

Analysis of the Artistic Elements of Broadcast Hosting Based on Media Speech Corpus

Zexuan Huang

Communication University of China, Beijing, China

Abstract. Whether it is speech synthesis research or speech recognition research, it depends on the construction of excellent speech corpus at the back end to some extent. The phonetic corpus includes three forms: text database, phonetic database and acoustic parameter database. Based on phonetic facts, it serves the basic research of language and the processing of language modernization technology through the functions of natural phonetic collection, phonetic annotation, retrieval and statistics. The purpose of the construction of the media corpus is to serve the research of the audio language of the media, especially the prosody research of the host language. Broadcasting and hosting is an important part of radio and television, and the diversity and richness of modern radio and television programs also put forward new requirements for broadcasting and hosting. Broadcasting determines the basic style, purpose and orientation of radio and television news programs. In the media industry, the status of broadcasting and hosting art is gradually highlighted. Under the background of media convergence, broadcasting and hosting art needs to be developed in more advanced technical fields, and the media voice corpus is an important product created by the development of science and technology.

Keywords: Media; Phonetic corpus; Broadcasting and hosting art; Data processing.

1. Introduction

With the development of corpus linguistics, it has been widely recognized that it plays an important role in the field of language research and natural language processing, and various forms of corpus construction have also made great achievements. The use of corpus in various fields has become very common. The media language corpus under construction is a comprehensive corpus system featuring audio media and radio and television corpora, which includes two closely related sub-corpora, the phonetic corpus and the text corpus. To build a framework of speech database, we should first consider its purpose. The speech database used for speech recognition is different from the speech database used for speech synthesis. Speech database for speech recognition is to provide training data for acoustic layer. It not only involves linguistics, but also needs to comprehensively reflect the acoustic characteristics of speech from the perspective of phonetics and acoustics, so as to provide complete data for establishing speech models. With the continuous development of science and technology, the mode of cultural transmission has also changed, from the original simple printing process to the current cultural feast of visual and auditory coexistence. Under the guidance of phonetics and linguistics, the design of a scientific and reasonable continuous speech database is of great support to the speech recognition and synthesis system.

We know that in the past, among all kinds of media, TV was the most important media to spread art for a long time, but today, with the continuous development of the times, the network has penetrated into thousands of households and is closer to people's lives. Moreover, the art of broadcasting and hosting needs to show its unique charm through a wide space. Corpus is an effective and indispensable research means and plays an irreplaceable role in speech science research and technical applications such as speech recognition and synthesis and speech analysis. Speech corpus plays an important role in corpus-based speech synthesis or recognition system, which directly affects the quality and efficiency of speech synthesis or recognition. In this corpus, sound and text coexist, original data, transcribed data and annotated data exist at the same time, and there are all kinds of complicated relations between them. The media voice corpus mainly provides services for the teaching and scientific research of the language art of broadcasting hosts, and it has very important practical significance for studying the art of broadcasting hosts.

2. Method

2.1 Design and Development of Phonetic Corpus

Language includes three basic attributes: form, meaning and sound. Phonetics is the most important attribute. Language first exists in the form of phonetics. There are languages with or without words in the world, but there is no language without phonetics. Phonetics plays an irreplaceable role in language. Although pronunciation is a kind of sound, it is essentially different from ordinary sounds. When we determine the pronunciation materials, according to the principle of pronunciation text selection, there are four parts: one is the single word part. The second part is vocabulary. The third part is the sentence material part. The fourth part is the natural dialogue part. After recording the pronunciation text, it is necessary to analyze and process the speech data to get the different features of the speech signal, which is the key to the design of the speech corpus and the necessary foundation for the later speech processing. For the recorded speech signal, the acoustic parameters of each segment need to be extracted. Intelligent speech technology based on speech synthesis, speech recognition and natural language understanding is a challenging and difficult high-tech field. Many researchers have made a lot of research achievements in this field.

The goal of speech synthesis is to make the computer emit speech with distinctive personality, natural fluency and emotion like human beings. Speech production is a process from the deep structure of conscious expression to the surface structure of sound wave output, including three stages: construction, conversion and execution, which also corresponds to the three stages of speech synthesis system (front-end processing, prosody analysis and synthesis), as shown in Figure 1.

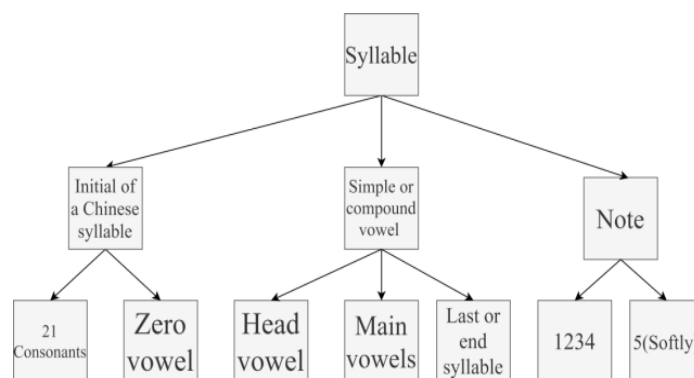


Figure 1. Schematic diagram of speech synthesis system

Language Corpus is a comprehensive corpus system which faces the field of broadcasting and hosting art and takes radio and television news corpus as its main content. It includes two closely related sub-corpora, namely, text corpus and phonetic corpus. In order to ensure that the training of acoustic model can describe each primitive and its co-pronunciation phenomenon more accurately, each phonetic unit in the corpus should reach a certain coverage and balance. Generally speaking, the larger the primitives, the better the effect and the more the number, and the speech library will be very large and difficult to cover all primitives; Small primitives, strong flexibility, and more splicing points will lead to the degradation of speech quality. The media corpus should be representative, and strive to reflect the actual use of audio language in radio and television. Its main purpose is to serve the research of audio language in media, especially the prosody research of audio language in media. In practical application, the appropriate primitive model should be selected after considering the scale of speech database and covering all kinds of language phenomena. A good corpus construction means that sparse training sets are needed to enrich the feature space of as many phonetic units as possible. The essence of this process is how to perfect a mathematical model about coverage. The coverage formula is as follows: let $F = \{1, \dots, N\}$, N be the characteristic factor number, any $\forall i \in F$, define the specific factor number F_i as a discrete set, $F_i = (F_1, \dots, F_\xi)$, where $\xi = |F_i|$. Define the feature space

$F = (F_1 \times F_2 \dots F_N)$, so that each feature value ρ corresponds to a specific vector. The formula is as follows:

$$f(p) = (f_1, \dots, f_N \in F), f_i \in (1 \leq i \leq N) \tag{1}$$

The addition model is another method that we consider. The eigenvalue corresponding to the eigenvector f is $D(f)$, and the formula is as follows:

$$D(f) = D_{\text{mean}} + \sum_{i=1}^N D_i(f_i) \tag{2}$$

Where D_{mean} is a constant, and $\sum D_i(f_i) = 0, 1 \leq i \leq N$. Therefore, at least $1 + \sum_{i=1}^N (\xi_i - 1)$ sample is needed to solve all the parameters.

More than 4 hours' corpus (including 14 male voices and 12 female voices) were selected from the media speech corpus for statistical analysis. The basic information of the selected corpus is shown in Table 1:

More than 4 hours' corpus (including 14 male voices and 12 female voices) were selected from the media speech corpus for statistical analysis. The basic information of the selected corpus is shown in Table 1:

Table 1. Specific distribution of selected annotated corpus

Linguistic form		Representative radio and television programs	Temporal distribution
Soliloquy	Broadcast	Point news, News broadcast, National news broadcast, etc.	1:08:56
	Explain	Stories, World Weekly, China Weekly, etc.	0:45:35
	Read aloud	Mainly literary works.	1:10:35
	Talk	Stories, World Weekly, China Weekly, Immediately etc.	0:50:53
	Comprehensive	News horizon, etc	0:18:03
Amount to			4:14:02

Even if these live actual sound or background music are filtered out, the extracted voice data, especially the pitch data, will be greatly distorted, so it is of little significance to compare and analyze this kind of voice data.

When synthesizing, the same kind of speech primitives appearing in different environments are processed by digital signals to adapt to different environments. With the development of waveform splicing system, later systems record multiple samples for each kind of phonetic primitives in the phonetic database. When synthesizing, choose different samples in different environments. According to the requirements of speech corpus for speech synthesis and speech recognition, and taking into account the characteristics of Chinese speech, the following criteria are put forward for sentence selection: 1. Select three-tone submodel as the basic unit for describing speech; 2. Covering all kinds of language phenomena to the maximum extent with as little corpus as possible; 3. In the speech synthesis and speech recognition system, the selected sentences should contain as many high-frequency words as possible; 4. Avoid data sparseness, so that each triplet appears a certain number of times. The corpus selection process is shown in Figure 2.

In this way, it is ensured that the sentences with many high-frequency words will be considered first when selecting the corpus. The establishment of phonetic corpus is a complicated problem, and there are still some problems to be improved for the later improvement of phonetic corpus, such as making full use of the existing inter-word tone sandhi rules to reflect the actual situation of tone sandhi and light tone as much as possible. For the lack of corpus, we can improve the utilization rate

of existing corpus in preprocessing. In view of the above reasons, the voice database should be an open database, so that changes can be added at any time to improve the database.

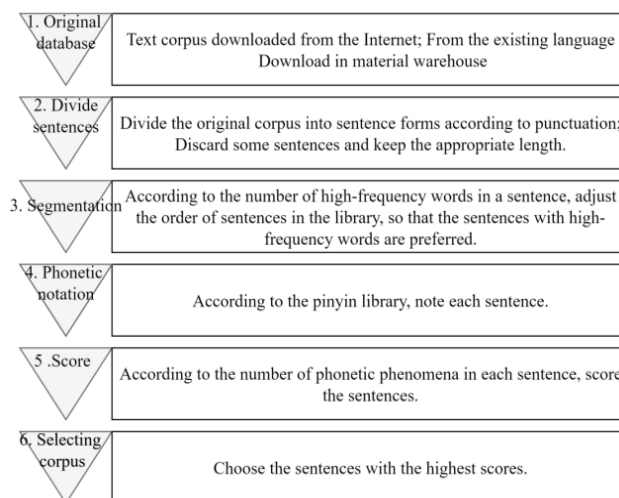


Figure 2. Corpus selection process

2.2 The Art of Broadcasting and Hosting under the Background of Media

It is more and more important to carry on the cultural function of the media. Compared with other types of cultural carriers, modern media has more diversified grasp of cultural characteristics, and it is an artistic aesthetic expression with comprehensive characteristics. In this society, everyone is an independent individual, and each person has his own characteristics, which is individual difference. In fact, the external image and the internal connotation of the broadcasting host are dialectically unified, and the external image is the expression and formation of the internal connotation, which is the solid foundation of the external image. The two support and unify each other, so it also includes the pronunciation ability of each person and the characteristics of his voice, which mainly refers to the pronunciation of each person. Television has promoted and created the spread of visual culture, and changed the original way of communication-print communication. Before TV media, people wanted to receive information, either by word of mouth or by reading books, which was inefficient and slow. However, the TV media has abandoned this drawback, and through the TV programs with sound and image signals, it spreads by video, which not only greatly improves the speed of information dissemination, but also greatly improves the efficiency of information dissemination. Thus, the artistic effects and themes of the envoys can be fully disseminated, and the artistry of the broadcast host can be perfectly displayed from the perspective of modern media. With the complexity of the media environment, the new media orientation has changed. Therefore, the broadcast host needs to invest more energy to analyze and study the audience market. Only by fully mobilizing the enthusiasm of the audience can we provide more choices for the audience under the background of rich choices of the current audience. Therefore, it's very important to study the artistic aesthetic connotation of the radio host in detail, for how to shape excellent radio host and build a successful radio program which is welcomed and loved by the audience. (As shown in Table 2)

Therefore, broadcasting and hosting programs should first grasp their visual and artistic image and artistic connotation. In the process of the development of TV livelihood news programs, the arrival of the new media era, on the whole, will have more advantages than disadvantages, which will contribute to the development of TV livelihood news programs, and promote the artistic development of broadcasting and hosting programs from the perspective of modern media in combination with the comprehensiveness of current communication laws. New media technology should be adopted to innovate the production and dissemination of TV news about people's livelihood, so as to promote the better development of TV news about people's livelihood. Once the host appears in front of the audience, it is no longer just himself, but a general artistic image. What he conveys is not only the knowledge structure and life experience of the host himself, but also the values embodied in a grand

social culture. As broadcast hosts, language is an important means for them to constantly create and communicate with the audience. Therefore, it is the most basic ability that broadcast hosts should master under the media background and one of the important elements of broadcast host art to use the acceptable voice of the audience and use appropriate language expression skills.

If you want to do a good job in broadcasting, you must firmly establish the goal of serving the people, establish correct social values, and improve your professional skills and self-cultivation. Grasp the pulse of the times and express the new voice of the times. Broadcasters must have rich knowledge reserves to improve their sensitivity to all kinds of people when they receive languages. Choose a reasonable way of language application. At the same time, it is necessary to further strengthen the understanding of current events by broadcast hosts, so as to ensure that the materials they cite have strong timeliness. Under the background of media, only with good appearance, rich connotation, high value orientation, sound and appropriate language expression skills that can bring good hearing experience to the audience, can the broadcast host greatly improve the probability of success of the column.

Table 2. Strategies to show the artistic charm of radio and television broadcasting hosts

Tactics	Main content
Create a good auditory image	Including the quality, accomplishment, connotation and ability of radio and television broadcasters. You should not only be articulate, good at communication and expression, but also have a keen mind and good psychological quality. You should be able to deal with any unexpected situation and control the radio and television programs freely, so as to bring the audience a pleasant hearing feeling.
Understand the social value orientation	Due to the differences in positioning, planning, organization and dissemination of radio and television programs, and the differences in social functions such as moral demonstration, government mouthpiece and social voice, radio and television programs put forward different social value orientations. Broadcasting art must realize social functions such as cognition, education, publicity and aesthetics in its creation, and establish its own artistic image in the social and cultural horizon.
Familiar with language pronunciation skills	In the host of radio and television programs, language should first give people a good aesthetic feeling in hearing, which can arouse the audience's basic understanding and intuitive evaluation of the cultural connotation and personality cultivation of radio and television hosts.
Apply the broadcasting law	The broadcasting host of radio and television stations must obey the basic laws of broadcasting programs of radio and television stations. In the context of a specific era, we should take care of the discretion and scale of the host of the radio program.

3. Result analysis

The design of the overall system structure is to divide the whole system into various functional modules reasonably, so as to correctly handle the relationship between modules and the inside of modules, as well as the data connection between them, and define the internal structure of each module. Includes five functional modules: user registration module, user administrator login module, recording module, database maintenance module and query download module. Database management system is a computer core subject that develops rapidly with the development of computer science and computer hardware in order to meet the demand of information society for data management technology. The main reason for choosing a database is whether it can support your application well. We can use the following aspects as reference factors for database selection: the convenience of database construction; Support for program development; Performance analysis of database; Data integrity constraint; Concurrency control function; Safety; Chinese character processing. A complete

phonetic corpus needs not only original phonetic data, corresponding pronunciation texts, but also corresponding annotation files. To improve the utilization value of speech corpus, the key is to mark the speech corpus completely. The tagging process of phonetic corpus is a formalization process of language knowledge. The generated preliminary voice annotation file can reduce a large number of tedious text input work and save manpower and material resources in the process of annotation of voice files. Fig. 3, fig. 4, and fig. 5 show an automatically generated preliminary markup file that is not aligned with the voice file.

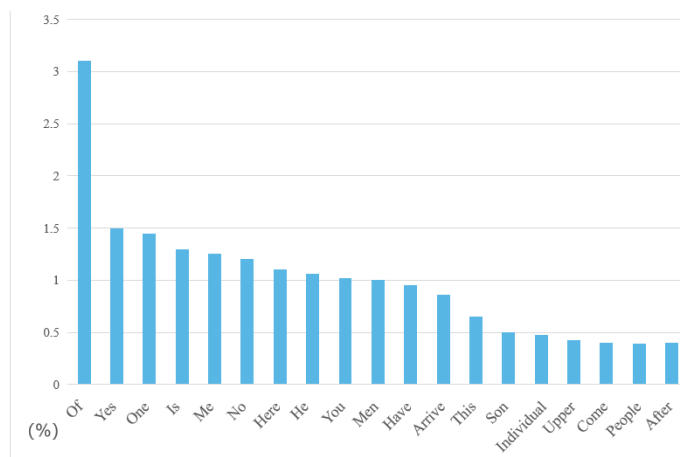


Figure 3. The frequency of the first 20 Chinese characters in the original corpus

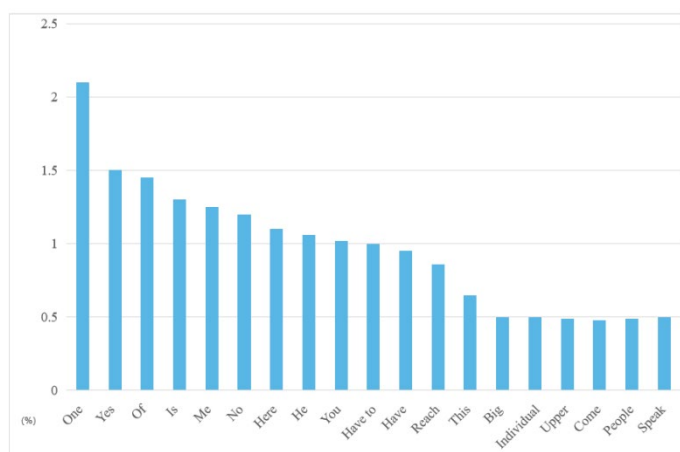


Figure 4. The frequency of the first 20 Chinese characters in phonetic corpus

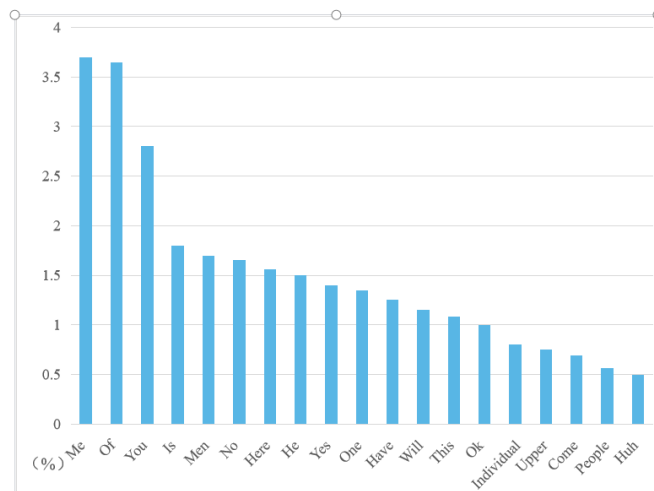


Figure 5. The frequency of the first 20 Chinese characters in the dialogue corpus

After screening, the corpus covers 30 common words in Chinese Common Words List and 30 common words in Modern Chinese Common Words List, with the coverage rates of common words and common words reaching, and respectively. For common words and phrases that cannot be covered by the corpus, we will automatically generate the corpus texts of these words and phrases, and store them in our corpus in the form of word stock or word stock. The establishment of database management system in phonetic corpus needs to store four kinds of materials: one is the speaker's attribute materials, such as the speaker's age, gender, education, mastery of Chinese, and my use of mother tongue. Second, pronunciation text materials, inputting and storing the pronunciation materials of the speaker and the corresponding dialect pronunciation, Putonghua international phonetic symbols and other text materials. The third is the actual voice data material, which is mainly used to save the original parameters of the recorded voice waveform graphics. Fourth, acoustic analysis parameter data, that is, preservation of acoustic parameters extracted from processed speech waveform. At the same time, in the process of spreading TV program information, the broadcast host should learn to make good use of his own artistic conditions, and combine the overall style and program characteristics of TV programs, so as to form a language art with strong influence and own personality. The improvement of radio and television director's ability in the new period requires directors to keep pace with the times and constantly improve their director's ability, innovation ability and comprehensive level, so that radio and television works can have higher quality and level.

4. Conclusions

With the rapid development of computer technology, speech recognition and speech synthesis technology have also made rapid progress, and more and more technologies have been applied to practical products. Products such as speech input system and computer-assisted language learning system are constantly emerging, providing superior services for people's study and life. Excellent speech synthesis and speech recognition system, a speech corpus with high information content and low redundancy is essential. It can be seen that speech corpus plays an important role in speech recognition, speech synthesis and speech research in other fields. Broadcasting is an important part of TV media, which has a very important influence on the style and orientation of TV programs. Its artistic image, aesthetic connotation and appeal have an important influence on the media effect. In the following work, we can further enlarge the scope and information of the original corpus, so that the selected corpus text can cover more phonetic phenomena. It is also possible to further expand the size of the screened corpus text, so as to expand the information content of the whole speech corpus. Only in this way can the broadcasting host accelerate the improvement of the social influence and penetration of TV media in the declining status of TV media and the fierce competition of multimedia technologies. More importantly, it reduces the dependence of media language corpus on third-party tools, and makes a big step towards the goal of a completely independent media language corpus management system.

In the aspect of voice recording and voice database management system, the workflow and design ideas of the recording module are described in detail, and the voice database management system is realized, which facilitates users to inquire and download corpus texts, voice files and annotation files, and tests the recording work and establishes a preliminary voice database. Language corpus is a comprehensive corpus which faces the field of broadcasting and hosting art and features the news corpus of radio and television. The first phase of its text corpus project has been completed and completed. With the continuous development of the Internet, the analysis of the artistic elements of broadcasting and hosting based on the network background is an important way to improve the quality of current broadcasting and hosting programs.

References

- [1] Song Jingjing. A Comparative Study of Chinese Style Vocabulary Based on Media Corpus. *Journal of Changsha University*, vol.no.6, pp.4,2015.
- [2] Yuan Jiahong. Collection, processing and research of large-scale phonetic corpus. *Linguistic Research*, vol.no.1, pp9,2017.
- [3] Jian Xu, Reyiman Tursun, hushur islam. Research and implementation of online multilingual speech corpus platform. *Information and Communication*, vol.no.4, pp.4,2018.
- [4] Yang Jian, Li Haiguang, Zhang Xiaoling. A Study on the Construction of Bai Phonetic Corpus. *Journal of Dali University*, vol. 2, no.2, pp.6,2017.
- [5] Zhang Jielan. The innovation of radio and television broadcasting hosting art in the new era. *New Media Research*, vol.2, no.004, pp.159-160,2016.
- [6] Li Xiaoxue. Exploration of the innovative space of broadcasting hosting art in the new era. *Popular Literature and Art: Academic Edition*, vol.no.2, pp.1,2016.
- [7] Liang Nan. On the Art of Broadcast Hosting-Making Broadcast Hosting more personalized, personalized and friendly. *Think Tank Times*, vol.no.12, pp.2,2017.
- [8] Wu Huiping. Innovative discussion on the art of radio and television broadcasting in the new era. *Western Radio and Television*, vol.no.13, pp.1,2016.
- [9] Oqir Asna, Buyin. The innovation of radio and television broadcasting hosting art in the new era. *Science and Technology Communication*, vol.no.15, pp.2,2016.
- [10] Dark Dream Qi. Innovation Space of Broadcasting and Hosting Art. *Educational Science: Full Text Edition*, vol.000, no.003, pp.00202-00202,2016.
- [11] Lu Ming. Innovative research on the art of radio and television broadcasting in the new era. *Information Weekly*, vol.no.31, pp.1,2019.
- [12] Cheng Yong. Discussion on the elements of broadcasting hosting art from the perspective of media. *New Media Research*, vol.no.6, pp.2,2016.
- [13] Nie Yuanfang. Exploration of innovative space of broadcasting hosting art in the new era. *News Research Guide*, vol. 8, no.11, pp.2,2017.
- [14] Du Qianqian. On the innovation of radio and television broadcasting hosting art in the new era. *Western Radio and Television*, vol.no.17, pp.1,2015.
- [15] Xu Jiaxin. Summarize the innovation and future development of broadcasting hosting art. *Naked*, vol.no.12, pp.1,2015.