

Research on Japanese Language Teaching Reform Based on Speech Visualization Technology in the Context of Artificial Intelligence

Qi Fang

Zhengzhou University of Science and Technology, 450064

ABSTRACT

With the rapid development and application of artificial intelligence technology, Japanese language education has ushered in new opportunities and challenges, Japanese language education responds to the needs of the times, and the deep integration with computer technology and artificial intelligence technology to realize the effective teaching of Japanese language education is imperative. Speech visualization is based on the theory of multimodal discourse, integrating visualization technology into speech teaching, effectively integrating visual and auditory modal information, and realizing the effective teaching of Japanese speech with sound visualization. In this paper, through combing the development history of speech visualization research, we discuss how to apply speech visualization technology to realize the reform of Japanese speech teaching, try to use human-computer interaction and human-computer synergy to improve the efficiency of Japanese speech learning, comprehensively improve the appearance of speech, and realize the self-evaluation and independent learning of Japanese speech.

KEYWORDS

Japanese phonetics; Phonological visualization; Effective instruction

DOI: 10.47297/taposatWSP2633-456923.20230402

1 Introduction

The Education Informatization 2.0 Action Plan calls for the active development of innovative research and demonstration of intelligent education based on emerging technologies such as artificial intelligence, big data and the Internet of Things, and relying on various types of intelligent devices and networks, so as to promote changes in the mode of education and ecological restructuring under the auspices of new technologies. This points out the development direction for the integration of information technology and education. With the application of artificial intelligence technology, big data, and Internet of Things technology in the cultivation of listening, speaking, reading, writing, and translating skills in Japanese language courses, Japanese language speech teaching has the technical conditions to realize intelligent teaching, but the teaching mode has not been transformed in line with the development of technology.

At present, China is in the critical period of comprehensive modernization of education, which puts forward higher requirements for the high-quality cultivation of Japanese language talents, and the breakthrough lies in comprehensively improving students' effective communication ability and independent learning ability. For this reason, it is the trend to integrate visualization technology into Japanese speech teaching, reform the traditional Japanese speech teaching mode, use human-

computer interaction, human-computer collaboration and other technologies to realize the differentiated and effective teaching of Japanese speech, and comprehensively improve the quality of Japanese language talents.

2 Pilot Study

In the keyword search of "speech visualization" in ZhiNET, excluding those with low relevance, a total of 38 theses were retrieved, including 6 master's theses and 5 core theses, the earliest of which appeared in the Research on Visualized Speech Information Retrieval Model published by Zhou Ning (2006), which proposes the visualized speech information retrieval model based on the automatic speech recognition technology, mainly used for speech program users to facilitate the extraction of information. information retrieval model based on automatic speech recognition technology, which is mainly used for voice program users to facilitate the extraction of information, and does not involve the field of education. Subsequently, over the past decade, the number of papers published has increased year by year and reached an annual publication volume of 6 in 2013, with a drop in 2014-2016, followed by a slow incremental trend.

As far as research tools are concerned, empirical research based on visualization tools occupies the majority, and Praat, which is the most used visualization tool, has the largest number of related literature, reaching 32 articles, followed by Cool Edit Pro, EMA, BetterAccent Tutor and Reading Assistant, which each involves one article of literature; non-empirical research is mainly focused on the effectiveness research of visualization assisted speech teaching and practical research on specific application of teaching. The non-empirical research mainly focuses on the effectiveness of visualization in assisting speech teaching and the practical research on specific applications in teaching, which provides a strong argument for the effectiveness and positive effects of visualization software applied to the teaching of speech.

From the viewpoint of research content, in recent years, the research on the visualization of phonological teaching aspects of the research content has gradually deepened and developed. Zhuang Muqi and Bu Youhong (2011) conducted a study on the visual teaching of suprasegmental phonology using BetterAccent Tutor software, and found that the visual phonological learning system can arouse the learners' awareness of stress and non-stress, display the visual interface of intonation, tone intensity, and interpretation, which can help them practice repeatedly against each other until the pitch, tone intensity, and duration are close to those of native speakers. Li Ying (2017) proposed the characterization of pronunciation errors at the level of suprasegmental segments and analyzed the causes from the point of view of experimental phonetics in order to address the problems such as the neglect of suprasegmental features formed in the speech stream. Dan Liu (2019) conducted an experiment on teaching English intonation using Praat speech analysis software, and found that students' intonation problems, such as narrow register and overuse of rising and falling tones, were corrected through the study.

In recent years, the field of Japanese language phonetics has been expanding its ideas and actively exploring effective teaching and learning methods. As far as research is concerned, the traditional method of imitating pronunciation, in which the teacher provides model pronunciation and the learner imitates it, is no longer effective in improving the learner's pronunciation (Toshiyuki Kono and Kan Matsuzaki 1998). After that, scholars developed rhythmic graphs (Toshiyuki Kono et al. 2004) and pitch arcs (Chieko Nakagawa et al. 2009) based on the concept of visualization, in addition to the development and use of online Japanese tone dictionaries, the pitch curve graphs, and the tone

rhythmic graphs to a certain extent help the students to be able to practise pronunciation in a more intuitive and more convenient way (Takako Toda 2017), and Jiaqi Liu (2018) used the experimental phonological method, by summarizing the characteristics of Japanese learners' acquisition of Japanese phonological segments and combining the theory of mediated language to study the phonological problems in Japanese language teaching. Although these phonological pedagogies possess different forms of expression, the basic idea is to try to guide learners to integrate visual and auditory modal information through phonological visualization, so as to improve the efficiency of phonological acquisition and teaching.

To sum up, research on speech teaching has made great development in recent years, the number of theoretical studies has been reduced, and more attention has been paid to the empirical research on speech, researchers use computer tools and visualization analysis software to obtain the real materials of speech intonation and analyze them, and obtain more scientific and accurate data, and the research on speech has shown the tendency of being deeper and more refined. From the research direction, the research on phonological teaching is no longer limited to the description and difference comparison of single tones, but pays more attention to suprasegmental phonemes, such as stress, intonation and rhythm, etc. However, from the research object, it is basically focused on the description and comparison of single tones. However, from the perspective of the research object, basically English majors are the main students, and less for other languages, among which there are involved in the phonological teaching of Japanese majors, the number of only 4 articles, for the Japanese phonological visualization teaching research is still in the preliminary stage, the relevant research is far from being enough, and the scope of the research object needs to be further expanded. From the application of the research results, the use of voice visualization teaching software for teaching the positive effect has been widely confirmed, a lot of literature on the results of the teaching concept can be drawn on, the specific teaching practice guidance is not strong, the teacher's practical application is not extensive, especially for the Japanese language speech teaching, the current application of visualization technology and the specific methods of Japanese language speech teaching is still not clear, the present The combination of research results and speech teaching practice and the realization of the path of research is still insufficient, and needs to be further explored and researched.

3 Definition of Speech Visualization

The concept of "Speech Visualization" was firstly proposed in the 1940's. In 1945, W. Koenig and his team firstly proposed the concept of "Visible Speech" by using classical acoustic spectrometer to analyze speech data. In 1945, W. Koenig and his team first proposed the concept of "Visible Speech" by analyzing speech data with classical spectrography, i.e., the process of displaying the frequency and time distribution spectrum of speech energy after converting complex speech data with the use of an acoustic spectrometer. Speech visualization was initially used in the fields of acoustics and medicine, and then extended to the field of education and teaching, resulting in visual speech teaching. Visual speech teaching is based on multimodal technology, through the computer and speech synthesis software, the sound information will be electronically and graphically, so that the auditory information is transformed into visual symbols, so that the voice becomes tangible and visible, so that the voice characteristics become intuitive and tangible, and through the analysis of the comparison, to obtain more perception, which is more conducive to the students to adjust the corrections of their own pronunciation.

Since the 1970s, many research institutions and scholars at home and abroad have developed a series of speech visualization software, such as Praat, Speech Analyzer, Better Accent Tutor, Wavesurfer and so on. This kind of speech software can record, save, recognize, label, analyze and edit speech, and provide visual interfaces of waveform, fundamental frequency, energy, etc. It is easy to download and highly operable.

4 Current Status of Japanese Language Phonetics

The syllabus of the basic stage of teaching Japanese language in colleges and universities has made a clear stipulation on the teaching of phonetics: the ultimate goal of Japanese phonetics is to require students to correctly master the pronunciation of the Japanese language, the tone of voice, the stress and intonation of sentences, and to be able to communicate correctly and fluently and to express their thoughts and feelings appropriately in the language practice. However, the reality is that many Japanese learners have a lot of problems in Japanese, which directly affects the overall improvement of the level of Japanese language education in China. The main reasons are: ① insufficient attention to voice teaching, there is a general emphasis on grammar, the phenomenon of light voice, due to class time and teacher concepts and other reasons, many teachers compressed voice teaching hours, there is no systematic and detailed explanations, students in the case of half-knowledge directly into the vocabulary, grammar, from the time of the study and the number of research papers, the study of Japanese voice teaching also lagged behind the study of vocabulary, grammar, which also can be verified from the side; ② voice teaching research is also lagging behind the research of vocabulary and grammar, this is also the case of voice teaching. From the side to verify; ③ the lack of professional knowledge of speech teaching, some teachers lack of professional and systematic knowledge of speech teaching, lack of confidence in language teaching, do not have enough ability to find out the speech problems or know what is true but not know why, resulting in the overall quality of speech teaching is not high; ④ teaching methods are outdated. At present, most of the Japanese speech teaching still takes the traditional teaching mode of imitating and following reading, in which the teacher gives a model pronunciation, the students imitate and follow the pronunciation, and the teacher corrects the mistakes. Due to the lack of systematic pronunciation methods and pronunciation guidance, it is difficult for students to recognize the difference between their own pronunciation and the model pronunciation, not to mention to improve their pronunciation.

5 Feasibility and Necessity of Applying Speech Visualization Technology

(1) Feasibility of applying speech visualization techniques

Foreign language teaching methods have been changing with the trend of the times and the updating and development of linguistic theories. In the 1980s, with the emergence of functional linguistics theory, more and more attention was paid to the communicative function of language, and Long & Robinson (1998) pointed out that Focus on Form, i.e., explicit teaching methodology that favors the form of the language, is effective, especially in the area of phonetics and phonology, and De Keyser (1998) pointed out that pronunciation problems of foreign language learners are more important. This is especially true for phonetics and phonology, and De Keyser's (1998) study pointed out that the pronunciation problems of foreign language learners can only be improved after a long period of intensive training. With the continuous development of computer technology, many excellent speech visualization software such as Praat, Audition, Better Accent Tutor, Speech

Anglyser not only realize the characterization of speech features, making waveforms, audio and other articulatory data visualized, but also can draw spectral graphs according to the needs of the speech spectrum, and by extracting, analyzing, and changing the By extracting, analyzing, and changing the acoustic physical parameters, we can produce synthesized speech with different pitch accents and intonation types to carry out discrimination and speech perception training.

(2) The need for applying speech visualization technology

At present, the most common method of phonics teaching is still imitating and following reading. Teachers provide model pronunciation, students imitate and practice reading, and teachers correct the pronunciation. However, more practical studies have shown that even if learners repeat the imitation practice, the pronunciation problem cannot be improved effectively (Kono and Matsuzaki 1998). From the perspective of cognitive brain science, the "perception-cognition-output" process of the human brain is very complex and prone to cognitive bias, and different learners have different cognitive strategies and knowledge construction due to the differences in geographic location and the influence of migration of their mother tongue, so it is not possible for them to listen to and recognize the pronunciation of the words, not to mention to improve the pronunciation of their own pronunciation problems. The visualization of pronunciation helps learners to strengthen the practice of poorly mastered pronunciation according to their own cognitive situation and acceptance level without the limitation of time and region, and to better evaluate and correct their pronunciation through visualization, which is a necessary way to realize differentiated and effective learning.

6 Path of Realization

The application of speech visualization technology is of great significance in promoting the reform of Japanese speech teaching and realizing the effective teaching of Japanese language under the human-computer cooperative technology. The realization of visualization teaching involves the development and promotion of the platform, the selection and use of software, the transformation of the teaching mode, the guidance of teachers and the careful design of teaching contents.

① Establish a comprehensive and systematic voice learning library. In recent years, the Japanese language phonics teaching model has evolved, and teaching resources are constantly being updated. The online Japanese Pitch Dictionary is designed to help Japanese teachers and Japanese learners to improve problems such as centering stress and sentence rhythms. By reading aloud the tutor, Suzuki-kun, the pitch stress and intonation characteristics of the text can be displayed, initially realizing voice visualization. After that, the first Japanese pronunciation course JPC was launched on the Moocs edx platform, which allowed more Japanese learners to learn Japanese pronunciation skills and voice characteristics for free. However, students could only see the difference between their own pronunciation and the pronunciation of the target language, and there was a lack of analyzing and evaluating pronunciation data, which prevented them from evaluating their own pronunciation problems, and from mastering tuning and correcting methods. Therefore, it is necessary to establish a comprehensive and systematic frontal voice learning library, and select different modes of visual voice analysis system to extract and analyze the voice data for different students' pronunciation problems, and adjust the parameters to adjust the corrected voice, so as to effectively improve the effect of voice learning.

② Visualization of voice teaching in a blended teaching mode. Speech visualization teaching suggests the use of online and offline blended teaching mode, breaking the limitations of voice

course hours, better integration and use of high-quality network resources and platforms, the teacher in the classroom before the students watch and learn teaching videos, familiar with the rules of pronunciation and voice characteristics, the classroom teacher explains and demonstrates the process of visualization of the pronunciation and the method of the voice waveforms, fundamental frequency and other visual interfaces and voice analysis data dynamically presented, students learn to identify and adjust the pronunciation, through repeated practice, submit recorded assignments, while students comment on feedback, and mutual assessment between them. Students learn to recognize and adjust pronunciation, through repeated practice, submit recorded assignments, teacher comments and feedback, and at the same time, students evaluate each other to cultivate students' ability to consciously listen to and recognize pronunciation, voice monitoring and self-assessment.

③ Combination of segmental and suprasegmental instruction. At present, the research on phonological teaching focuses more on the guidance on the length of sound, stress and rhythm at the level of suprasegmental segments, emphasizing the coherence of rhythm and speech flow, but the guidance at the level of segments is still indispensable. Especially in the early stage of phonological learning, phonemic and phonemic pronunciation guidance can be carried out through the lip shape and mouth shape diagram, tuning cross-section diagram, Praat spectrogram, and so on, and gradually transitioned to the level of suprasegmental teaching, which can be practiced through the extraction of pitch, duration, frequency and other parameters of the sound intensity, and even change the parameters of the audio, such as tuning and chorusing. This requires teachers to flexibly choose the content displayed in the speech map according to different teaching contents, and choose to analyze the extracted parameters in order to carry out more comprehensive and systematic voice teaching instruction.

7 Conclusion

Visual speech technology is not widely used in Japanese language speech teaching at present, but with the rapid development of human-computer interaction and human-computer collaboration technology, visual human-computer interaction through speech visualization technology will surely shine in the near future. Deeply embracing artificial intelligence technology, integrating the advantages of speech software technology, and developing a speech visualization platform, in addition to the need for modern teachers who are skilled in the use of software technology and have the ability to develop the platform. Teachers are the core of the teaching process, and it is imperative to improve teachers' informatization literacy and cultivate a high-quality teaching force. The author will carry out subsequent research based on this to realize the high-quality development of Japanese language education.

References

- [1]Yang Y. A study on the effectiveness of phonological visualization Praat software applied to middle school English phonics teaching[D]. Harbin Normal University,2023.
- [2]Yang Haoxin.Research on the application of speech visualization technology in English speech teaching under the guidance of POA theory[J]. Teacher,2022(22):45-47.
- [3]Liu Jiaqi,Li Xin. Effective Japanese Speech Teaching Based on Acoustic Visualization Technology[J]. Contemporary Foreign Language Research,2022(03):149-56.
- [4]Chen Yiping,Zhang Zhenyu. An empirical study of the effectiveness and generalization of visual feedback paradigm on English phonological learning[J]. Higher Architectural Education,2022,31(01):171-80.
- [5]Xing Nan. Listening and Speaking Teaching Practice Enabled by Speech Visualisation Technology - An Example of

- Preparation for the Imitation Reading Section of the English Listening and Speaking Test of the Guangdong College Entrance Examination[J]. *English Learning*,2021(51):21-24.
- [6]Yang Haoxin. Research on the effectiveness of speech visualisation technology in assisting English intonation acquisition[J]. *Heilongjiang Education (Theory and Practice)*,2020(07):89-90.
- [7]Wang Y. A study on visualisation teaching of English monophthong segments - based on migration theory and experimental phonetics approach[J]. *Journal of Jiamusi Vocational College*,2020,36(03):163-64.
- [8]Liu Dan. Research on visualisation teaching in English phonetics course - taking intonation teaching as an example[J]. *Cross-linguistic cultural research*,2019(00):291-300.
- [9]Liu, Rhinoceros. Analysis of Chinese Japanese learners' phonological problems and instructional countermeasures based on Praat software--taking the vowel /e/ as an example[J]. *Journal of Jilin Radio and TV University*,2019(04):31-33.
- [10]Chen H. A study on the rhyming features of college students' English read aloud discourse based on speech visualisation[J]. *Journal of Southwest University of Science and Technology (Philosophy and Social Science Edition)*,2019,36(01):82-88.
- [11]Yang Hui,Guo Qing. The auxiliary role of speech visualisation software Praat in higher vocational English speech teaching[J]. *Campus English*,2018(36):100-01.
- [12]Liu Jiaqi. A study on the visualisation of teaching Japanese phonological segments - based on mediational theory and experimental phonology method[J]. *Japanese Language Learning and Research*,2018(04):77-84.
- [13]Sun Jianying. Research on the Effectiveness of Speech Visualisation Praat Software Assisting English Speech Teaching [D]. *Jiangsu University*,2018.
- [14]Liu Jiaqi. Practice and Prospects of Japanese Speech Teaching-Theory, Methods and Achievements[J]. *Higher Japanese Language Education*,2018(01):43-55+194.