TUTORIAL

From articulatory phonetics to the physics of speech: Contribution of Chiba and Kajiyama

Kikuo Maekawa*

Department of Language Research, National Institute for Japanese Language, 3–9–14, Nishigaoka, Kita-ku, Tokyo, 115–8620 Japan

Keywords: The Vowel: Its Nature and Structure, Tsutomu Chiba, Masato Kajiyama

PACS number: 43.70.Aj, 43.70.Bk

1. INTRODUCTION

Sixty years ago, shortly after the beginning of the Pacific War, a classical work of phonetics was published in Tokyo, *The Vowel: Its Nature and Structure* by Tsutomu Chiba and Masato Kajiyama [1].¹

This monograph can be regarded as a classical work for two main reasons. For one, it opened the way to calculate a vowel spectrum from the data of the three-dimensional vocal tract shape for the first time in the history of phonetics. For another, it gave strong impetus to the foundation of what we today call the acoustic theory of speech production [2,3].

2. CONTENTS OF THE BOOK

The Vowel consists of four parts and thirteen chapters. Part one, entitled "The Action of the Larynx," is devoted to analysis of the voice source. The main apparatuses used for the study were the stroboscopic motion picture of the larynx and the Fourier analysis of vowel signals. The dynamic aspects of the vibration of the vocal folds observed under four voice registers ('soft', 'sharp', 'ordinary', and 'falsetto') were compared to the corresponding vowel spectra.

Part two, "The Mechanism of Vowel Production" includes an historical survey of the theories of vowel production (chapter 5). This is interesting and important reading material for today's researchers because the historical backgrounds stated here are now almost forgotten. The theories surveyed involve those of Willis, Wheatstone, Helmholtz, Herman, and Rayleigh-Trendelenburg.

This survey is followed by the authors' investigation of the action of a resonator (chapter 6) and the behavior of the vocal tract as a resonator (chapter 7). A remarkable characteristic of these chapters is the introduction of electronic circuit theory into the investigation of vocal tract resonance. This is clearly a precursor to the post-war work of Fant, Stevens, and others [2,3,7,8].

Part three is entitled "The Measurement of the Vocal Cavity and the Calculation of Natural Frequencies." Vocal tract shape was measured by a combination of X-ray photography, palatography, and laryngoscopic observation of the pharynx (chapter 9). This was the most up-to-date technique at the time of publication, and it remained so until MRI-based measurement became available.

In chapter 10, formant frequencies were calculated on the assumption that the measured vocal tract configuration of a given vowel could be classified into one of ten ideal resonators, viz., A) uniform tube closed at one end, B) pipe open at both ends, C) non-uniform tube open at one end (with wide pharynx), D) non-uniform tube open at one end (with narrow pharynx), E) slightly tapered pipe open at both ends, F) closed pipe with a small open end, G, H, I) resonators with a large enclosure and one or two small openings (i.e., various types of Helmholtz resonators), and J) double resonator. (See Fig. 1(a) in Motoki's paper in this issue (p. 208)).

The vocal tract shape of [i] and [a], for example, were approximated to types H) and J), respectively; the natural frequencies calculated from the approximated vocal tract coincided fairly accurately with the values obtained by Fourier analysis (290 and 2,700 Hz from the vocal tract data and 280 and 2,800 Hz from Fourier analysis, in the case of [i], for example). Moreover, the authors synthesized vowel sounds using replicas of the vocal tract and a larynxtone emitter, and compared the spectra of synthetic and natural vowels.²

The success of the calculation of formant frequencies from vocal tract data is the highlight of the whole study and without doubt the main source of the long-standing scientific value of the book.

^{*}e-mail: kikuo@kokken.go.jp

Although the calculation done in chapter 10 was limited to the first two formants, the authors tried to extend their calculation in the following chapter to the higher formants and to the prediction of formant frequency shift due to the location of the vocal tract constriction. It is in chapter 11 that we encounter the famous figure showing the distribution of volume velocity in the uniform vocal tract (Figure 93 on page 147); this figure was cited by Fant on page 85 of [3] and adopted by many subsequent monographs of acoustic phonetics, including very recent ones.

Part four is entitled "A Subjective Study of the Nature of a Vowel" and is devoted to the issues of vowel perception. After an introduction to the physiology and psychology of hearing (chapter 12), the authors' attention centered on two issues: the reconciliation of so-called "transient" and "steady state" theories of the vowel on the one hand, and the relation between changing pitch and formant structure on the other.

Based upon the experimental results obtained by frequency filtering and velocity-altered playback of a gramophone disc, the authors advocated their "space-pattern theory" of vowel perception. They wrote, "A vowel is characterized by its relative formants, provided the centers of the formants are situated within certain frequency regions fixed for a given vowel." (p. 193–4. Italics theirs). This frequency region was named "characteristic frequency region" and was supposed to be fairly wide.

Part four is fundamentally different from the rest of the book in that the problems dealt with in chapter 13 remain largely unsolved even by today's researchers.

3. THE AUTHORS AND THE LABORATORY

Tsutomu Chiba (1883–1959) studied English literature at Tokyo Imperial University. From 1913 to 1916, the Japanese government sent him to England to study English philology. During this period, Chiba also learned phonetics probably by attending the lectures of professor Daniel Jones at the University of London. After his return, Chiba taught phonetics at the Tokyo School of Foreign Language (TSFL)³ where he became a professor in 1919.

In 1928, TSFL received governmental funding for the improvement of its activities and decided to invest it in establishing Japan's first phonetics laboratory. The laboratory was established in 1929, and Chiba was nominated as its director. The laboratory was equipped with an anechoic chamber, X-ray apparatus, condenser microphones, and an electromagnetic oscillograph in addition to traditional devices like a kymograph and Helmholtz resonators. The laboratory was one of the most well equipped phonetics laboratories in the world.⁴

Chiba's contribution to the laboratory was twofold: for

one, it was his steady policy to introduce natural science, namely physics, into the study of phonetics, and for this purpose he continuously employed young physicists to help him. The interdisciplinary nature of the TSFL laboratory was in advance of the standard practice in the world at that time. For another, Chiba concentrated the laboratory's activity towards the study of vowels. Judging from his writings other than *The Vowel*, Chiba's main research interest was in the cross-language comparison of vowels, especially between Japanese and English. His experience of teaching English phonetics, however, led him to cast doubt on the scientific validity of the vowel diagram and the notion of the highest point of the tongue being the classificatory criterion of vowels.

From a retrospective point of view, concentration upon vowels rather than consonants was a very smart decision, because it was only to vowels that manual Fourier analysis (corresponding to today's DFT analysis) could be applied in a satisfactory manner, not to mention the dynamic nature of consonants. To sum up, Chiba was an innovative research manager with indomitable will.

Masato Kajiyama (1909–1995) was a physicist educated at Tokyo Imperial University. After spending one year as an assistant in the physics department where he was educated, he joined the phonetics laboratory of TSFL in 1933. His first work in the laboratory was the comparative analysis of accent in various languages, which was published as Chiba's monograph [4].⁵

Kajiyama then started his endeavors to reach a full understanding of the nature of vowels and continued in this direction until the publication of *The Vowel*. His contribution to *The Vowel* was both clear and enormous. He was the only physicist in the laboratory after 1933. Moreover, we have direct evidence of his contribution: the five volumes of laboratory notebooks that Kajiyama wrote during the years 1933–1940, which were preserved by Koh-ichi Kajiyama, Masato's son, and found in 2001 due to the present author's research. Figure 1 is a picture of a page from one notebook in which Kajiyama examined the distribution of particle velocity and sound pressure in a uniform pipe closed at one end.⁶

A good way to understand Kajiyama's contribution is the comparison of the laboratory's first monograph [5] and *The Vowel*. Both of them analyzed vowels using nearly the same experimental facilities, i.e., X-ray photographs and oscillograms. The fundamental difference between the two publications lies in the treatment of vocal tract shape. In [5], the vocal tract was analyzed as a two-dimensional shape (the mid-sagittal area of a vocal tract was measured from X-ray photographs), while in *The Vowel*, the vocal tract was treated as a three-dimensional tube. As a matter of course, three-dimensional treatment is the theoretical prerequisite for the physical treatment of the vocal tract as

K. MAEKAWA: CONTRIBUTION OF CHIBA AND KAJIYAMA



Fig. 1 Photograph of Kajiyama's laboratory note. Courtesy of Koh-ichi Kajiyama.

a resonator. It is not an exaggeration to say that the research policy of the first monograph was articulatory phonetics, but that the study of *The Vowel* was guided by physical principles. It was this new research policy that required Kajiyama's painstaking efforts.

Despite his great contribution, Kajiyama's position in the laboratory was not a favorable one. Throughout the years in TSFL he was employed as a part-time worker with poor payment. This is partly because in pre-war Japan there were no teaching positions for physicists in a college like TSFL whose primary mission was to train specialists in foreign languages. In addition to this, Japan in the 1930's was in a deep recession that ultimately led the country to war in China and the Pacific area.

In 1944, Kajiyama left TSFL to teach physics at the Military Academy of the Imperial Army, but this position was terminated naturally by the dissolution of the Imperial Army the following year. It was only in 1946 that Kajiyama found a stable position in Musashi Institute of Technology, where he worked as a professor of physics until his retirement in 1980. In this position, Kajiyama continued research on acoustics but never returned to the field of speech study. In 1968, when the International Congress of Acoustics was held in Kyoto, the organizing committee invited Kajiyama to join the conference, but in vain.⁷



Fig. 2 Chiba (right) and Kajiyama (left) operating the stroboscope in the phonetics laboratory of the Tokyo School of Foreign Languages in the summer of 1933. Courtesy of Koh-ichi Kajiyama.

Chiba, on the other hand, retired from TSFL in 1945 and became a professor of English at Sophia University in 1950. He then showed a keen interest in aphasiac speech. His last research article was entitled, "Phonetic study of speech centers" [6]. He became Japan's first neurolinguist, so to speak.

The remaining history of the laboratory needs only a short mention. In 1944, TSFL moved to a locality in northern Tokyo to evacuate from U.S. strategic bombing. But the new locality was not safe at all. Due to the incendiary bombing in the spring of 1945, the phonetics laboratory together with all experimental materials collected since 1929 was destroyed and lost forever. As far as I know, the only remaining material from the laboratory is a copy of an X-ray photograph and some photographs belonging to Kajiyama, which were reserved with the laboratory notebooks mentioned above. Figure 2 is a snapshot of Kajiyama (left) and Chiba (right) operating their stroboscope in the summer of 1933.

4. CONCLUDING REMARKS

The Vowel: Its Nature and Structure is a landmark of speech research. Before the publication of *The Vowel*, there were three trends in the study of speech sounds: articulatory phonetics, anatomical as well as medical studies of human vocal organs, and physical analysis of musical sounds as represented by Helmholtz's work. Studies in each of these trends were carried out more or less independently until the 1930's. The importance of *The Vowel* lies in the fact that all these trends were merged into a single science. Moreover, *The Vowel* was a precursor to the electronic treatment of speech sound that has played a leading role in speech science during the last half of the 20th century.

Postscript: See [7] and [8] for reminiscences on The

Vowel by the constructors of the acoustic theory of speech production. See [9] for the reception of *The Vowel* in England. The current paper is largely dependent upon the bibliographical and field research done by the present author and Kiyoshi Honda upon the background of *The Vowel* [10]. Lastly, [12] will help readers to understand the issues discussed in *The Vowel*.

ACKNOWLEDGMENT

I thank Koh-ichi Kajiyama who kindly gave permission to reproduce his father's photograph and notebooks. I also thank Dr. Kiyoshi Honda of ATR Human Information Science Laboratories through whom my understanding of the physical and anatomical issues discussed in *The Vowel* was deepened by a considerable amount.

REFERENCES

- [1] T. Chiba and M. Kajiyama, *The Vowel: Its Nature and Structure* (Tokyo-Kaiseikan, Tokyo, 1942).
- [2] K. N. Stevens and A. S. House, "Development of a quantitative description of vowel articulation," J. Acoust. Soc. Am., 27, 484–493 (1955).
- [3] G. Fant, Acoustic Theory of Speech Production: With Calculation Based on X-Ray Studies of Russian Articulation (Mouton, The Hague, 1960).
- [4] T. Chiba, Research into the Nature & Scope of Accent in the Light of Experimental Phonetics (Fuzanbo Publishing Company, Tokyo, 1935).
- [5] T. Chiba, Research into the Characteristics of the Five Japanese Vowels Compared Analytically with Those of the Eight Cardinal Vowels (Nichibei-Press, Tokyo, 1931).
- [6] T. Chiba, "Phonetic study and speech centres," Bull. Phon. Soc. Jpn., 97, 1–7 (1958).
- [7] G. Fant, "T. Chiba and M. Kajiyama, pioneers in speech acoustics," J. Phon. Soc. Jpn., 5(2), 4–5 (2001).
- [8] K. N. Stevens, "The Chiba and Kajiyama book as a precursor to the acoustic theory of speech production," J. Phon. Soc. Jpn., 5(2), 6–7 (2001).
- [9] J. Laver, "Some future directions of phonetic research," J. Phon. Soc. Jpn., 5(1), 46–48 (2001).

- [10] K. Maekawa and K. Honda, "On *The Vowel: Its Nature and Structure* and related works by Chiba and Kajiyama," *J. Phon. Soc. Jpn.*, 5(2), 15–30 (2001).
- [11] T. Arai, "The replication of Chiba and Kajiyama's mechanical models of the human vocal cavity," J. Phon. Soc. Jpn., 5(2), 31–38 (2001).
- [12] K. Honda, "Evolution of vowel production studies and observation techniques," *Acoust. Sci. & Tech.*, 23, 189–194 (2002).

Notes

- ¹ There is confusion about the publication year of *The Vowel*. The year printed on the back of the title page was 1941. On the back cover of the book, however, it was printed in Japanese era *Showa 17-nen*, *1-gatsu*, which corresponds to January 1942. See [10] for more details on the publication year of the book.
- ² Arai reports the frequency characteristics of the vowels that were synthesized based upon the vocal tract data in *The Vowel* [11].
- ³ Currently, TSFL is known as Tokyo University of Foreign Studies.
- ⁴ See [10] for the photograph image of these research facilities.
- ⁵ Though this work was published as Chiba's monograph, Kajiyama's name was mentioned in the preface.
- ⁶ A picture of another page concerning the theory of a double-resonator can be seen in [10].
- ⁷ Personal communication with Osamu Fujimura.



Kikuo Maekawa was born in Kyoto in 1956. He received his M.A. in linguistics from Sophia University in 1982. After teaching Japanese linguistics at Tottori University for five years, he moved to the phonetics laboratory of the National Institute for Japanese Language (formerly known as the National Language Research Institute until 2000) in 1989, and remained there to this day. From 1993 to 1994,

he visited the Ohio State University as a visiting researcher. His research activities include experimental phonetics, sociolinguistics, and corpus linguistics. Currently he is the head of the Second Division of the Department of Language Research.