOWCHART OF SPEECH ANALYSIS

Fig. 2 shows a flowchart of speech analysis system. The flowchart consists of 5 steps as follows.

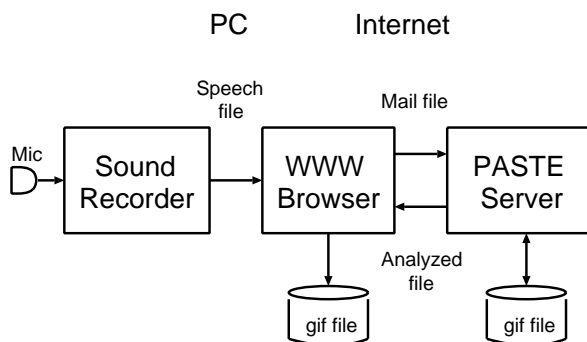


Figure 1: Block diagram of speech analysis system.

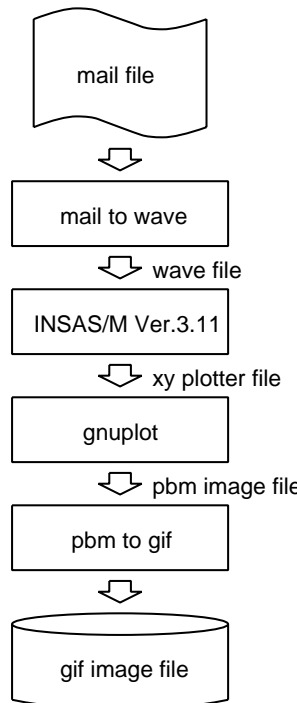


Figure 2: flowchart of speech analysis system.

1. Read a mail file
2. Conversion of base64 code in E-mail to raw file
3. Speech analysis by INSAS/M[2] Ver.3.11
4. Conversion of X-Y plotter image to PBM image
5. Conversion of PBM image to GIF mage file

### 3. SPEECH ANALYSIS

#### 3.1 FILE FORMAT

Table 1 shows file format for speech analysis.

Table 1: File format.

Encode	Base64
File name	English.WAV
File format	RIFF, .WAV
Sampling Frequency	11.025 kHz
No. of channels	1 channel
No. of bits	16 bits
PCM	Linear
Maximum length	3 seconds

#### 3.2 ANALYSIS CONDITIONS

Table 2 shows conditions for speech analysis.

Table 2: Analysis conditions.

Frame period	10 ms
Window	Hamming(30 ms)
Analysis	13order LPC
Formant extraction	Peak picking (Bandwidth $\leq$ 800 Hz)
Fundamental frequency	LPC residual ( $\theta \geq 0.3$ )
F0 range	
male	50-200Hz
female	200-400Hz
child	300-600Hz

#### 3.3 ANALYSIS COMMANDS

Table 3 shows commands for speech analysis.

#### 3.4 SPEECH ANALYSIS

Fig.3 to Fig.11 show examples of speech analysis. Fig.6 to Fig.11 are examples for language learning.

Table 3: Analysis commands

Setup commands	F0 range
male (default)	50 ~ 200Hz (male)
female	200 ~ 400Hz (female)
child	300 ~ 600Hz (child)
none	50 ~ 200Hz (male)
Analysis commands	analysis
lc (default)	LPC locus
sc	LPC section
f0	F0 locus
accent	accent
f1f2	F1-F2 plane (no vowel distributions)
f1f2j	F1-F2 plane (distributions of Japanese vowels)
f1f2e	F1-F2 plane (distributions of English vowels)
f1f3	F1-F3 plane (no distribution)
f1f3e	F1-F3 plane (distributions of /r/ and /l/)
none	LPC locus
sc option	analysis
integer	frame number
none (default)	maximum point

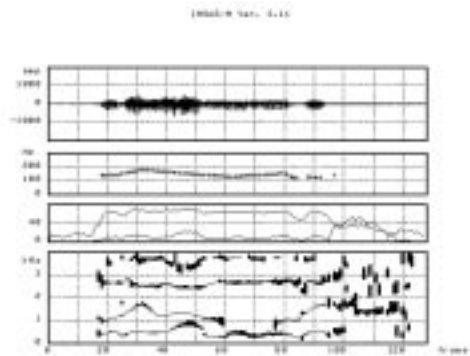


Figure 3: An example of formant locus analysis.

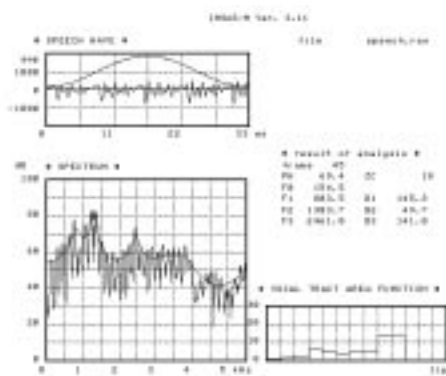


Figure 4: An example of section analysis.

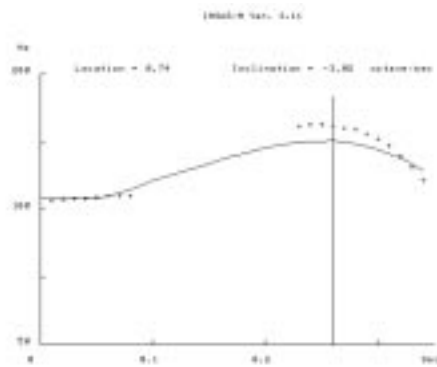


Figure 7: An example of Japanese word accent 2.

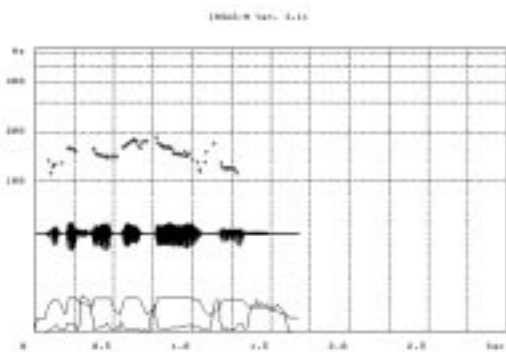


Figure 5: An example of F0 locus analysis.

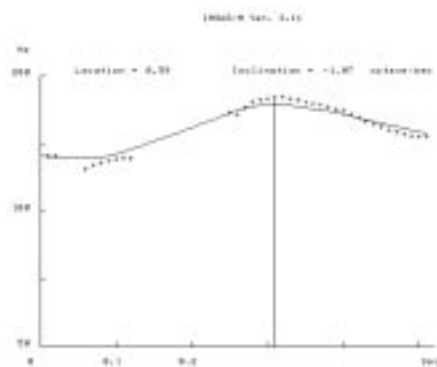


Figure 8: An example of Japanese word accent 0.

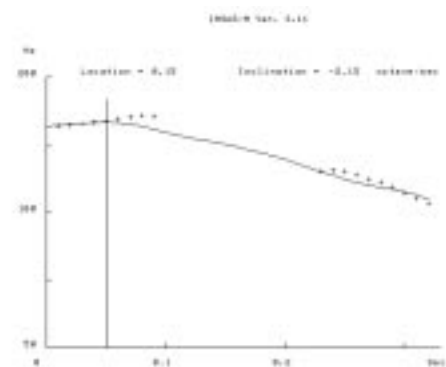


Figure 6: An example of Japanese word accent 1.

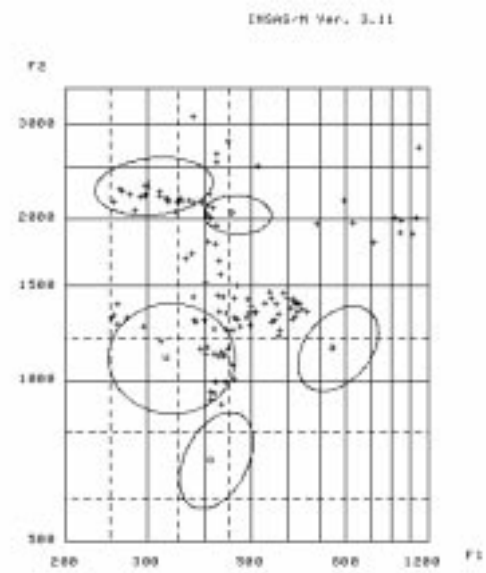


Figure 9: An example of F1-F2 plane for Japanese vowel analysis.

#### 4. EVALUATION OF THE SYSTEM

This system is open on March 5th, 1999. Now, total no. of access counts is about 12,000 for a top page. Fig. 12 shows access counts for 16 months at the Web page (<http://sp.cis.iwate-u.ac.jp/sp/paste/>).

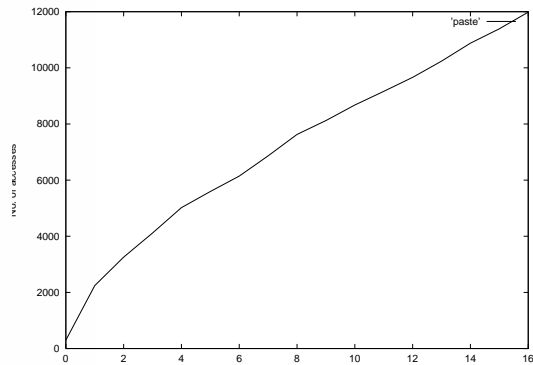


Figure 12: No. of total accesses at the top page.



Figure 10: An example of F1-F2 plane for English vowel analysis.

#### 5. CONCLUSION

This paper described a system for speech analysis on the Internet using E-mail and Web technology called as INSAS/M.

A user's speech waveform attached in E-mail is analyzed and converted to an image file with the system. Then, the user can easily see the analyzed result in a web page with a browser.

The system can be used at any time, in anywhere and for anyone so that it is very suitable for a phonetician and a busy language learner in the world.

Now, total no. of access counts is about 12,000 for 16 months at the top page. But it is necessary to improve the system.

#### ACKNOWLEDGEMENT

This project is partially supported by the grant-in-aid for scientific research(12558019) in Japan Society for the Promotion of Science.

#### REFERENCE

- [1] Masaru Kumagai and Jouji Miwa: "Public Analysis System for Speech Using Techniques of E-mail and Web", Report of Speech Technical Committee, The Institute of Electronics, Information and Communication Engineers (IEICE), SP99-21, pp.25-32 (June 1999). (in Japanese)
- [2] Jouji Miwa, "A Practical Introduction to Speech Processing Using Personal Computer", Shokodo, Tokyo, pp.1-186 (July 1991). (in Japanese)

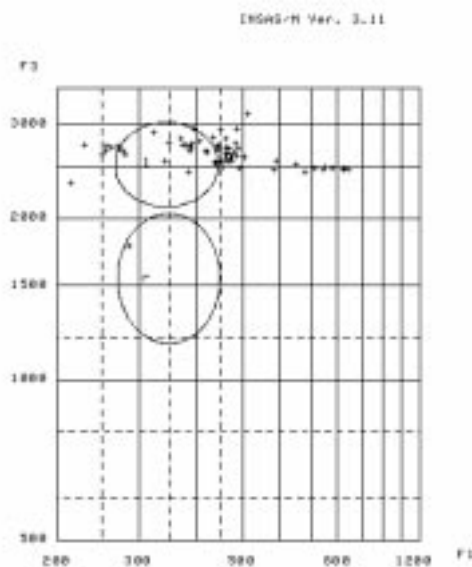


Figure 11: An example of F1-F3 plane for /r/ and /l/ learning.