Characteristic Difference Analysis of Vowel /A/ Phoneme Variants

Hongbing Zhang

Criminal Investigation Police University of China, Shenyang, China

zhanghongbing@cipuc.edu.cn

Keywords: Voiceprint Identification; Vowel; Phoneme Variant; Acoustic Characteristics

Abstract: In Chinese pronunciation, the characteristics of vowels and their formants are often the first and key parameters in speaker identification. In this paper, Praat software is used to analyse the phoneme variant of vowel /a/, and morphological analysis and data statistics are carried out on studying the acoustic characteristics of the vowel /a/ in different speakers' pronunciation, such as the phonographic formant and energy, etc., to conclude the variation rule of the acoustic characteristics of the vowel /a/.

1. Introduction

Chinese is a language dominated by vowels, so vowels and their characteristics of formant are often the first and key parameters in speaker identification. Due to the change of opening degree when pronouncing a vowel caused by the changes in the language style and the speaker's emotions, psychology and other factors, there are differences in the pronunciation and actions of the vowels. We can take /a/ which is of the most obvious opening degree as the research object to analyse the influence degree and law of the acoustic characteristics of vowels. In addition, vowels are affected by closed transition in syllables, and the variation of their acoustic characteristics in different syllable combinations is also worth studying[1].

In this paper, the vowel /a/ is chosen as the research object, the spectrums of /a/ under different allophone are taken quantitative and statistical analysis. Also, data such as the resonance peak, length, intensity are collected, the distribution of high pitch, formant, proportion of fundamental frequency, duration, and the acoustic characteristics of and the transitional characteristics research summary of the main consonants. All of those are aiming at providing a new method for the voice print appraisal and exploring the specific application value of vowel variation regularity of words in court.

2. Analysis of vowel /a/ allophonic variation

The allophonic variation of vowels is mainly due to the pronunciation habit or the influence of other vowels or consonants, which makes the action of the vocal organs different from the normal state.

2.1 Pronouncing /a/ in different OPEN-MOUTHED sizes

When different people pronounce the same sound, due to personal pronunciation habits, their open-mouthed sizes are also different. Even when the same person pronounces the same sound, they will be affected by many subjective factors such as mood, speaking habit, intentional disguise, or by objective factors such as degree of pitch, different tone in different language habits and so on[2]. The effect of the different open-mouthed size can be attributed to the phenomenon of phonemic variation by changing the shape of vocal organs. When the size of our mouth’s changes, the most intuitive feeling is that the tongue position has changed, the intensity of the airflow has changed. The tongue position directly affects the position of the first formant, and the intensity of the airflow is closely related to the sound strength. Therefore, when pronouncing vowel /a/, the characteristic differences of different open-mouthed sizes mainly focus on the first formant and sound intensity.

2.2 The effect of the close juncture

Within syllables, the vowel /a/ is affected by the sound before and after it, which results in variants, such as /ia/, /uai/, /ai/, etc[3]. During the transition between different sounds of the same syllable, due to the deformation of the vocal organs, various transitions are formed when the shapes of formants change, and the formant tilt occurs at the transition, which is definitely different from the flat formant morphology when a single vowel is pronounced, and the allophonic variation occurs.

3. Experiment and analysis

In this experiment, three syllables of /a/ in different open-mouthed sizes and twenty syllables of /a/ with other vowels (i, u, o) or consonants (f, l, h, n,) (/a/is put in the first, middle or last of the syllable) were selected as the corpus[4]. In order to study the application value of allophonic variants of vowel /a/ to the greatest extent, homogeneous groups with similar conditions such as age and living environment were selected as the corpus recorder. In the experiment, five males and five females were selected, ten of whom were fully familiar with the corpus before recording, had sufficient recording experience, and had good pronunciation conditions.

In the experiment, the broadband spectrum of each syllable from each speaker were extracted. We can observe the sonogram features of vowel /a/ and artificially collect the formant frequency of the vowel /a/ in each syllable (including the starting value, the median and the closing value), analysis the acoustic features of /a/ under different allophone such as the distribution of high pitch, formant frequency, resonant amplitude curve and duration, and the transition characteristics of the main vowels. Meanwhile, the statistical software SPSS was used to analyse the formant frequency values[5].

3.1 Results and analysis of vowel /a/ in different open-mouthed size

After the comparison of ten speakers' vowel /a/ with different open-mouthed sizes, it can be found that there is indeed a change in the position of the formant, especially the position change of the first formant is the most obvious, and there is a certain rule between the change of the first formant and the change of the open-mouthed size: the greater the opening degree is, the higher the first formant is. The smaller the degree is, the lower the first formant is.

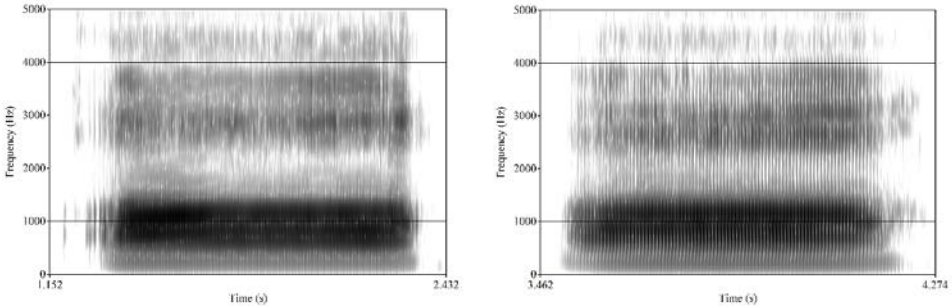

Figure 1. Speaker S1 pronounce/a/in large and normal open-mouthed size

Formant frequencies of /a/ in different open-mouthed sizes in the samples were measured and counted, and the mean values of each level of formant frequencies emitted by ten speakers under different opening degrees /a/ were listed in the table. On the whole, the average values of the first four /a/ formants under the normal opening degree of different speakers are not significantly different, which are roughly 785Hz, 1250Hz, 2750Hz and 3575Hz.In the case of deliberately changing the opening degree, the frequency of /a/ from the ten speakers varied greatly.

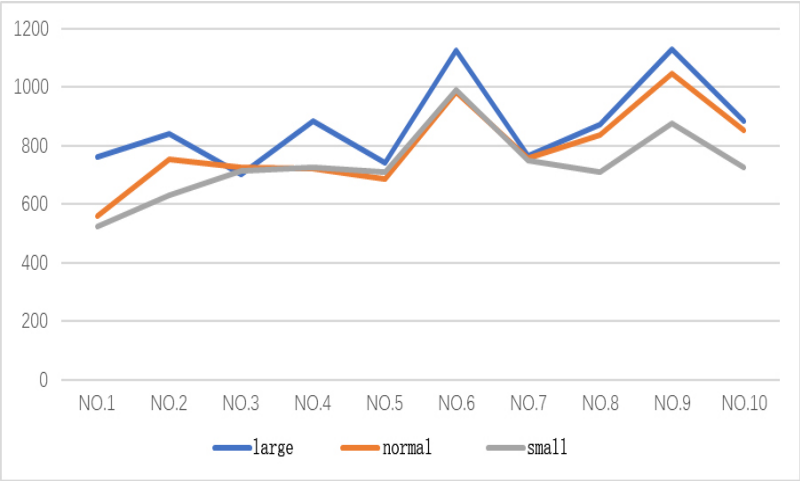

Figure 2. Average F1 values of /a/ in different open-mouthed sizes from 10 speakers

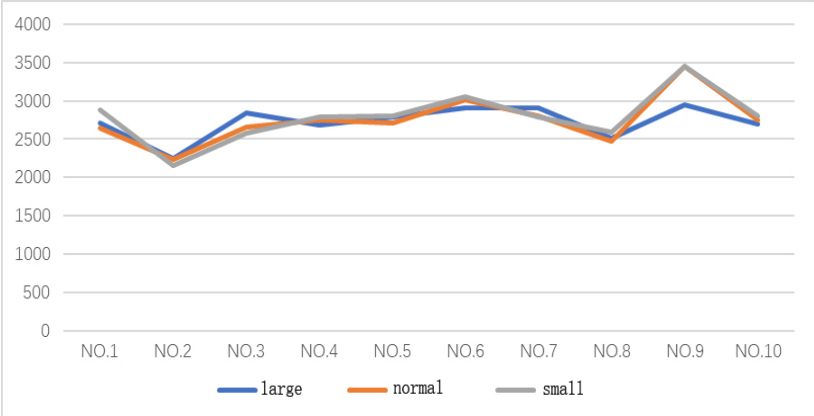

Figure 3. Average F3 values of /a/ in different open-mouthed sizes from 10 speakers

By comparing the formant data of vowel /a/ under different opening conditions for the same person, it can be found that the opening sizes have the most obvious influence on the F1 values. With the increase of the opening size, the value of F1 also increases[6]. Taking the speaker NO.2 as an example, the F1 frequency of /a/ under the condition of large opening size is 841.8Hz, and the frequency of the first formant under the condition of normal opening size is 752.1Hz, with a difference of nearly 100Hz.F1 is 630.1hz under small opening condition, which is also about 100Hz different from that under normal opening condition[7]. It confirmed the rule that the larger the size of opening is, the higher the formula 1 is, the smaller the size of opening is, the lower the formula 1 is.

3.2 Analysis of vowels /a/ in different pronunciation positions

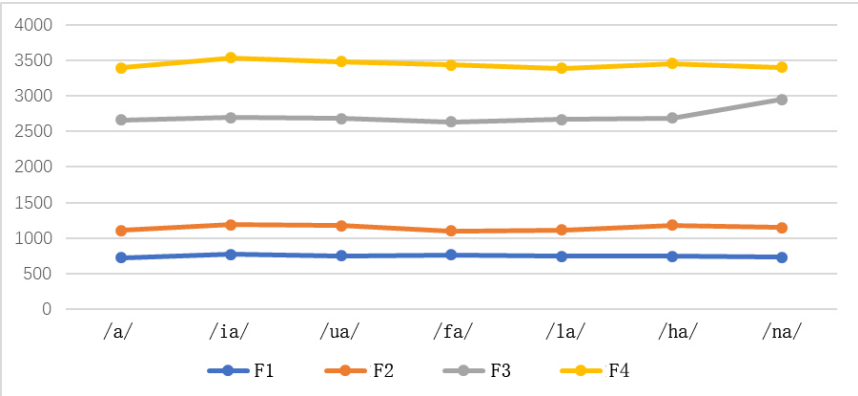

Figure 4. /a/ is located at the front end of the syllable

According to the above data, it can be found that due to the influence of the subsequent sound after the vowel /a/, transition phenomenon occurs at the end, which also affects the mean value of

the formant. Taking the speaker S3 as an example, the mean values of F, F2 and F4 of /ai/ and /an/ are all higher than the formant of the single sound /a/, while F3 is lower than the formant of the single sound.F1 and F2 of /ao/ and /ang/ are lower than the formant values of the unit sound /a/, while F3 and F4 are higher than the unit sound /a/.Therefore, when the vowel /a/ is located at the front of the syllable, there is a regular variation of the formant[8].

When the vowel /a/ is at the end of the syllable, the trend of the left formant hardly changes. In /la/, the left part of spectrum of /a/ shows transition, F1 shows a downward trend, F2 shows an upward trend, and F3 is almost unchanged. The transition is most obvious in /ya/.F1 has a downward trend, while F2 and F3 have an upward trend.

Although it is difficult to find the influence of the sound before /a/ in the spectrogram, the analysis of the formant value shows the variation of the formant of the vowel /a/. It can be found from the chart that the formant of /a/ at the end of the syllable increases at all levels, especially F2.

By observing and analysing the spectrogram in which vowel /a/ is in the middle of syllables, it is found that the transition between different sounds and vowel /a/ is obvious, which affects the formant of vowel /a/ at all levels. Taking the /uai/ of speaker S2 as an example, F1 of the vowel /a/ has a downward trend from left to right, and the mean frequency of its formant is bound to decrease[9].F2 has a downward trend on the left and an upward trend on the right, and the value of its formant does not change much.F3 tends to rise on the left and right sides, and the frequency of F3 is a little higher than that of unit /a/.

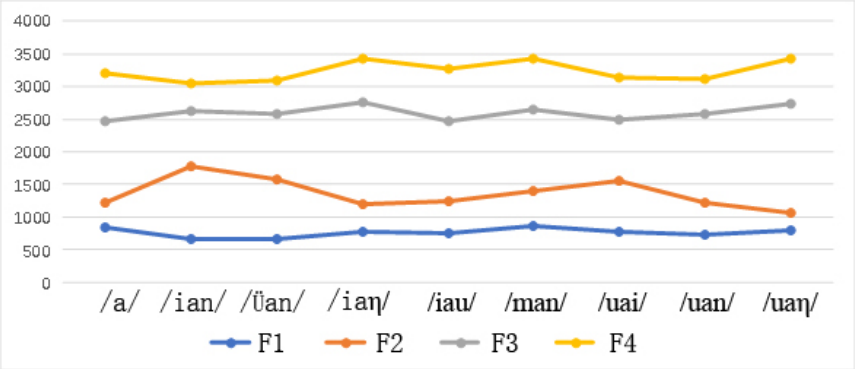

Figure 5. /A / formant data of corpus at the front of syllable

By analysing the formant data of ten speakers, it can be found that when /a/ is located in the middle of the syllable, F1 decreases and F2 increases, especially the three corpuses of /ian/, /an/ and /Üai/, F2 increases by nearly 30%. F3 and F4 all change, but there is no obvious rule.

Conclusion

Chinese is a language dominated by vowels, so vowels and their characteristics of formant are often the first and key parameters in speaker identification[10]. However, the phenomenon of allophonic variation has been neglected in judicial phonetic examination. Therefore, this article selects comparatively representative vowel sound /a/ as the research object, analyses the specific circumstances of allophonic variation, and its variation patterns, designs experiment researching the allophonic characteristic differences of vowel /a/, and collects the experimental data. Then, we can finally get the specific situations of allophonic variation of the vowel /a/ and the differences of the acoustic characteristics caused by the variation, and discuss the specific application value of vowels in speaker identification in court from the analysis of the acoustic characteristics.

Acknowledgements

This work was supported by Research on Dynamic Feature Extraction of Speaker Based on Deep Learning (Special funds for basic scientific research fees of Central Universities No.3242019012) a nd Research on the standardization of judicial application of audio-visual evidence in the context of Litigation Reform (No.2017LLYJXJXY040).

References

[1] Wang Yingli, Li Jingyang, Cao Honglin. A review of voiceprint identification techniques[J]. The Police Technology ,2012(4):54-56.

[2] Wang Yingli. Voice print identification technology [M]. Beijing: people's publishing house, 2013: 208.

[3] Eriksson A. Aural/Acoustic vs. automatic methods in forensicphonetic casework[M]// NEUSTEIN A, PATIL H.A. In ForensicSpeaker Recognition: Law Enforcement and Counter-Terror-ism. New York: Springer, 2011: 41-69.

[4] GOLD E, FRENCH P. International practices in forensic speak-er comparison[J]. International Journal of Speech Language andthe Law, 2011, 18(2): 293-307.

[5] Reynolds D A,Quatieri T F,Dunn R B.Speaker Verification Using Adapted Gaussian Mixture Models [J].Digital signal processing,2000,10(1-3):9-14.

[6] Cortes C,Vapnik V.Support-vector Network[J].Machine Learning,1995,20(3):273-297.

[7] Yangtian Jin, Hong Wang, Guoliang Ouyang.The influence of speech opening on formant frequency[J].Journal of Criminal Investigation Police University of China, 2014(3):48-51.

[8] Zehong Zheng.On the variability of vowel phonemes in Putonghua syllables[J].Journal of Polytechnic University,2011(2):103-105.

[9] Honglin Cao, Jianwei Liu.The acoustic difference between normal and loud speech and its influence on voiceprint identification[J].Journal of evidence science, 2009(6):754-764.

[10] Bing Li. Typology of harmony of vowels[J]. Journal of ethnic Chinese, 2001(2):32-46.