# Voice Similarities between Sisters*

**Do-Heung Ko****

## ABSTRACT

This paper deals with voice similarities between sisters who are supposed to have common physiological characteristics from a single biological mother. Nine pairs of sisters who are believed to have similar voices participated in this experiment. The speech samples obtained from one pair of sisters were eliminated in the analysis because their perceptual score was relatively low. The words were measured in both isolation and context, and the subjects were asked to read the text five times with about three seconds of interval between readings. Recordings were made at natural speed in a quiet room. The data were analyzed in pitch and formant frequencies using CSL (Computerized Speech Lab) and PCQuirer.

It was found that data of the initial vowels are much more similar and homogeneous than those of vowels in other positions. The acoustic data showed that voice similarities are strikingly high in both pitch and formant frequencies. It is assumed that statistical data obtained from this experiment can be used as a guideline for modelling speaker identification and speaker verification.

**Keywords: Voice Similarities, Vocal Tract, Speaker Verification**

## 1. Introduction

Virtually, no previous studies have been made on voice similarities in the Korean language. It is assumed that the similarities of voice among family members are high when measured both perceptually and acoustically. If two brothers' or sisters' voices are very similar acoustically, that is because they share common physiological characteristics. The voices, therefore, are sometimes strikingly similar among brothers or sisters. It is presumed that these common physiological features come from the same biological mother. It is not difficult assume that the length of the vocal tract, the shape of the articulatory organs and the sizes of cavities are affected genetically.

---

* This was supported by a grant(R01-1999-00229) from Korea Science & Engineering Foundation.

** Division of Language and Literature, Hallym University

It is well-known that the major factors affecting voice qualities are (1) the length of the vocal tract, (2) the thickness and elasticity of the vocal folds, (3) the shape and function of articulators, and (4) the sizes of cavities. It is said that fundamental frequency (Fo) and possibly amplitude are usually affected by (1) and (2) while Formant frequencies are primarily affected by (3) and (4).

This paper deals with voice similarities between sisters who are supposed to have common physiological characteristics from a biological mother. Statistical data obtained from this experiment can be used as a guideline for modelling speaker identification and speaker verification.

## 2. Experimental Procedures

### 2.1 Subjects and Speech materials

Nine pairs of sisters who are believed to have similar voices participated in this study. The subjects, in their early 20's, were raised and educated in either Seoul or Chunchon. Unfamiliar listeners were instructed to listen to the recordings and rate their similarity on a numerical scale (1-10). Results showed that most of them indicated very high perceptual scores (8 out of 10), except one pair of sisters. The speech samples obtained from this pair of sisters were eliminated in the analysis because their perceptual score was relatively low (4 out of 10). In this experiment, the initial vowels were analyzed because the data of the initial vowels were much more similar and homogeneous than those of vowels in other positions. The words were measured in both isolation and context, and the subjects were asked to read the text five times with about three seconds of interval. The test word list is as below:

/a/ in kaji 'eggplant' /u/ in munje 'problem'
/i/ in iyagi 'story' /o/ in moksori 'voice'

In context, the carrier sentence is : 'This is called _____.'

In analyzing pitch patterns, different lengths of sentences were compared as below:

(1) Youngho-ga maeu aphayo.
'Youngho is very sick.'
(2) Oje hakkyo-e kassossoyo.
'Yesterday, I went to school.'
(3) sesang-un cham arumdapkunyo.
'The world is so beautiful.'

(4) nori kongwon-enun saramdul-i mani moimnida.
'Children's parks are always crowded.'
(5) onul taehangno-eso yeoguk-ul poassoyo.
'We went to see a drama at Taehangno.'
(6) Chunchon-e yeorum-un maeu teowoyo. '
'It's very hot in summer in Chunchon.'

### 2.2 Acoustic Analysis

Recordings were made at natural speed in a quiet room. The data were analyzed in pitch and formant frequencies (F1, F2, F3) using CSL (Computerized Speech Lab) and PCQuirer. To control the overall duration for the test words or sentences, a pair of sisters was asked to repeat them again if needed.

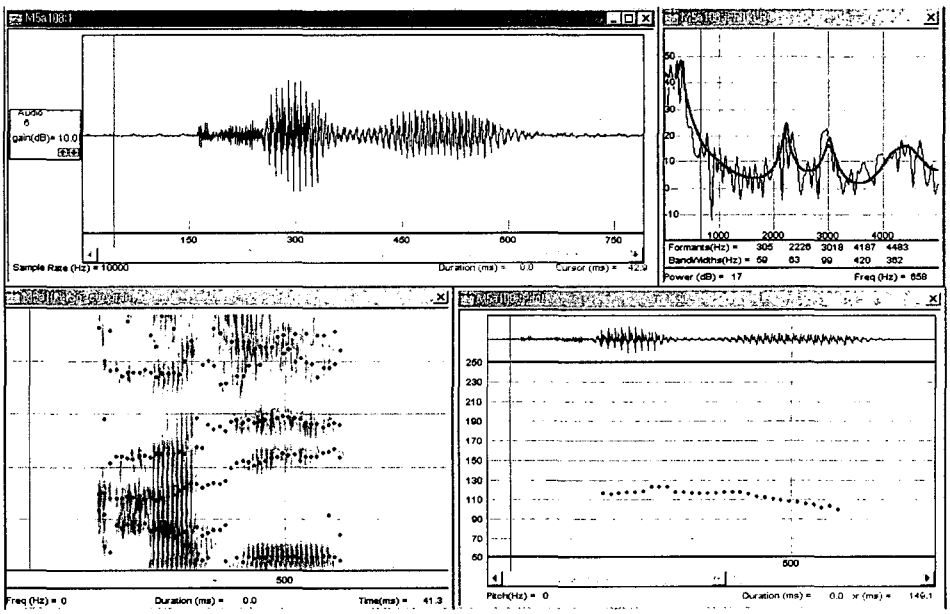

Figure 1. Samples of Acoustic Measurements

## 3. Results and Discussion

### 3.1 Duration

Even though recordings were made at natural speed, their speaking rates were different for some pairs of sisters as we can see in Figure 2. In case of pair 1, the difference of utterance durations was 317 msec which is considerably long. However, their pitch patterns were surprisingly similar.

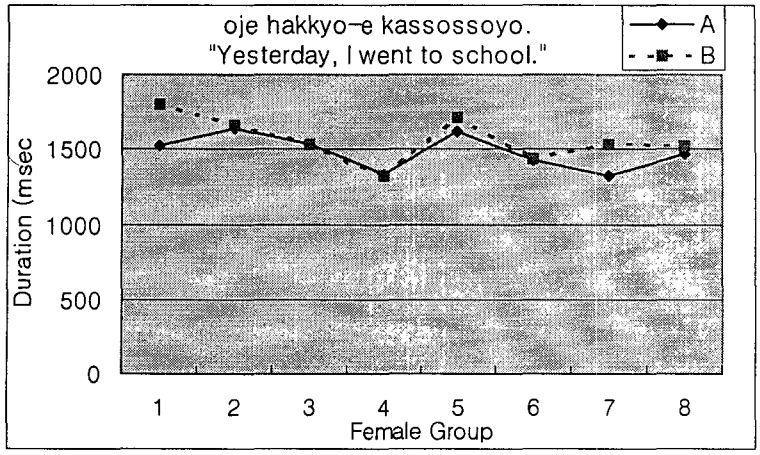


Figure 2. A Sample of Utterance Duration

3.2 Pitch

As we can see in Figure 3 and 4, pitch patterns were very similar. In a sentence which has four peaks, the mean peak values, and the rate of declination were strikingly similar in all pairs of sisters. It is assumed that pitch is the one of the most important cues for voice similarities because pitch is the perceptual correlate of fundamental frequency (Fo).

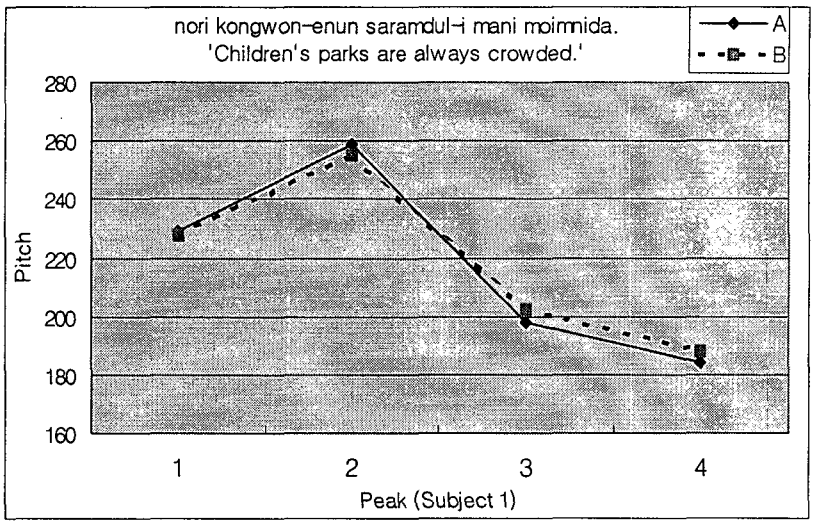


Figure 3. Comparison of Pitch Variation in a Pair of Sisters

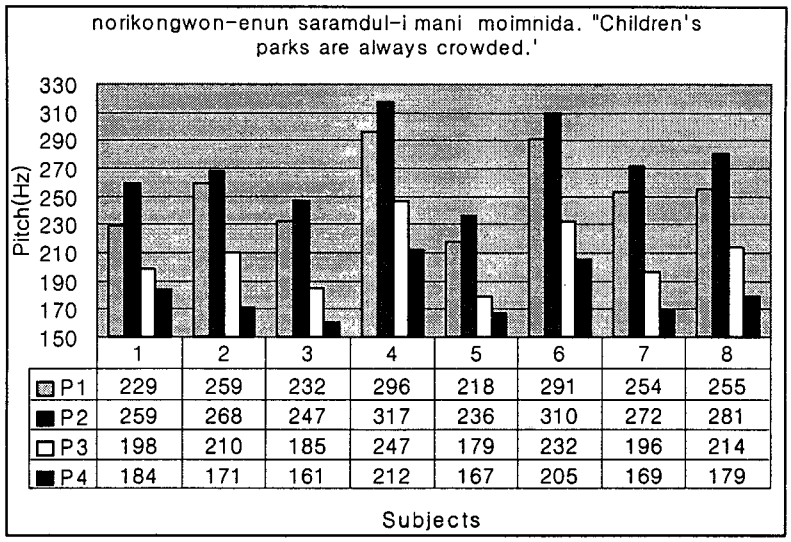


Figure 4. Comparison of pitch variation in eight pairs of sisters

3.3 Formant Frequencies

Figure 5 shows that the mean value of formant frequency in initial position was fairly similar in pairs of sisters. In addition, as can be seen in Table 1, the mean value of pitch in isolated word was also similar in most cases. The mean values of formant frequencies in the same pair were very consistent. It should be noted, however, that the mean values of F1 were more consistent than those of F2 and F3 throughout the data. This means that F2 and F3 could be more reliable for speaker verification.

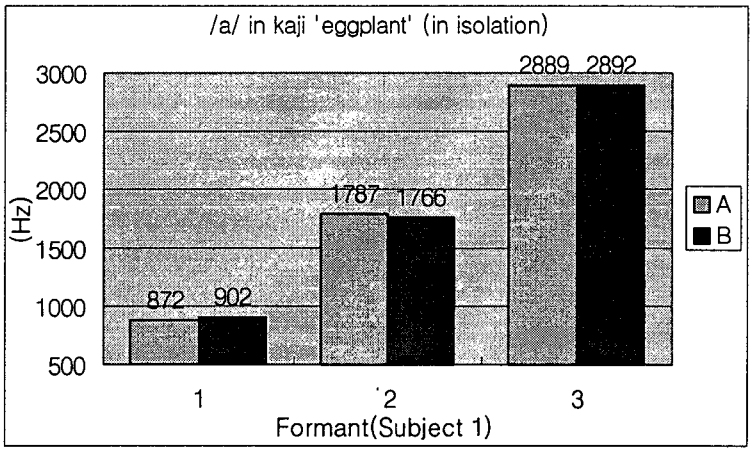


Figure 5. /a/ in kaji 'eggplant' in isolation for a pair of sisters

Table 1. Pitch and formant frequencies in eight pairs of sisters

| | Pitch Mean | SD | F1 Mean | SD | F2 Mean | SD | F3 Mean | SD |
|---|---|---|---|---|---|---|---|---|
| Subject 1A | 175 | 2 | 872 | 23 | 1787 | 37.5 | 2889 | 15.2 |
| 1B | 191 | 7.8 | 902 | 26.7 | 1766 | 14.1 | 2892 | 23.4 |
| 2A | 199 | 5 | 893 | 28 | 1601 | 53 | 2727 | 12.6 |
| 2B | 195 | 5.3 | 831 | 49.9 | 1700 | 104.7 | 2700 | 121 |
| 3A | 190 | 5.3 | 839 | 23.7 | 1899 | 38.5 | 2561 | 45.2 |
| 3B | 180 | 7.2 | 799 | 26.7 | 1858 | 3.8 | 2501 | 41.4 |
| 4A | 212 | 2.1 | 852 | 23 | 1613 | 30.6 | 2648 | 44.5 |
| 4B | 218 | 1.4 | 893 | 12 | 1669 | 56.7 | 2606 | 36.7 |
| 5A | 172 | 0.5 | 907 | 28.5 | 1739 | 42.1 | 2688 | 19.8 |
| 5B | 171 | 1 | 923 | 10.9 | 1777 | 10.9 | 2711 | 41.9 |
| 6A | 214 | 2.4 | 868 | 18.1 | 1932 | 36 | 2898 | 45.1 |
| 6B | 208 | 2.4 | 830 | 12.1 | 1945 | 26.2 | 2867 | 63.5 |
| 7A | 193 | 2.1 | 909 | 23.4 | 1793 | 54.3 | 2941 | 30.5 |
| 7B | 197 | 4.8 | 931 | 14.6 | 1766 | 41.6 | 2903 | 44.1 |
| 8A | 184 | 3.2 | 786 | 29.1 | 1926 | 20.8 | 2758 | 45.3 |
| 8B | 192 | 4.5 | 797 | 17 | 1915 | 17.6 | 2770 | 26.8 |

As in Figure 6 and 7, the Average values of F1 and F2 in isolation and context were similar in each pair of sisters except subject 6. In examining formant frequencies, the initial vowels were analyzed because the data of the initial vowels was much more similar and homogeneous than those of vowels in other positions. In many cases, vowels in other than initial position were not acoustically similar.

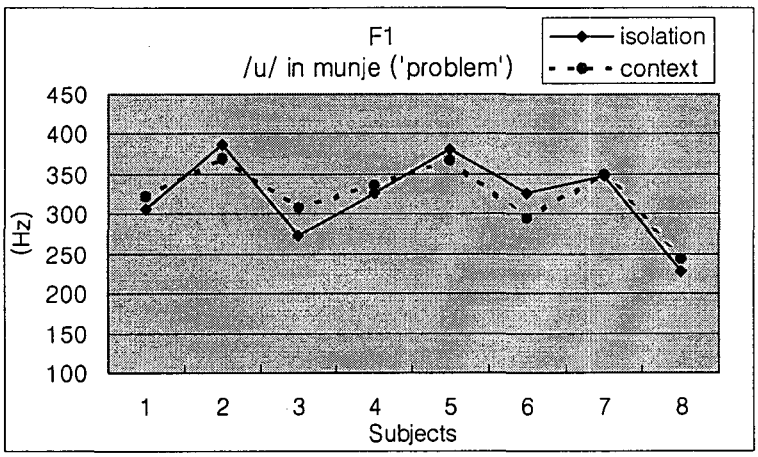


Figure 6. Comparison of F1 Average in Isolation vs. Context

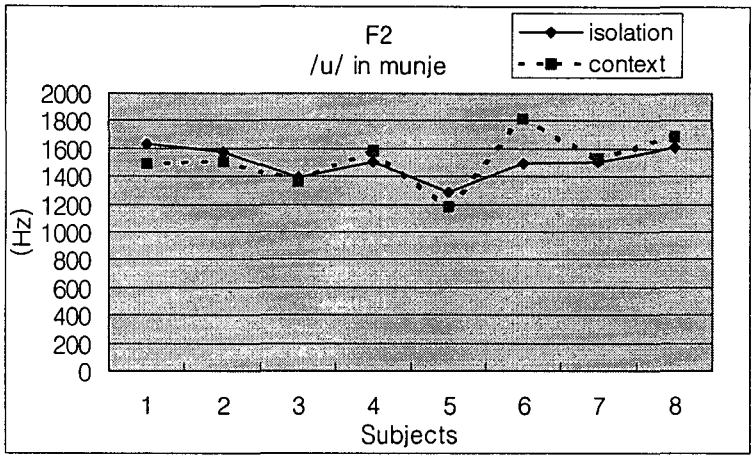


Figure 7. Comparison of F2 average in isolation vs. context :

## 4. Concluding Remarks

Voices are as distinctive as our faces. Indeed, no two voices are exactly alike. It should be well known that the shape of an individual's vocal tract is partly genetic, partly learned. In other words, uniqueness or individuality in voice is a product of both physiology and learning. The learned component of the equation could be called vocal habits (Kushner & Bickley 1995). These would include items such as rhythm and rate of speech and vowel pronunciation. Thus, it is not a surprising fact that family members sound alike because they share genes and environments.

In this experiment, the author focused on how pitch and formant frequencies play a role in distinguishing similar voices of eight pairs of sisters because the two acoustic parameters are believed to be the most important in the fields of speaker identification and speaker verification. According to Kuwabara (2001), voice individuality has been found less sensitive to pitch frequency than for formant manipulation. Although this study was bound to comparisons of the two parameters, intonation patterns as a learned component could be other important cues for voice similarities.

## Acknowledgement

This work was partially supported by a grant from Hallym University in 1999.

## References

Ko, Do-Heung. 2000. "An analysis of phonetic parameters for individual speakers." *Speech Sciences* (Korean Association of Speech Sciences), 7(2), 177-189.

Kushner, Rochel E. & Corne A. Bickley. 1995. "Analysis and perception of voice similarities among family members." ASA 130th Meeting, St. Louis, MO.

Kuwabara, Hisao. 2001. "Voice Individuality and related acoustic Parameters." *Proceedings of ICSP 2001* (International Conference on Speech Processing), Vol. 2, 829-832.

Yang, Byunggon. 2001. "Perceptual Experiment on Number Production for Speaker Identification." *Speech Sciences* (Korean Association of Speech Sciences), 8(1), 7-19.

Received: July 27, 2001.
Accepted: August 19, 2001.

▲ Do-Heung Ko
Division of Language and Literature, also
Division of Speech Pathology and Audiology
Hallym University
1 Okchon-dong, Chunchon
Kangwon-do 200-702, KOREA
Tel: +82-33-240-1205
E-mail: dhko@hallym.ac.kr