# Acoustic Analysis for Chest-to-Head Register Transition in Singing Voice

Reda Elbarougy Department of Computer Science Faculty of Computer and Information Sciences, Damietta University New Damietta, Egypt

# ABSTRACT

The most common registers discussed in the singing are chest and head register. The proficiency of singers is measured by how smoothly they cross from one register to another, the more smoothly transition is the best the singer is. The purpose of this study is to investigate the acoustic characteristics of the chest-to-head register transition region. In addition, to describe in objective terms the acoustic mechanisms employed by singers during transition region. Therefore, spectral features that reflects voice quality were used, these features are: the first four harmonic amplitudes H1, H2, H3, H4. In addition, vocal tract formants F1 and F2 were used. However, there are many drawbacks for measuring spectral features especially for singing voices, due to, it is sensitive to different pitches and sensitive to different vowel qualities. To overcome these problems a feature extraction scheme was proposed. To validate the proposed method, a singing database of singing utterances for vowel [a] recorded by three groups of singers; Altos, Mezzo, and Soprano was used. Singers sung each utterance such that they start with sung vowel [a] using chest voice then change to head voice at the end. Extracting the 6 features and observing the change happened near the transition area. It was found that, near the higher edge of the chest register, the characteristic feature for vowel [a] is: Tuning of F1 on H2 and of F2 on H4. Thus, H2 and H4 are the prominent harmonics of the chest register's range. During the register transition, the tuning of F2 shifts from H4 to H3, leading to a characteristic reduction of the level of H4, together with an increase in that of H3 in the head register.

## **General Terms**

Speech processing, acoustic mechanisms, pattern recognition.

# Keywords

Register transition; spectral patterns; chest register, head register; singing voice.

# 1. INTRODUCTION

Registers in the singing voice have been considered a series of consecutive and homogeneous tones going from low to high [1], produced by the development of the same mechanical principle, and whose nature differs essentially from another series of tones equally consecutive and homogeneous produced by another mechanical principle [2, 3, 4]. Thus, register is a sequence of tones produced by a particular vibratory pattern of the vocal folds in the human voice [5]. Therefore, all the tones within the same register are subsequently of the same nature. To sum up, vocal registers are perceptually distinct regions of vocal quality which can be sustained over some ranges of pitch [6, 7].

Registers may be recognized by the production mode, by perception of listeners, and often in the singer 's self-

perception [8]. The changing in pronunciation for voice quality in vocal register breaks leads to perception of different vocal registers. The change could be smooth, or sudden. [9, 10,11] Classical educated singers are characterized by smooth change, because they use mixing of nearby registers. In this case, the smooth change is challenging to be perceived. While, abrupt change happens in some cases such as modern style singers, or non-singers. The best instances for register breaks are vocal jumps in young men suffering from voice mutation.

The most common registers discussed in singing are chest register, and head register [12]. The chest register is usually used in normal speech and encompasses lower frequencies. Head register is used to sing notes with F0s much higher than that of normal speech [13], in the range of 400-800 Hz. Singing in head register leads to the perception of a lighter, softer voice quality.

In order to realize if the singers learn how to use different register without register breaks, it is very important to distinguish sudden, or smooth change. The goal of this research is to investigate the acoustic characteristics of the register transition between chest-to-head register in a singing voice. To achieve this task, it is important to describe in objective terms the acoustic mechanisms employed by singers during this transition region.

The rest of this research paper is organized as follows. Section 2 introduces some related works, section 3 explains motivation and objectives. Moreover, section 4 presents the sinning dataset. Then section 5 gives the proposed method for investigating the characteristics of register transition. In section 6 the results of our method are introduced. Finally, section 7 describes our conclusions.

# 2. RELATED WORK

In the Western classical music tradition, discussing the proper transition between chest and head register is essential. This transition is required to reach the higher pitches by passing from the lower part to the upper part of the male full singing voice. Most untrained singers exhibit noticeable timbre differences when moving from tones from a lower range to a higher range and vice versa. The perceptible differences in register stem from physiological differences in production configuration that affect both airflow and the nature of vocal fold vibration during phonation.

The upper vocal tract, from the vocal folds to the lips, has several resonances (F1, F2,..., Fi) i.e. formant frequencies which, in normal speech, produce spectral peaks. Values of the frequencies of the first two or three of these are associated with different vowels [14]. Taff [15] found that formant frequencies of vowels do change near the register transition [16]. These changes varied significantly in degree depending on vowels being sung and singers. For example, the baritone and tenor voices usually produced comparable results with the transition point being commonly a one third lower for baritones than for tenors. Moreover, the transition points open vowels (e, o, and a) were found to be slightly higher than for closed vowels (i and u) for in both voices [15].

In the early last century, the sound homogeneity of a register was already recognized as depending on definite acoustic features. It was assumed that these features are caused by particular mechanisms of tone production as vocal fold vibration, glottal shape, and airflow, and that neighboring registers may intersection to a definite degree. Accordingly, for long period the source of registers was regard as principally to the larynx. Integrated theory on registers in 1970 was based on acoustic phonetics and spectrum analysis took into concern that not only adjustments of laryngeal, but also adjustments of additional resonance characterized the register transitions.

## 3. MOTIVATION AND OBJECTIVES

In order to answer the research question of this paper, the perceptive classification can be accompanied by measurements of voice source parameters such as glottal open quotient [1, 15, 14]. Vocal breaks at register transitions usually associated with amplitude jumps and pitch [3, 4]. In spite of broad experimental examinations using acoustic signals; [19] and various problems concerning register transitions still open: what is the objective characterization of register transition region? This could be obtained only by measurements of acoustical parameters. To achieve this goal this study, try to investigate what kind feature should be used? Answering this could help us to know if the trained singer could pass this region smoothly, do a violation during it. Voice quality is the best choice to study the change happened in this region.

## 4. SPEECH MATERIAL

A singing database of total 20 singing utterances for vowel [a] was collected from 6 singers; four Altos, one Mezzo, and one Soprano. These singing utterances contain trials of sustained pitches on vowel [a], from 440 Hz to their highest sustainable note. To check the recoded singing utterance, two listeners; one male and one female were asked to detect the register transition point between the chest and head register for the whole database. Using a listing experiment subjects were asked to determine time t0 of the transition point, the steps of listening experiment is presented in Fig. 1.

Subjects were asked to listen and try to find the location of the transition point. They can listen to any utterance two or three times, until they can distinguish the transition point. Finally, after detecting the transition points for each utterance in the database. To concentrate on the transition point it is important to focus our analysis on a short speech segment. Therefore, in this study, a speech segment of 3 second length that contains the transition point in the middle was used, i.e. the transition point t0 is the middle of this segment. The first half of this segment is the chest register which starts at the beginning of the segment to the middle and the second half is the head register which start from the middle and ends at the end of the used segment.

The purpose of this experiment is to collect a dataset of singing utterance such that the transition point in the middle to apply some analysis and see the effect of the middle region.



Fig. 1: Steps for the experiment to prepare the dataset, listening test to detect the register transition point.

Figures 3-4 show the spectrogram for a three segments taken from the three singer groups: Altos, Mezzo, and Soprano. It is clear from these figures that; the first half is chest register and the second half is head register. The transition region is located between chest and head register. The blue dotted line shows the register transition point. It is clear from the spectrogram that this region has its specific characteristics.







Fig. 3: Spectrogram for one sample from Mezzo singers, left half is chest register, right half is head register.



Fig. 4: Spectrogram for one sample from Soprano singers, left half is chest register, right half is head register.

#### 5. PROPOSED METHOD

This section introduces the proposed method for determining the characteristics of the transition region between chest and head registers. In this section, the proposed analysis method is presented. It is important to employ the most related acoustic features for register transition. Pitch, duration and loudness are acoustic features of vocal sound that are fully developed in music. Register is not included among these parameters, but belongs rather to sound quality, which is physically represented in spectral makeup, the balance of the various overtones of a given sound. Voice quality conveys both linguistic and paralinguistic information and can be distinguished by acoustic source characteristics. Therefore, the voice quality is explained in more details in the next subsection.

#### 5.1. Voice quality related features

The register transition in sinning voices is reflected in the singer vocal utterances. Listeners know this and can recognize this transition based on vocal cues alone. Voice quality is primarily associated with the spectral properties of the speech signal. Spectral measurements can provide cues to relevant aspects of voice quality, such as H1-H2 where H1, H2 are the amplitudes in (dB) of the first harmonic, the second harmonic. Fig. 5 shows the speech spectrum and the positions for harmonic amplitudes H1, H2 and format frequencies F1, F2.



Fig. 5: Speech spectrum |S(f)|in dB, showing harmonic amplitudes: H1=|S(F0)|, H2=|S(2F0)| in dB, and format frequencies F1, F2 in Hz.

H1-H2 is an indicator of the relative length of the opening phase of the glottal pulse. H1 and H2 are related to the open quotient (OQ) using Eq. as reported by Fant in [20, 21, 22].

$$H1-H2 = -6 + 0.27 \exp(5.5 \times OQ)$$
(1)

It is expected that H1-H2 to be small and positive or negative for creaky voices, large positive for breathy voices. Moreover,

formants make a significant impact on the acoustical power of a tone, if one of the harmonics falls adjacent to its center frequency, brought about a change in this understanding of registers. Utilizing these findings, the formant frequencies F1, F2, and the first four harmonic amplitudes: H1, H2, H3, H4 were used as a spectral features of the singing voices. These features reflect voice quality to investigate the physical characteristics of the register transition from chest-to-head.

#### 5.2. Acoustic features extraction

The analysis for acoustic characteristics was done for each register individually i.e. for chest and head register, respectively. As mentioned in the introduction of the study, formant and harmonic amplitudes are the most important acoustic feature for understanding the register transition. Fig. 6 presents the proposed method for feature extraction. Firstly, formant frequencies F1 and F2 were measured by using Snack Sound Toolkit [17], using as defaults the covariance method, pre-emphasis of .96, window length of 25 ms, and frame shift of 1 ms. Moreover, the first four harmonics frequencies F0, 2F0, 3F0, and 4F0 in Hz, and their magnitudes H1, H2, H3, and H4 in dB, were measured.



Fig. 6: Diagram of the proposed feature extraction method

Then, F0 was extracted by using STRIGHT [18]. Then, harmonic amplitudes were computed pitch-synchronously, to find the maximum of the spectrum level nearby the peak locations as estimated by F0, the standard optimization techniques were used. Finally, spectral analysis using FFT was used to determine harmonic structure. Thus, H1, H2, H3 and H4 are calculated through integer multiplication of the F0 value obtained from the STRIGHT.

#### 6. RESULTS AND DISCUSSION

The acoustic characteristic for singing voice around the transition point was analyzed and extracted as described in the previous section. The trajectory of the first four harmonic amplitudes were shown in Fig. 7. This figure, shows the trajectories for H1, H2, H3, and H4, where the horizontal axis is the time in ms and the vertical axis is the harmonic amplitude in dB. The experimental results reveal that, the tarnation point is characterized by sudden decrease in the

amplitude of both H2 and H4. Moreover, and a sudden increase in H1 amplitude, and gradually increase of H3.

Fig. 8 displays the first four harmonic frequencies in the left panel and the first two formant frequencies in the right panel. From this figure, we can easily observe that in the chest register F1 is approximately equal 2F0 i.e. F1 tuning in H2 and also, F2 is approximately equal 4F0 i.e. F2 tuning in H4. While in the head register F1 decreased and no longer tuning H2 which make the amplitude of the second harmonics decreasing, and the amplitude of the first harmonics increasing at this point. Moreover, F2 and H4 have similar characteristics.



Fig. 7: The whole trajectories for the first four harmonic amplitudes: H1, H2, H3, H4 in dB for singers passing from chest to head register.



Fig. 8: Trajectories for the first four harmonic frequencies (left), trajectories for the first two formants frequencies (right) in Hz.

Therefore, we can conclude that the second harmonic H2 is the dominant harmonic in the chest register the reason behind this is because the first formant F1 supports the second harmonic H2. On the other hand, in the head register we found that the first formant is decreased and no longer support the second harmonic and begin to support the first harmonic therefore, H1 become the dominant harmonic in the head register.

The difference between the first formant frequency (F1) and the second harmonic frequency (2F0); Distance(F1,2F0), as well as the distance between the frequencies of the second formant and the fourth harmonic Distance(F2,4F0), were calculated at every time t in the speech segment as shown in Fig. 9. It is clear from this figure that the characteristics of the chest register is completely different from that for the head register. In order to extract the difference between the chest and head register, the average value for this distance was calculated for each register individually.

Let F1(t) be the value of first formant frequencies at time t in the speech segment, t=1,2,...,T, T=3000 ms is the length of the speech segment,  $t_0$ =1500 ms is the register transition point in the middle of the speech segment. Moreover, 2F0(t) be the value of second harmonic frequencies at time t in the speech segment.

Then the average value of the distance between the first formant and the second harmonic in the chest register Av\_chest can be determined by the following equation:



#### Fig. 9: The distance between the first formant frequency F1 and the second harmonic frequency 2F0(above), the distance between F2 and 4F0 (below).

In a similar way, the average value of the distance between the first formant and the second harmonic in the head register Av\_head can be determined by the following equation:

$$Av\_head = \frac{\sum_{t=t_0+1}^{T} |F1(t) - 2F0(t)|}{t_0}$$
(3)

The most important finding from our investigation is the average distance, Av\_chest, Av\_head for all speech segments for all singers, it is observed that there is common characteristic for each group Altos, Mezzo, and Sopranos. For example, for most of Altos, we found that Av\_chest, Av\_head

in both cases Distance(F1,2F0) and Distance(F2,4F0) are very close. In case of Distance(F1,2F0). It is found that the mean value for Av\_chest is 1 Hz and this mean value is decreased in the head register to -165 Hz. While in case of Distance(F2,4F0), the mean value is decreased from -126 Hz in the chest register to -386 in the head register.

Table 1. The mean value for the average distance between
the first formant frequency F1 and the second harmonic
frequency 2F0, for chest and head register.

Singer Group	Chest Register	Head Register
Altos	1	-165
Mezzo	-461	-532
Soprano	-613	-671

 Table 2. The mean value for the average distance between

 the second formant frequency F2 and the fourth harmonic

 frequency 4F0, for chest and head register.

Singer Group	Chest Register	Head Register
Altos	-126	-386
Mezzo	-977	-1147
Soprano	-1245	-1350

Finally, to show the effectiveness of the proposed method for acoustic feature analysis and see all different changes in all harmonic amplitudes. Therefore, mutual trajectories for the first four harmonic amplitudes H1, H2, H3, H4 in dB are drawn in the same figure as shown in Fig 10.



Fig. 10: : mutual trajectories for the first four harmonic amplitudes H1, H2, H3, H4 in dB.

### 7. CONCLUSIONS

The aim of this study is to find objective characteristics of the register transition region between chest and head registers. Therefore, six features were used as follows: the first two formants frequencies F1, F2 and the first four harmonic H1, H2, H3 and H4.

This study confirmed the existence of definite spectral features, for open back vowels, that characterize chest and head registers and the transition between them. In chest H2 resonated F1 and H4 resonated F2. The singers' intuitive adjustments to maintain an optimal resonance are reflected in a tendency of the F1 and F2 frequencies to increase with rising pitch in order to keep being tuned on H2 and H4. F1 follows the rising F0 only until a critical time where a register transition must be done; otherwise the highest tones cannot be reached.

During the register transition the spectral energy of H2 is reduced since it drops the resonance of F1, whereas now F2 is tuned to H3. Simultaneously, H4 decreases its amplitude due to it is not supported by F2 any longer. This pattern is not gotten in register violation's case, where F1 and F2 keep following H2 and H4.

## 8. REFERENCES

- N. Henrich, "Mirroring the voice from Garcia to the present day: Some insights into singing voice registers," Logopedics Phoniatrics Vocology, 31, 3-14, (2006).
- [2] M. Garnier, N. Henrich, J. Wolfe, and J. Smith, "Vocal tract adjustments in the high soprano range," Journal of the Acoustical Society of America, 127(6), pp. 3771-3780, (2010).
- [3] B. Roubeau, N. Henrich and M. Castellengo, "Laryngeal Vibratory Mechanisms: The Notion of Vocal Register Revisited," J. Voice, 23, pp. 425-438, (2009).
- [4] J. Svec, H. K. Schutte and D. G. Miller, "On Pitch Jumps between Chest and Falsetto Registers in Voice: Data from Living and Excised Human Larynges," J. Acoust. Soc. Am., **106**, pp. 1523-1531, (1999).
- [5] F. Marek, F. Šram, and G. S. Jan, "Voice registers, vocal folds vibration patterns and their presentation in videokymography," Proceedings of the 33rd International Acoustical Konference - EAA Symposium, Acoustics High Tatras, (2006).
- [6] G. Mark and G. Juana, "Romance Phonetics and Phonology," Oxford Scholarship Online, (2019).
- [7] G. L. Salomão, "Relationship between perceived vocal registers and glottal flow parameters: preliminary results," *Speech Prosody 2008*, Campinas, Brazil, (2008).
- [8] K. Neumann, P. Schunda, S. Hoth, and H. A. Euler, "The interplay between glottis and vocal tract during the male passaggio," Folia Phoniatr Logop, 57, pp. 308–327, (2005).

- [9] N. Henrich, J. Smith, and J. Wolfe, "Vocal tract resonances in singing: Strategies used by sopranos, altos, tenors, and baritones", Journal of the Acoustical Society of America, **129**: pp. 1024-1035, (2011).
- [10] N. Henrich, B. Roubeau, M. Castellengo, "On the use of electroglottography for characterisation of the laryngeal mechanisms," In proc. Stockholm Music Acoustics Conference, Stockholm, Sweden, Aug. (2003).
- [11] L. Thurman, G. Welch, A. Theimer, C. Klitzke, "Addresing vocal register discrepancies: an alernative, science-based theory of register phenomena," Second International Conference The Physiology and Acoustics of Singing National Center for Voice and Speech 6–9 October 2004 Denver, Colorado, USA, (2004).
- [12] H. Matthew, and L. Linda, Voice Secrets: 100 Performance Strategies for the Advanced Singer, *Rowman & Littlefield*, (2016).
- [13] F. W. Graham, M. H. David, and N. John, The Oxford Handbook of Singing, Oxford University Press, (2019).
- [14] G. Fant, "Speech Sounds and Features," *MIT, Cambridge, Mass*, (1973).
- [15] G. Fant, "Acoustic Theory of Speech Production," The Hague, Mouton, (1960).
- [16] E. C. Rubyetta, "A study of vowel-sound intensities in singing," thesis in music education," August, (1974).
- [17] K. Sjolander, "Snack sound toolkit," KTH Stockholm, Sweden, http://www.speech.kth.se/snack, (2004).
- [18] H. Kawahara, A. D. Cheveign, and R. D. Patterson "An instantaneous-frequency-based pitch extraction method for high quality speech transformation: revised TEMPO in the STRAIGHT-suite," Proc. ICSLP'98, Sydney, Australia, December, (1998).
- [19] J. Sundberg, G.L. Salomão, and K. R. Scherer, "What does LTAS tell about the voice source?" The Voice Foundation Annual Symposium: Care of the Professional Voice, Philadelphia, USA, (2018).
- [20] Fant, G. "The voice source in connected speech," Speech Commun. 22, 125-139, (1997).
- [21] D. G. Miller., "Registers in Singing: Empirical and Systematic Studies in the Theory of the Singing Voice," University of Groningen, (2000).
- [22] A. Nicole, "Teaching Registration in the Mixed Choral Rehearsal:Physiological and Acoustical Considerations", (2011).