A GERMAN DATABASE OF PATTERNS OF PATHOLOGICAL VOCAL FOLD VIBRATION

Manfred Pützer & Jacques Koreman

Abstract

This paper presents a survey of a database of pathological voice qualities, collected in collaboration with the Department of Phoniatics and ENT at the Caritas clinic St. Theresia in Saarbrücken. The major part of this paper is devoted to a medical description of pathological voice qualities, but the recording procedure as well as the goal of the database collection will also be described. In particular, we hope that in the future it will be possible, on the basis of the present database, to relate properties of recorded voice signals to phoniatic diagnoses of voice pathologies.

1. Introduction

In this article, we shall present the collection of a database of pathological voice qualities, a description of the voice qualities it contains and the possible uses of the database. The database was collected in a collaborative project of the Department of

Phoniatics and ENT at the Caritas clinic St. Theresia in Saarbrücken and the Institute of Phonetics of the University of the Saarland. So far, 95 organic and non-organic pathological voices have been recorded. New recordings of pathological voices will be added to the database with the continuation of this project. The aim of the project is to define the relevant properties of different pathological voice types.

2. Methodology

The collection of the database has combined research methodologies from speech science with phoniatic methods. Methods from speech research which were used are electroglottography (EGG) and recording of the sound pressure waveform (microphone signal). Both signals were recorded onto DAT tape simultaneously in a quiet room. The signals were recorded for a read text and for the vowels /i:/, /a:/ and /u:/ at normal, high and low pitch. The phoniatic investigation consisted of video recordings of the vocal folds, using a laryngoscope and additional stroboscopy. In addition to these measurements, the patients were asked to fill out an information form with their name, age, sex and smoking habits; the phoniatic diagnosis from the standard investigation was also noted down on the information form.

3. Contents of the database

The database consists of recordings of different types of voice pathologies, which we shall list below. We can divide the voice pathologies into organic (section 3.1.) and non-organic voice pathologies (section 3.2.). Within the organic voice disorders, benign pathologies (section 3.1.1.) are distinguished from malignant ones (section 3.1.2.). The former can be neurological (section 3.1.1.1.) or non-neurological (section 3.1.1.2.). The non-organic voice disorders or functional dysphonias in section 3.2. can be split up into hyper- and hypofunctional dysphonia, phononeurosis and dysodia (Biesalski & Frank, 1994; Boenninghaus, 1993; Wirth, 1995). In addition, two voices with combined symptoms have been recorded into the database.
3.1. **Organic voice disorders**

3.1.1. **Benign voice pathologies**

3.1.1.1. Neurological

*a. Vocal fold paralysis (presently, 14 speakers)*

Vocal fold paralysis is caused by paralysis of the recurrent nerve (inferior laryngeal nerve) and often occurs as a complication in thyroid operations. It can occur as a one-sided or double vocal fold paralysis. Since the inner laryngeal muscles do not function, the paralyzed vocal fold(s) cannot be brought into phonation position. Instead, the exterior laryngeal muscle (cricothyroid muscle), which is not affected by the paralysis, draws the vocal fold into median or paramedian position by tension towards the middle line. In the case of paralysis of both the inferior and superior laryngeal muscles, the paralyzed vocal fold remains in intermediary position, with or without excavation. The phonation conditions deteriorate with the extent of the paralysis. If only the superior laryngeal muscle is paralyzed, phonation is possible; the paralyzed vocal fold is usually lax and shows a slight excavation of the free rim of the vocal fold.

*b. Parkinson's disease (presently, 2 speakers)*

Typical of the Parkinson's disease is the inability to produce a sufficient adduction of the vocal folds, which leads to a breathy, soft, monotonous voice quality. A tremor of the vocal folds and possibly of the epiglottis, velum and tongue is often present in addition to an adduction deficiency. Phonation duration is short.
3.1.1.2. Non-neurological

a. *Chronical laryngitis (presently, 8 speakers)*

Chronical laryngitis typically shows in varying harshness or breathiness. The inflammation of the vocal folds and a layer of epithelium reduce the ability of the vocal folds to vibrate, lowers the average pitch and reduces higher frequencies. The swelling of the vocal folds and their irregular thickness reduce the mobility of the edges of the folds, leading to irregular vocal fold vibration or even an inability to phonate.

b. *Reinke oedema (presently, 4 speakers)*

The Reinke oedema of the vocal folds is a special form of chronic hyperplastic laryngitis, characterized by the collection of oedema fluids directly under the vocal fold epithelium. The voice, which has a pressed quality, is harsh or breathy and has an exceptionally low pitch ("sexy voice"). The low pitch is caused by the flapping movements of the hyperplastic mucosa, which is blown apart during phonation. Closure is irregular and caused by contact of the hyperplastic mucosa instead of the real vocal fold masses. The closure phase is relatively long.

c. *Vocal fold nodules (presently, 5 speakers)*

Vocal fold nodules, for instance caused by continuous phonation at high loudness levels (shouting by children or singing) indicate abuse of the vocal folds. In this case, the basic symptomatic of a hyperfunctional dysphonia exists. The presence of nodules is related to a breathy and harsh voice quality. Large vocal fold nodules, especially if they lie exactly opposite each other on the vocal folds, cause an incomplete closure of the vocal folds in front of and behind the nodules ("hourglass glottis"), reduce the mobility of the vocal fold edges and lead to a smaller amplitude of the vocal fold vibrations.
d. **Leucoplacia and pachydermia (presently, 10 speakers)**

Both in the case of leucoplacia and of pachydermia we are dealing with hyperplasia of the epithelium layer of the vocal folds, leading to greater hornification in the top layers of the epithelium.

The ability of the vocal folds to vibrate is directly related to the thickness of the epithelium: the thicker and therefore stiffer the epithelium, the more restricted the mobility of the vocal fold edges, resulting in a reduction of the ability of the vocal folds to vibrate. The voice is characterized by a low pitch and a harsh or breathy voice quality.

e. **Benign tumours**

**aa. Chondroma (presently, 1 speaker)**

A chondroma is a tumour in the cartilagenous parts of the larynx, usually starting at the cricoid cartilage. This subglottal curvature is covered by an unchanges smooth mucosa and initially creates harshness and problems with breathing (dyspnoea) and swallowing. With increasing size of the tumour, the vibration of one or both vocal folds may be affected, either by the mechanical changes in the vocal fold(s) or because the tumour causes a lesion of the recurrent nerve.

**bb. Laryngocele (presently, 1 speaker)**

A laryngocele is an expansion of the Sinus Morgagni, which can be directed inwards (towards the larynx) or outwards (between the edge of the thyroid cartilage and the tongue bone). If the laryngocele extends into the larynx, it pushes the false vocal fold or the aryepiglottic fold forward, causing harshness, breathing problems and problems with swallowing.

### 3.1.2. Malignant pathologies

The malignant pathologies are divided into hypopharyngeal and laryngeal carcinomas (Kleinsasser, 1988).
a. Hypopharyngeal carcinoma (presently, 2 speakers)

The hypopharyngeal carcinoma is found in the piriform fossa, in the posterior wall of the hypopharynx or in the postcricoid space. It can extend into the larynx and thyroid, in which case it can cause a harsh voice quality.

b. Laryngeal carcinoma

aa. Supraglottic carcinoma (presently, 4 speakers)

A supraglottic carcinoma is positioned in the laryngeal epiglottal area, the false vocal folds or the laryngeal ventricle. It initially causes a feeling of pressure in the larynx. When it begins to affect the vocal folds, it leads to a harsh and breathy voice quality.

bb. Vocal cord carcinoma (presently, 9 speakers)

The relatively early appearance of symptoms (harshness) usually enables an early start of the treatment of vocal fold carcinomas. One of the vocal folds has a deep red colour, is thick, irregular and ulcerous. Initially its mobility is not affected. If it is, this indicates that the carcinoma extends towards the arytenoids.

c. Subglottic carcinoma (presently, 1 speaker)

Subglottic carcinomas are rare. They cause breathing problems and harshness if the vocal folds are affected directly.

3.2. Non-organic pathologies (functional dysphonias)

Functional dysphonias are characterized by disturbances of the voice quality and of the vocal achievements, without any observable organic changes in the larynx. After a while, however, functional pathologies can lead to secondary, organic changes.
a. **Hyperfunctional dysphonia (presently, 9 speakers)**

Hyperfunctional dysphonia, the most common functional dysphonia, can have multiple causes: the patient may have a weak constitution, which affects the efficiency of his or her voice; speaking habits, e.g. ineffective use of the voice, may play a role; or straining the voice by speaking with too much effort may cause hyperfunctional dysphonia. Psychogenic factors, related to psychosocial stress and depression for instance, are also a possible cause of hyperfunctional dysphonia. The symptoms of a hyperfunctional voice are a harsh, weak, pressed or even diplophonic voice. The voice quality also shows breathiness, which depends on the vocal effort and is typically present at least in later stages of hyperfunctional dysphonic voices. The voice is often loud and has a high average pitch. Hard voice onsets and a poor articulation of vowels (too dark) are part of the pathological symptoms. Often the loudness modulation is very limited, while the frequency modulation is relatively good. The articulation is characteristically close, with minimal jaw and lip movements.

b. **Hypofunctional dysphonia (presently, 3 speakers)**

Hypofunctional dysphonia is in most cases caused by weak laryngeal muscles, leading to incomplete closure of the vocal folds. The voice is breathy, noticeably weak and shows little dynamic modulation, especially with regard to pitch modulations. Hypofunctional speech is therefore monotonous. Due to the incomplete closure of the vocal folds, more air is used ("wild air"). Stretches of phonated speech are short and many breath pauses are necessary. Average pitch is high and voice onsets are weak and breathy. Secondary hypofunctional dysphonia can be caused by a general weak physical state.

c. **Phononeuroses**

Phononeuroses are caused by acute or chronical psychological disturbances or to compensate a primary hyperfunctional dysphonia.

aa. **Psychogenic dysphonia (presently, 7 speakers)**

Psychogenic dysphonias, which show a hypo- or hyperfunctional laryngeal mobility ("motility"), show a characteristic breathiness which is independent of vocal effort. Also characteristic are quick changes between pressed and breathy voice. Like
psychogenic aphonia, psychogenic dysphonia can occur after a sudden fright, stress or seemingly unbearable psychological strain.


bb. *Spastic dysphonia (presently, 6 speakers)*

Spastic or spasmodic dysphonia is a disturbance of phonation due to glottal cramps, which mostly affects speech, but not isolated voicing. Deep neurological problems are no longer taken to be the cause of spastic dysphonia ref. Strong spastic cramps of the breathing and voice muscles during speech lead to an interruption of voicing, when the muscle tension and the extremely high subglottal pressure can no longer be kept up. The voice sounds tortured and crushed Laver 80 oder Vieregge 96 and can suddenly change to falsetto or into dyplphonia.

d. *Dysodia (presently, 6 speakers)*

Dysodia is a functional pathology of a singer's voice, which is characterized by a disturbance of the voice quality and by a limited vocal efficiency. The natural or acquired vocal capabilities are lost and the voice is easily strained. Voice onset is difficult and weak. Additional symptoms are breathiness and a hoarse phonation. No voicing can be produced at a low pitch and higher pitches can only be produced with much effort. Loudness modulations and flexibility of the vibrato voice are limited. The voice is easily fatigued.

e. *Mixed symptoms*

aa. *Mutation falsetto (presently, 1 speaker)*

Mutation falsetto voices have a higher pitch than children's voices. They normally only occur for boys, and the falsetto voice is higher than before the mutation. In a few cases, deep humming tones, harshness and diplophonia occur. The cause of the pathology can be local, sensorial or psychological.

bb. *Glottal sulcus (presently, 2 speakers)*

A glottal sulcus limits the mobility of the vocal folds. There is a groove in the middle of the vocal folds on one or both sides, dividing the vocal fold(s) into an upper and a lower half. The changed form of the vocal fold(s) causes more turbulence when the
A German database of patterns of pathological vocal fold vibration

air flows out through the glottis, changing the speaker's voice quality. Besides congenital forms, glottis sulcus can also be caused by inflammations or it may be the final state of the vocal folds after hyperfunctional dysphonia.

4. Future work with the database

The database has been collected with the aim to relate properties of the recorded signals to the different diagnoses and to the related auditory symptoms. Since we want to find such relationships through statistical analysis (see Laver et al., 1992; Ptok, 1990), it is very important to have a large number of speakers to represent each voice pathology. In addition, healthy (male and female) speakers have been recorded as a baseline for comparison. Further voice signals will be collected in a longitudinal study.

In a first step, an auditory judgment of the recorded voices will be made by a phoniatrician and a phonetician. The phoniatric judgment consists of an investigation of the larynx and of the vocal folds. The phonetic judgment will consider the different categories of voice quality which are described by Laver (Laver, 1980) and Vieregge (Vieregge, 1996).

In a second step, a separate analysis of the acoustic and electroglottographic signals is provided. The EGG signal will be analyzed using the EGG program from CSL; the microphone signal will be analyzed using CSL’s MDVP (multi-dimensional voice program). The resulting analysis parameters from the two programs will be evaluated as to their usefulness for the statistical distinction between voice pathologies, as well as to their physiological interpretability. A pilot study with this goal which compares speaker groups with different degrees of vocal fold adduction is presented in this volume (Koreman & Pützer, 1997). It is likely that in a next step, other algorithms need to be developed to supplement the available EGG and MDVP parameters.

Different analyses of the signals will be made to find the most relevant parameters to support the phoniatric diagnosis and to measure the effects of therapy.

Results from laryngoscopy and stroboscopic video recordings will be compared with the parameters computed from the microphone and electroglottogram signals. The parameter values will be compared with the distribution of similar data for
healthy male and female speakers. At a later stage, we intend to use a neural net for
the classification of voice pathologies on the basis of the signal parameters. Such a
system could be used as a support tool for routine check-ups by a phoniatrician or to
evaluate the effectiveness of therapy.

5. Acknowledgments

The authors would like to thank Dr. Manfred Just from the Department of Phoniatics
and ENT at the Caritas clinic St. Theresia in Saarbrücken for discussions and for
diagnosing the speakers collected in the database.

6. References

Stuttgart: Georg Thieme Verlag.

Auflage).* Berlin: Springer-Verlag.

Thieme Verlag.

deficiencies. *PHONUS 3.* Saarbrücken: Institut für Phonetik, Universität des
Saarlandes.


Stimmqualität nach Kehlkopfoperationen bei Malignompatienten. *Sprache
Stimme Gehör* 14, 11-17.