Acoustic Parameters Database of Vocal Characterization for Study of Auditory Pathologies

Ma. Julia Calderon Sambarino[1], Sergio Suárez Guerra[1]

[1] Computing Research Center, National Polytechnic Institute,
Juan de Dios Batiz s/n, P.O. 07038, Mexico
julia.sambarino@gmail.com, ssuarez@cic.ipn.mx

Abstract. The voice can be characterized in an objective way by means of extraction and analysis of its parameters. The voice acoustic parameters allow identifying patterns of normality and using them to recognize pathologies related to tract-vowel system. This paper presents the summary of a voice library for extraction of parameters and its storage in database. From them, we show a preliminary study of auditory pathologies using the fundamental frequency and their variations.

Keywords: Acoustical parameters, hypoacusic, voices library, F0, jitter, variation coefficient, vocal characteristics database.

1 Introduction

Acoustic voice analysis allows extracting objective parameters that identify it. The algorithms for extraction of such parameters are implemented by software, however an alteration of one or more parameter may its origin in different and difficultly identifiable physical subsystems due to a non lineal behavior of voice producer mechanism, that's why there's no exact model that define it.

Human voice has characteristics that describe it in an objective way and they allow comparing it with normality patterns. According to Wirz [1] the alterations of the voice in hypoacusic are maximum in subjects with more intense and earlier auditory losses; namely, the variation of acoustic parameters is greater.

Vocal quality from the defined the acoustic parameters are defined may be altered by lack of auditory, neuromuscular or phonorespiratory control. That is, in hypoacusics fundamental tone can be displaced to deep or acuity, intensity can be increased or diminished and tone is generally characterized by being opaque due to a poor harmonic production.

To take these characterizations into an objective plane, it does not only requires to study of pathological voices and acoustic parameters, but also of normal voices that allows us to find an index of "normality" with those define the grade of variability of tone and what it means.

To be able to identify particular characteristics that differentiate voices of subject with normal audition of those with auditory pathologies is necessary to have a previously recorded library of voices for their analysis. The most used pathological voices

© A. Gelbukh, C. Yáñez Márquez, O. Camacho Nieto (Eds.)
Advances in Artificial Intelligence and Computer Science
Research on Computing Science 14, 2005, pp. 381-390

library is Kay Elemetrics [2], but they don't have an audiometric study of the subjects, so that, we cannot associate it with auditory pathologies.

Previous works, like Kawahara, Kato and Williams [3] look for relationship of auditory pathologies with speech phenomenon but guided to feedback effect and only supported in fundamental frequency as analysis parameter. De Bodt, Hernández and Van De Heyning [4] propose relationship that exists among some parameters of vocal stability with intelligibility in language, starting from their own group of voices. Divenyi and Haupt [5] make correlation analysis among capacity to understand speech starting from audiometric tests, but they have not been generated voice library that gather medical information that gives them a universality character, not for auditory pathologies or of language studies, but like investigation reference in any area.

In this work, we talk about design, obtaining and possible use of a voice library of subject with pathologies that go from slight, moderate and deep deaf to light neurological damage, as well as subjects without pathologies and normal audition, considering in each case, the clinical file provided by a medical institution.

Starting from study of gathered voices can define investigation works like determination of voice alterations (variations in acoustic parameters) of hypoacusic subject starting from acoustic analysis in relation to normal subjects to identify, by artificial intelligence tools, existence of auditory deficiency, and other investigations that remain open.

2 Acoustic Parameters as Vocal Characterization Elements

The objective evaluation, by means of vocal stability parameters analysis, is a characteristic that can be attributed to voice since 70's [2] and it has improved diagnostic and rehabilitation, as well as works related to recognition of language and in development of tools of artificial intelligence.

The most important parameter to achieve voice description is fundamental tone expressed by fundamental frequency (F0) since it allows characterizing sound in phonation. The variations or interferences of F0, produced by neurosensorial, phonorespiratory or muscular problems are reflected in vocal emission. From F0 we can measure stability and carry out statistical calculations that characterize voice with more precision (modal, whisper, rough, etc).

An important derived parameter of F0 is Jitter, which measures uncertainty of frequency of fundamental tone, namely, is a measure of variability for period of time of F0, cycle-to-cycle, during its production.

From F0, cycle-to-cycle measurement, we can compute, among other statistical parameters, standard deviation and variation coefficient regarding F0 mean, and variation percentage and interference average, to make a deeper analysis of frequency and amplitude, to obtain a formal characterization of voice production.

Summary of vocal characterization parameters calculated to create an acoustic parameters database from voices library is presented in Table 1 and the formulas used to obtain some of the important parameters are presented in Table 2.

The voice parameters were obtained with EXPARAM system, which can extract values described with a sign of at least 20 mseg, but for a real analysis of fundamental tone and its variability, it is recommended to have a voice sign of at least 2 secs.

Table 1. Parameters calculated for acoustic parameters database from voices library.

Vocal characterization parameters	Description
F0 mean	Fundamental frequency of sign under study. Global measure of voice height.
F0 maxim	Higher value of fundamental frequency F0
F0 minimum	Lower value of fundamental frequency F0
Standard deviation	Width of variation in Hz of F0 with regard to mean.
Variation coefficient	Allow us relative standard deviation in comparison with F0mean.
Jitter mean	Measure of sum of difference between two consecutive cycles of vibration.
Jitter factor	Allow us relative Jitter mean in comparison to F0mean.
Jitter ratio	Mean of variations of periods among consecutive vibration cycles what is relative to value of half period of observed sign.
Perturbation relative average	Mean of variation of three serial periods with relationship to half period of observed sign.

Table 2. Formulae used for acoustic voice parameters extraction

Parameters	Formula
F0 mean where $F0_i$ is F0 cycle-to-cycle interval	$F0_m = \frac{1}{N}\sum_{i=1}^{N} F0_i$
Standard Deviation	$Desv.est. = \sqrt{\frac{1}{N}\sum_{i=1}^{N}(F0_i - F0_m)^2}$
Variation coefficient	$c.v. = 100 * \frac{Desv.est}{F0_m}$
Mean Jitter	$Jitter_m = \frac{1}{N-1}\sum_{i=1}^{N-1}\lvert F0_i - F0_{i+1}\rvert$
Jitter factor	$Jitter_{factor} = 100 * \frac{jitter_m(Hz)}{F0_m}$
Jitter ratio	$Jitter_{ratio} = 1000 * \frac{jitter_m(ms)}{T0_m}$
Perturbation relative average	$Promrelpert = \frac{\frac{1}{N-2}\sum_{i=2}^{N-1}\left\lvert \frac{T0_{i-1} + Tq_i + Tq_{i+1}}{3} - T0_i \right\rvert}{T0_m}$

3 General Purpose Voice Library

According to the voice characteristics, we can extract reliable parameters from sustained recording of /a/ vowel.

Kay Elemetrics [2], one of the most used voice library, outlines two types of recordings, one with a paragraph reading and another with vowel / a / emission. However, Kay Elemetrics looked for to relate voice parameters with speech pathologies and physical problems that rotated around it that is why clinical files are only based on organic and physical characteristic of vocal tract characteristic.

Our position intends to include a voice coming from any pathology, from light neurological damage, physical located damage, to deaf, that is why voice library allows studying interferences coming from different causes. The difficulties in obtaining samples are basically: 1) to have a clinical file emitted by a medical institution that identified pathology and 2) that subject was capable of interact in obtaining a satisfactory sample. As library also includes normal voices, university student's volunteers were looked for not with family antecedents neither defined pathologies.

The procedure for obtaining the library consisted in several steps; the first one was experiment explanation. Later on we carried out capture from relative data to clinical file. Distances and angles were adjusted between subject and microphone and recording began. We allowed the subject to rest for five minutes and repeats recording until completing at least two successful recordings or a maximum of five fail intents. It is considered a successful recording when a vowel was emitted at least during three seconds in a sustained way. In figure 1 distribution of sample is shown by gender.

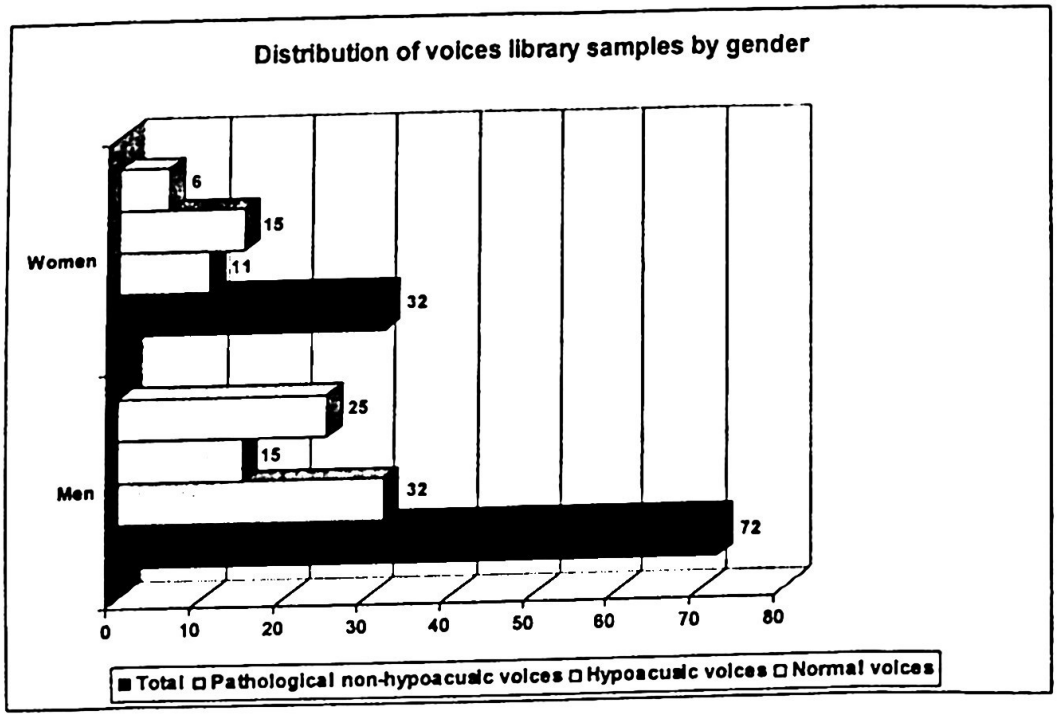

Fig. 1. Distribution of voices library sample by gender

To achieve a sample, clinical data was required using a template designed for such purpose according with the Mexican Official Norm for clinical file. Recordings were

carried out with a digital recorder and an unidirectional microphone with a frequency range between 100 and 17 000 Hz located at a distance lower than 50 cm. and not smaller to 20 cm. of subject's mouth using a pedestal to regulate an appropriate inclination to each subject. Recordings were transferred to a computer using an adapting audio cable connected to a sound card and stored in wav format with a sampling frequency of 44,100 Hz, 16 bits mono.

Lastly we built the database with tables for clinical files of those subject with pathological voices and another for normal voices, they were also related files recorded with subjects and their pathologies. In figure 2 shows final distribution obtained with a total sample of 104 subjects. Figure 3 shows the database design that incorporates the voices including a table with relative acoustic parameters to each sound file.

The total of gathered voice files was analyzed to define if they fulfilled defined characteristics for its use, and they should be eliminated when coming from subjects that were not able to support vocal emission for established time or they could not cooperate appropriately. The final distribution of obtained files is shown in Figure 4.

4 Work in Progress

One of direct sequels of auditory loss, mainly in event of severe hypoacusics is a problem of acquisition of language. But when deaf is light or moderate, quality of vocal emissions, as well as reception of same ones are altered, this is due to deterioration in auditory feedback process (feed-back). We have defined voice of hypoacusic like it tightens, page, lacking, whispering, rough, monotonous, rhythm less and high tone [1].

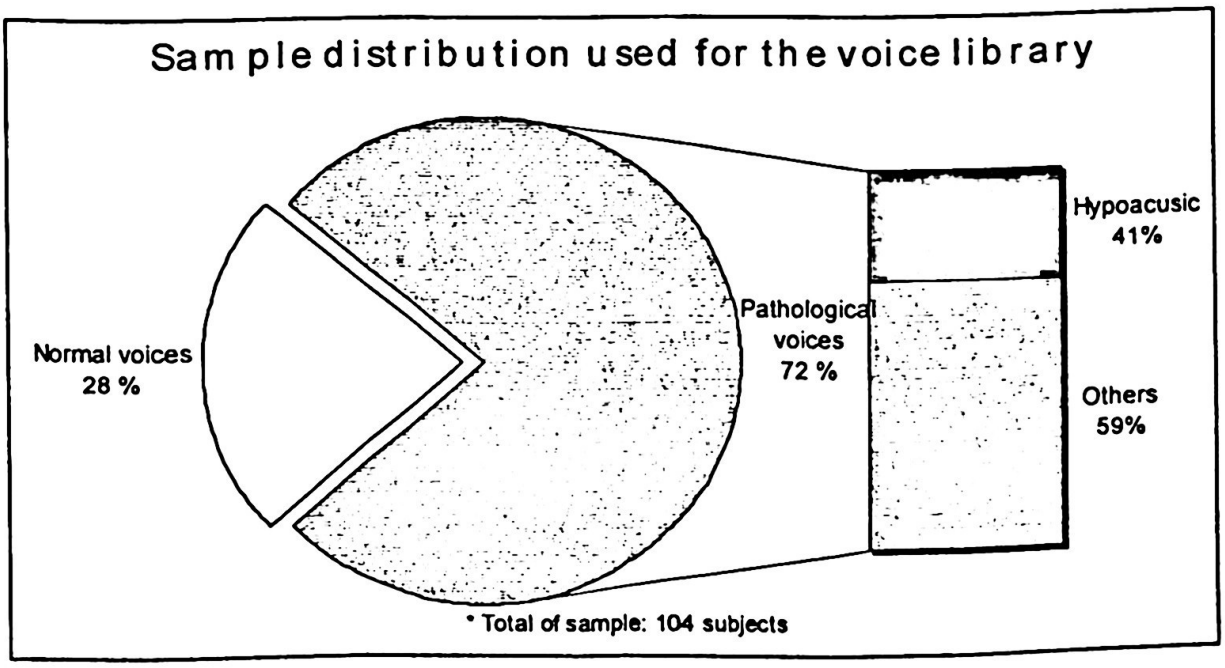

Figure 2. Sample distribution used for obtaining library of voices

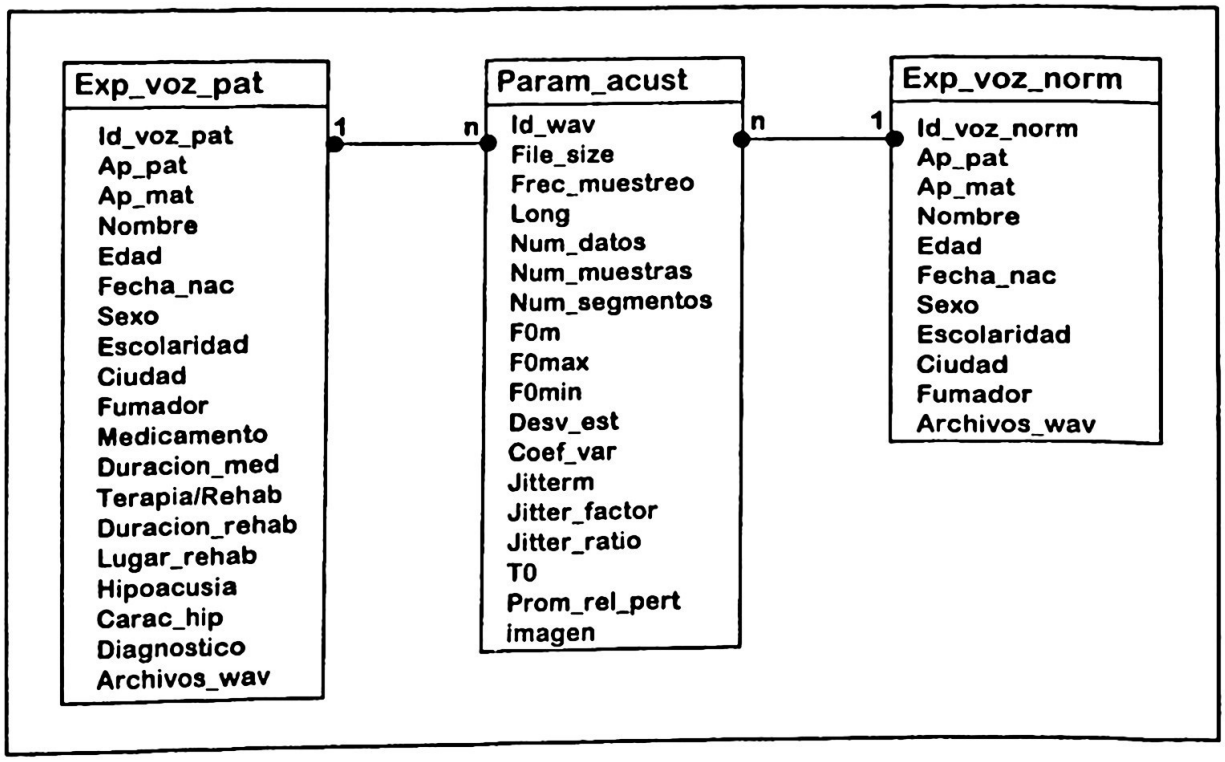

Figure 3. Database design incorporates library of voices and acoustic parameters of each file

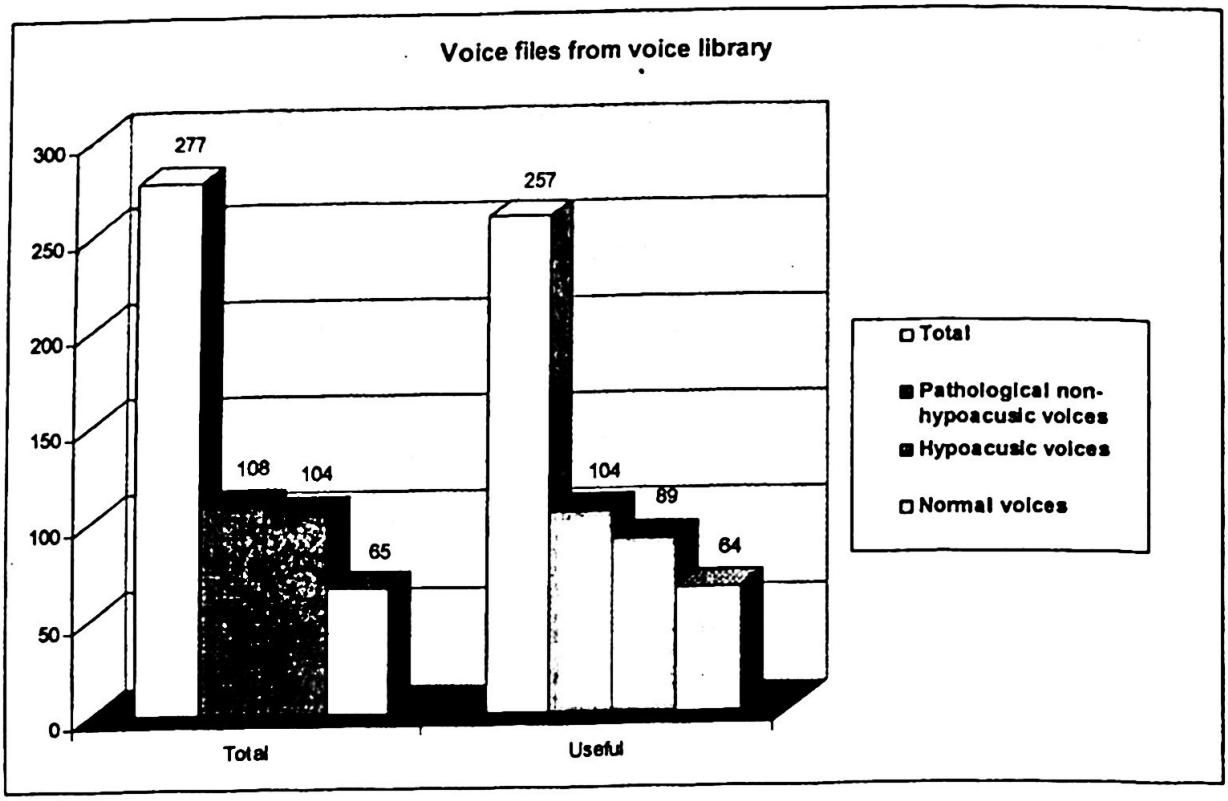

Figure 4. Total distribution of 277 obtained voice files of a universe of 104 subjects.

To define these subjective characteristics, we outline as proposal identify parameters obtained from voice acoustic analysis that are related with existence of auditory problems obtained from clinical analysis using voices library and tools of artificial intelligence to correlate them. We believe there is some kind of relationship between parameters of language (voice characterizations) and auditory deficiency.

One of the goals of this work is to provide to specialists an alternative of diagnostic and a tool that supports this diagnosis, besides opening investigation field for procedures of use of voice acoustic analysis inside diagnosis and rehabilitation of subject with audition and language problems.

Analyzing the voices of the library, we found clear differences between subject with auditory pathologies and subject of normal audition we use Exparam [15]. Figure 5 shows analysis carried out on a hypoacusic subject and a normal audition subject.

As it is seen in Figure 5, graph of left side, shows stability in voice generation, according to fundamental frequency, while graph of right side, shows a series of irregularities originated by flaw in feedback process that hearing provides for voice generation. The data shown are confirmed in the calculations presented in figure 6, where superior window shows resulting data of Figure 5 graph of a normal audition subject, while inferior window shows calculations that were of graph of hypoacusic subject.

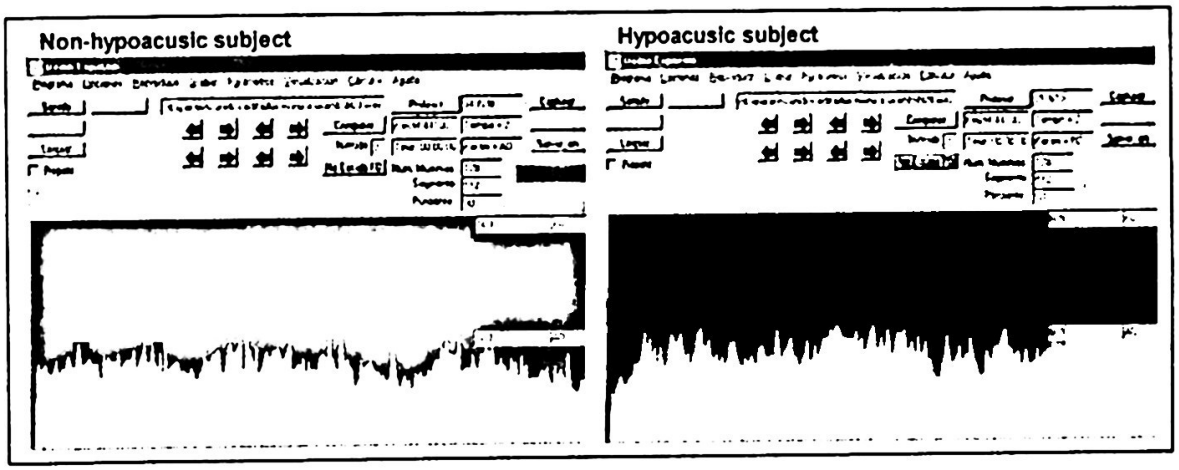

Figure 5. F0 representation of a normal audition subject, and of a hypoacusic subject, both cases with emission of /a/ vowel.

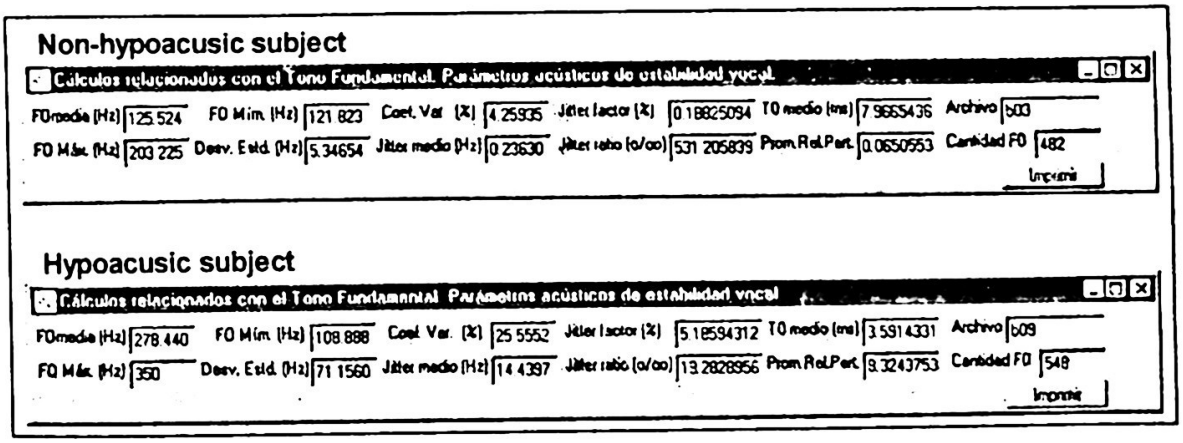

Figure 6. Calculations obtained from voice analysis in same subjects shown in figure 5.

According to the figure 6, it stands out difference between a F0 125.5 in a normal audition subject, and one of 278.4 in a hypoacusic subject, as well as coefficient of variability of 4.25 in a normal voice and of 25.5 in a hypoacusic voice. Regarding mean jitter, normal audition subject possesses a value of 0.236, while subject with pathological voice sample a mean jitter of 14.43.

As a preliminary analysis, we can conclude that in normal voices of the library, 73% of subjects' presents a smaller coefficient of variability at 10, while in hypoacusic voices, 27% hardly presents a smaller coefficient of variability to 10 and 61% it presents it between 10 and 30. Regarding mean jitter, we takes as normal values below 3 [6] that is proven in normal voices library where 78% of sample is presented below this value, while in the deaf pathological voices, 38% is only adjusted to the normal range, and 62% it surpasses this values, like it is shown in Table 3.

Table 3. Analysis of two parameters of vocal stability obtained starting from voices library.

Variation coefficient						
	Normal voices		**Pathological non-hypoacusic voices**		**Hypoacusic voices**	
Range	*Value*	*%*	*Value*	*%*	*Value*	*%*
<10	47	73.4	51	49	24	27
> 10 < 20	5	7.8	32	30.8	35	39.3
> 20 < 30	7	10.9	15	14.4	21	23.6
> 30 < 40	5	7.8	6	5.8	9	10.1
> 40	0	0	0	0	0	0

Mean jitter						
	Normal voices		**Pathological non-hypoacusic voices**		**Hypoacusic voices**	
Range	*Value*	*%*	*Value*	*%*	*Value*	*%*
<3	50	78.1	39	37.5	34	38.2
> 3 < 10	8	12.5	36	34.6	32	36
> 10 < 20	4	6.3	18	17.3	19	21.3
> 20 < 30	2	3.1	7	6.7	4	4.5
> 30 < 40	0	0	1	1	0	0
> 40	0	0	3	2.9	0	0

In previous chart 21.9% of cases is presented in that normal voices overcome value of 3 in mean jitter, this can originate for problems of noise in voice sign, problems of voice production like age or vocal training, as well as non critical problems in phonologic system, but they not necessarily outline a defined pathology.

5 Conclusion

Voice analysis in diagnosis and medical rehabilitation field is a complex subject, so that, constitutes an investigation area in development. Acoustic parameters of vocal

stability allow us to find relationships with defined pathologies. Therefore, it is necessary to have a voice library, besides pathological voices, normal voices, to find patron and relationships among them.

Simple pathologies rarely originated by a single cause are presented, in must of cases exist one or more problems related to the same pathology, this indicates a difficulty that turns out to be able to pre-establish premises, models and results.

In this work a voice library is presented that gathers 104 subject, with normal and pathological diverse characteristic. Of total of subjects they were obtained two voice files at least with emission of / a/ vowel, which were revised to determine if they fulfilled the duration requirements and sign quality. At end, voice library conformed to for 257 files of normal and pathological subject voices. All files were analyzed with Exparam system for parameters of vocal stability parameters, with which was integrated a database.

Among existent pathologies in those subjects of study, we defined hypoacusics like analysis matter, for what segmented sample and they were carried out preliminary analysis of vocal characterization acoustic parameters results, finding relationships between coefficients of variability and mean jitter.

Acknowledgments

For summary of voices library we thank participation of:

- Directive and students of Normal School of Special Education "Graciela Pintado de Madrazo". Vhsa. Tabasco, Mexico.
- Directress Mtra. María Valentina Contreras Pérez, teachers and students of Multiple Attention Center No. 12. Frontera, Tabasco, Mexico.
- Directresses Psic. Ma. del Pilar Rovira Balán, Mtra. Germana Emma Balán de la Cruz, personal and students of Psycopedagogic consultantship Center S.C. Vhsa. Tabasco, Mexico
- Students of Superior School of Computation. National Polytechnic Institute. Mexico City

This investigation has support of the national council of science and technology in Mexico CONACYT, through the program 35533-A and of National Polytechnic Institute, through Institutional Program Investigating Formation PIFI 2005779.

References

1. Wirz S. The voice of the deaf in Voice Disorders and their Managenment, 2nd, Singular, San Diego, ed M Fawcus ,Ca. 1992.
2. Kay Elemetrics Corp. Disorders Voice Database. Massachusetts Eye and Ear Infirmary, Voice and Speech Lab. Boston, Ma. USA 1994
3. Kawahara, H., Kato, H., Williams, J. Effects of auditory feedback on F0 trajectory generation. Human Information Processing Research Laboratories.

4. De Bodt, M., Hernández-Díaz Huici, M., Van De Heyning, P. Intelligibility as a linear combination of dimensions in dysarthric speech. Jorunal of Communication Disorders. USA. 2002
5. Divenyi,P y Haupt, K. Audiological correlates of speech undestanding deficits in elderly listeners with mild-to-moderate hearing loss. Ear and Hearing Vol. 18. pp 100-113. 1997
6. Ingo R. T. Workshop on Acoustic Voice Analysis. National Center for Voice and Speech. USA 1994.
7. Series Mundo electrónico. Inteligencia Artificial, conceptos, técnicas y aplicaciones. Boix-areu Editores. España 1987
8. Russell, S., Norvig, P. Inteligencia Artificial, un enfoque moderno. Ed. Prentice Hall. México. 1996.
9. Organización Mundial de la Salud. Guía para escribir un protocolo de investigación. http://www.asalfa.org.ar/htm/articulos/art_19.html
10. ISO. ¿What are standars? Http://www.iso.ch/iso. 2002
11. Secretaría de Economía. Normas Mexicanas. Normal oficial mexicana del expediente clínico. http://www.economia-nmx.gob.mx. 2002
12. Malcom J. Crocker. Handbook of acoustics. Ed. Wiley-Interscience. USA. 1998.
13. Calderón, J. Sist. Para el estudio, medición y eval. del conducto auditivo ext. del oído mediante estímulos sonoros. Tesis de Maestría en Ciencias. CIC- IPN. 2003.
14. Meléndez, J. Colomer, J. Inteligencia artificial. Automática e Instrumentación. No. 340. Mayo 2003.
15. Suárez S. Exparam. Voice analysis software. México. 2005