Studies in Chinese Phonetics

Aijun Li and Chunyu Ge

Contents

Abstract

The experimental study of Chinese phonetics started in early twentieth century in the context of a long philological tradition in China. Its development was interrupted by the war and political events during the first half of twentieth century, but resumed and eventually flourished after late 1970s. The works of Chao Yuenren, Wu Zongji, and other phoneticians have laid the foundations for in-depth studies of Chinese phonetics. Over several decades, studies of Chinese phonetics have covered all three fundamental aspects of speech, i.e., speech

A. Li (✉)
Institute of Linguistics, Chinese Academy of Social Sciences, Beijing, China
e-mail: liaj@cass.org.cn

C. Ge
Department of Linguistics and Modern Languages, The Chinese University of Hong Kong, Hong Kong, China
e-mail: chunyuge@cuhk.edu.hk

 1
Z. Ye (ed.), *The Palgrave Handbook of Chinese Language Studies*,
https://doi.org/10.1007/978-981-13-6844-8_21-1

production, acoustic phonetics, and speech perception. Particular attention has been paid to significant phonetic features of Chinese, such as tone and intonation. Phonetic studies of Chinese have also stimulated interests in interdisciplinary research. Speech pathology, forensic phonetics, speech acquisition, and phonetic learning of Chinese as a second language have been explored vigorously. Further studies in Chinese phonetics along these lines with integration with other disciplines and state-of-the-art technologies will not only deepen our understanding of Chinese, but of human language in general.

Keywords

Chinese · Phonetics · Production · Acoustics · Perception · Acquisition

With the study of sounds in ancient China emerging in the fourth century, the philological tradition has continued since then in what is known as historical Chinese phonology. The earliest explication of the sound system in Chinese could be traced back to *Qieyun*, a Chinese rhyme dictionary created by LU Fayan 陆法言 in 601 during the Sui dynasty (581–618). The book contains over 12,000 character entries, divided into five volumes, corresponding to the four tonal categories. These entries are divided into about 200 final rhyme groups, with their pronunciations given in the *fanqie* formula. No clear descriptions of tonal values were offered in the book, except highly impressionistic terms such as "light and shallow" and "heavy and hollow," found in the only extant preface to *Qieyun*. It is worth mentioning that SHEN Yue 沈约 (441–513) of the earlier Six Dynasties period (220–589) was the first one who recognized the existence of the four tonal categories in the *Treaties on Four Tones* that he compiled. The book is probably the earliest work on tones, but is no long available. Ancient Chinese scholars made enormous effort at classifying the Chinese language according to the syllable structure. The Chinese characters, which are entirely monosyllabic, are analyzed by the scholars and arranged in diagrams called *Dengyun Tu* (等韵图). In this kind of diagram, a syllable is represented by its initial on the column, and its final on the row. Initials are grouped according to their places of articulation, and the finals according to the height of the nucleus. This method represents the ancient scholars' attempt to understand the phonetics of Chinese by introspection. Another point to note is a story recorded by the scientist, SHEN Kuo 沈括 (1031–1095), in Song dynasty. The story said the local administers employed an instrument, which can be put into a dumb person's throat and let him pronounce in the court. This is an early and interesting application of phonetics in the ancient Chinese court.

For about a thousand years, Chinese historical phonologists mainly focused on the classification of tones and did not make much significant progress in identifying tonal values. The situation started to change with the arrival of missionaries during the Ming and Qing periods (1368–1911). They worked with a small number of Chinese scholars to transcribe Chinese sounds with Roman alphabet, including Chinese dialects, and produce Chinese textbooks with phonetic transcriptions.

Most philologists, however, immersed themselves in work of tonogenesis or evolution of tones, disregarding the actual phonetic values of tones. They made no attempt to discover principles of speech production, but it is fair to say that the system of sound changes and evolutionary paths they developed from the Middle Chinese onward is a remarkable accomplishment, which, in terms of its theoretical significance, is on a par with the distinctive feature theory as espoused in the SPE model of generative phonology.

The development of phonetics was driven by the intention to romanize Chinese characters. The intellectuals in the young Republic of China believed that the complex Chinese characters are hard for ordinary people, and caused the poverty of the nation. Romanization of Chinese characters was their first step in public education. A thorough and detailed understanding of the Chinese languages, especially the dialects spoken by most of the population, are extremely in need. While phonetic studies before the 1920s relied heavily on skills of hearing and imitating speech sounds within the purview of articulatory phonetics, LIU Fu 刘复 (aka LIU Bannong 刘半农) and CHAO Yuen-Ren 赵元任, who are the pioneers of the movement of New Culture in China, were among the first to use kymograph in the measurement of frequency of tones in the early nineteenth century. Their pioneering efforts started a new phase in the experimental study of speech in China.

This chapter seeks, in four sections, to provide a synopsis of the centennial history of the experimental phonetic studies in China, surveying important findings, scholars and milestones. It starts with a review of the overall development of studies in Chinese phonetics, followed by an overview of phonetic research in speech production, acoustic phonetics and speech perception. Application-oriented research is described in the third section on speech pathology, forensic phonetics, speech acquisition and cognitive development in children, and speech learning of Chinese as a second language. The fourth section offers concluding remarks. Due to the space limitation, this introduction of phonetic research itself and application-oriented research is succinct and highly selectively. Readers are referred to the references for more details about the research discussed in this chapter.

History of Phonetic Research in China

This section gives a comprehensive summary of the history of phonetic research in China, which began almost at the same time as that in Europe. The adoption of the experimental approach in studying Chinese concentrated mainly on lexical tones in the early days. Later researches were largely conducted in a few prestigious research institutions, most notably the Institute of Linguistics and Beijing University. The research areas and topics have been greatly expanded in the past few decades, and a wide array of equipment and experimental paradigms have been exploited.

Fig. 1 Prof. Liu Fu was recording Chinese dialectal sounds

The Beginning of Experimental Phonetics

Experimental study of phonetics in China began in the 1920s and had grown rapidly for two decades marked by the May Fourth Movement in 1919 till the outbreak of the Anti-Japanese War in 1937. In 1924, Liu Fu did research on Chinese tones with Tone Inference Scale he designed (Fig. 1 and 2), which can measure tone frequency on waves drawn with kymograph. He completed his doctoral thesis “An Experimental Study of Chinese Tones” (*Étude expérimentale sur les tons du chinois*) in Paris, which was later revised and published in China under the title *Experimental Study of the Four Tones*. Liu’s thesis is the first one in experimental phonetics ever written by a Chinese. In 1927, WANG Li 王力 also finished his doctoral thesis in experimental phonetics entitled “An Experimental Study of the Sound System in Bobai Dialect” in France, which was published in 1932 (Wang 1932). Liu Fu returned to China in fall 1925 and brought with him equipment to conduct phonetic experiments before setting up the Phonetics and Music Lab in the Institute of Liberal Arts at Peking University. He died of an acute illness at the age of 44 in Beijing on July 14, 1935, after a linguistic field trip in Inner Mongolia.

LUO Changpei 罗常培 took over as director of the lab shortly. He advocated using “experiments to compensate for the inadequacy of hearing” and was instrumental in prompting a shift in the study of Mandarin from the articulation-oriented tradition to the experimentation-based approach in modern phonetics (Fig. 3). In his research on Chinese historical phonology, Luo adopted the comparative method, commonly used in the study of Indo-European languages, and worked on materials from loan words and modern dialects. Among his work, *An Outline of General Phonetics*, coedited by Luo Changpei and WANG Jun 王均 (1957), laid out principles of phonetics in its description of the speech data in Chinese, and is still regarded one of the best readers in phonetics.

The founding of Academia Sinica in Nanking in 1928 was an important milestone in promoting scholarly research in sciences and humanities. Chao Yuen-Ren was appointed Head of the Language Section at the Institute of History and Philology and

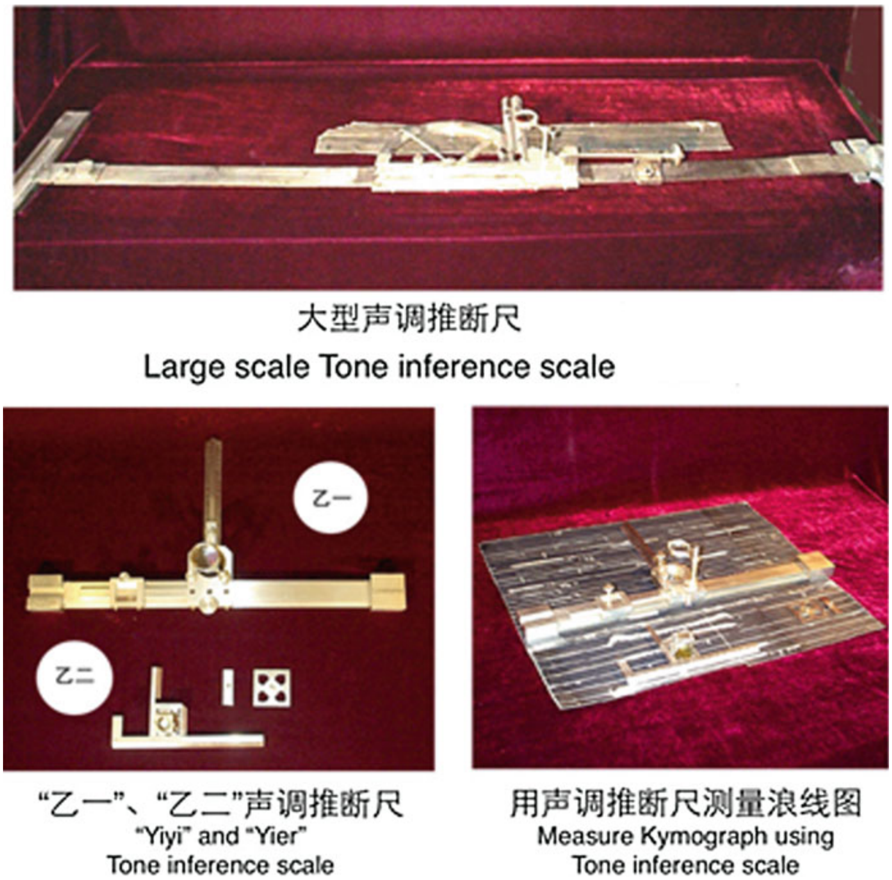

Fig. 2 Tone inference scales Yiyi (type I) and Yier (portable type II) invented by Liu Fu

put investigating Chinese dialects and establishing the phonetics lab high on the agenda, with priority given to the development of phonetics as a field. As illustrated in Fig. 4, Chao also enthusiastically conducted field work in various Chinese dialects. The field work was summarized in a seminal work entitled *Studies in Modern Wu Dialects* (Chao 1956), which was also initially published in 1928. In this monograph, he transcribed Wu Chinese dialects in an adopted version of IPA. Tones were transcribed in the time-pitch graph using a sliding pitch pipe. The book was considered the epitome of fieldwork on Chinese dialects. In 1930, Chao invented the “tone letters” used in IPA for the transcription of tones (Chao 1930/2002: 713–717). Chao’s phonetic work was also manifested in his contribution to phonological theory. In a classic paper on structural linguistics entitled “The non-uniqueness of phonemic solutions of phonetic systems” (Chao 1934/2002: 750–795), Chao argues that a phonemic analysis serves multiple functions and there is no unique solution that serves all functions. Chao established the phonetics lab in 1935

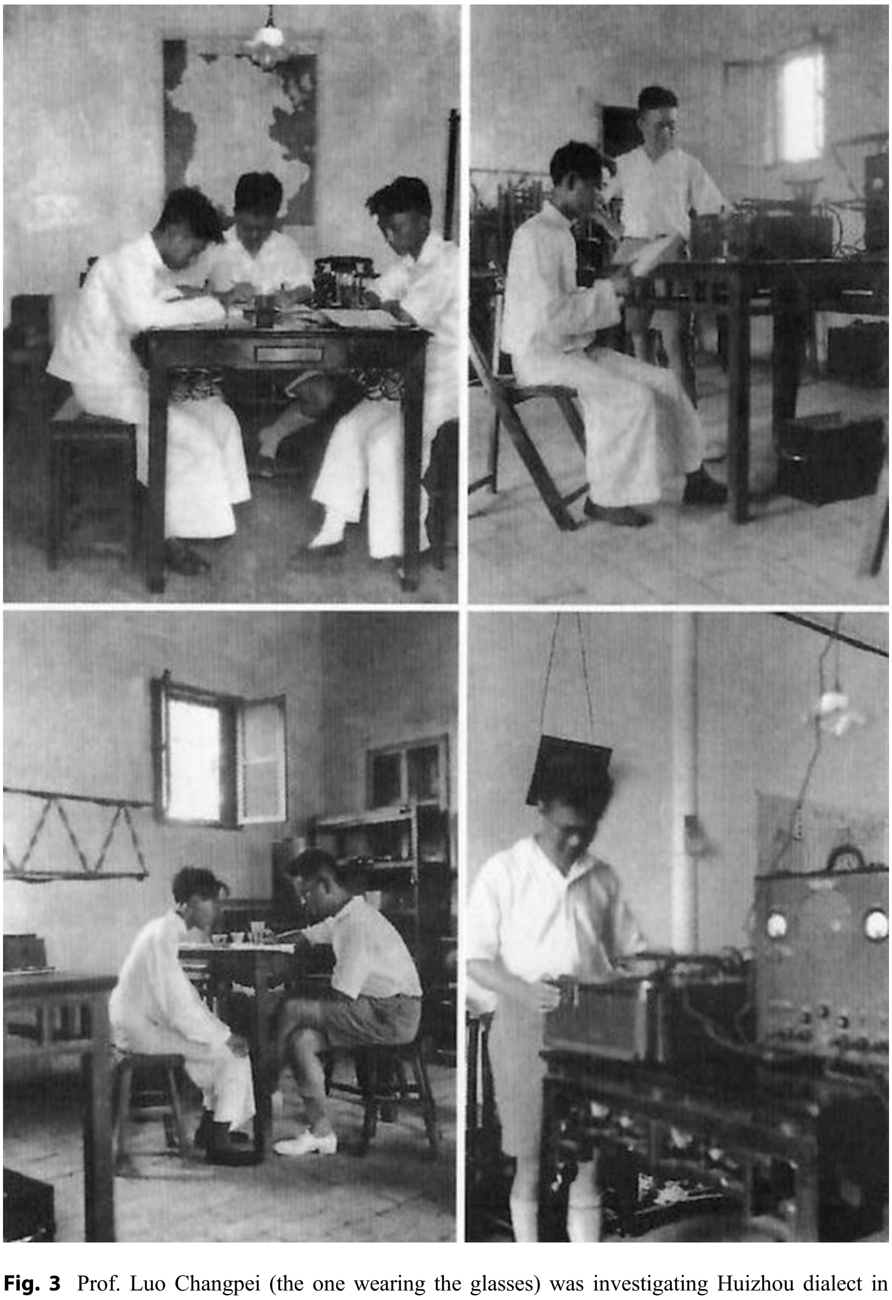

Fig. 3 Prof. Luo Changpei (the one wearing the glasses) was investigating Huizhou dialect in Anhui Province (1932)

Fig. 4 Chao Yuen-ren doing a field work in Southern China (1936)

and invited WU Zongji 吴宗济 as his assistant to carry out phonetic study. Their work was cut short due to the outbreak of the Anti-Japanese War.

Formation and Development of Modern Phonetics

After a period of stagnation due to the war, phonetic research began to recover. The formation and development of modern phonetics as a field of scholarly inquiry can be traced back to the founding of Institute of Linguistics in 1950, which was then part of the Division of Philosophy and Social Sciences at Chinese Academy of Sciences. The development can be divided into four periods: 1949–late 1960s, 1970s–1980s, 1990s–2010, and 2011–present.

Period 1: 1949–Late 1960s

Shortly after the founding of the Chinese Academy of Sciences, at the request of then premier ZHOU Enlai 周恩来, Luo Changpei – then director of Institute of Liberal Arts at Peking University – was charged with the founding of the Institute of Linguistics in June 1950 and served as its first director. Within the institute, a phonetic research group, the predecessor of the phonetics lab, was assembled on the basis of the Phonetics and Music Lab at Peking University. Wu Zongji joined the Institute at the invitation of Director Luo Changpei in 1956, whose arrival set in motion the upcoming experimental phonetic research in the Institute. In its early days, the phonetics lab was far from well-equipped for the development of the field in that the equipment mainly comprised of a kymograph, tone inference scales Yiyi and Yier (Fig. 2), tuning forks, sliding pitch pipe, and wire recorder. The Institute made a decision in 1958 to purchase speech analysis equipment such as frequency meter, oscillograph, and sound spectrograph in order to boost research in acoustic phonetics. Wu Zongji also designed a palatography device to identify which parts of the mouth are used when pronouncing different sounds. The focus of research in this period was on the articulatory and acoustic properties of consonants and vowels in *Putonghua* (Standard Chinese), which was made possible by research tools in acoustics and physiology, and in particular, availability of X-ray photography through collaboration with a local hospital. Major findings include a simple method to calculate vowel formants proposed by Wu Zongji (1964), and static X-ray pronunciation data, palatograms, and tongue position diagrams of phones in Putonghua, all produced by BAO Huaiqiao 鲍怀翘 and Wu Zongji. Figure 5 illustrates the X-ray photos and palatographs acquired by Wu. The five-volume *Experimental Study of Putonghua* and *Putonghua Sound Diagrams* published in 1963, compiled by ZHOU Dianfu 周殿福 and Wu Zongji (1963), were also important accomplishments of this period.

Zhou Dianfu undertook studies in diction. He worked with the famous Chinese crosstalk performer HOU Baolin 侯宝林 and the Peking opera master HAO Shouchen 郝寿臣, and lectured to young performers.

LIN Maocan 林茂灿 specialized in the study of tones in Putonghua. He carried out extensive acoustic experiments on tones in monosyllables and polysyllables using the visible pitch display he designed. His paper "Visible pitch display and the acoustic properties of tones in Putonghua" (Lin 1965) represented an important step forward in uncovering the acoustic properties of tones via visualization of tonal contours.

LIN Tao 林焘 from Peking University paid close attention to the relationship between phonetics and syntax/semantics. The issue that he was concerned with was the relationship between neutral tone and syntactic structures in modern Chinese (Lin 1962).

Period 2: 1970s–1980s

The second period was characterized by a full resumption of phonetic research, followed by remarkable progress made after a hiatus due to the Cultural Revolution (1966–1976). Wu Zongji and Lin Tao provided effective leadership respectively in their phonetics labs. In 1978, Lin Tao reopened the phonetics lab in the Chinese department at Peking University while Wu Zongji assumed the directorship of the phonetics lab in Institute of Linguistics, now affiliated with the newly created Chinese Academy of Social Sciences in 1977.

For the first time, large-scale studies of the phonetic system in Putonghua were underway in the phonetics lab in Institute of Linguistics. Under the leadership of Wu,

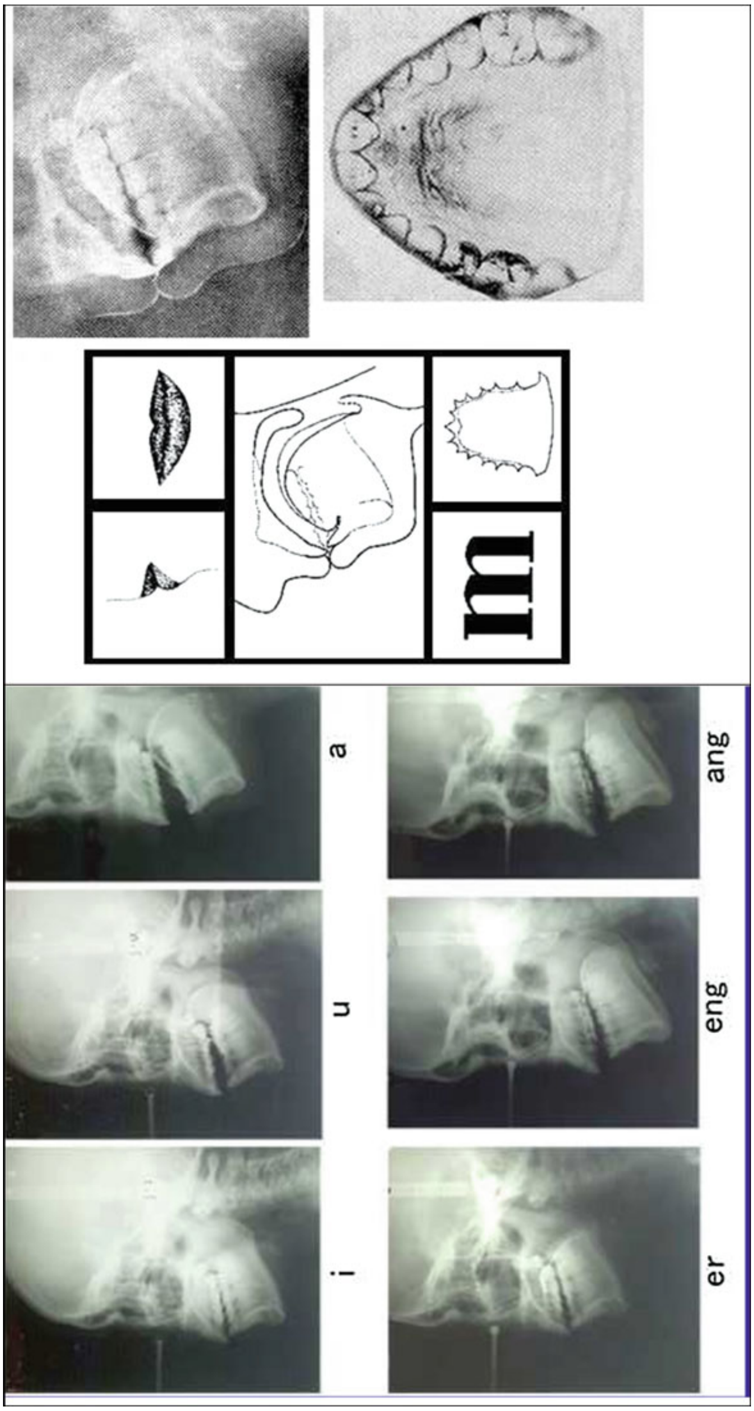

Fig. 5 Examples of X-ray photos of Chinese finals (left) and comprehensive diagram of articulation of [m] (right). (From Zhou and Wu 1963)

researchers in his lab used a sound spectrograph to conduct experiments on monosyllables and disyllables in Putonghua for the analyses of the acoustic properties of subsyllabic components of initials, finals, and tones. Research findings were produced in abundance, which were summarized in a series of "*Notes on experimental phonetics*" published in the core journal *Zhongguo Yuwen* (Chinese Language) over a period of 5 years (Wu et al. 1979). Another major work was *Monosyllable Spectrograms of Chinese Putonghua*, edited by Wu (Wu 1989), which became a valuable reference for study of syllables in Putonghua.

Lin Tao and William S-Y. Wang 王士元 (1984) presented an early study on tone perception, in which by varying frequency and duration values of the second syllable in disyllabic combinations, they examined how listeners perceived the tone in the first syllable.

Rule-based speech synthesis was a highlight in this period. A serious interest in synthesis was developed, thanks to the detailed descriptions of acoustic features of Putonghua. YANG Shun'an 杨顺安 was responsible for development of the first ever rule-based formant synthesizer for Putonghua.

Research began to be reported at international conferences too. In 1979, WU Zongji attended the 9th International Congress of Phonetic Sciences (ICPhS) held in Demark and presented a simple method to calculate vowel formants. He was elected to the ICPhS permanent council at the conference, joining the most esteemed phoneticians of his time.

During this period, extensive work on both acoustics and physiology of speech was carried out. While a linguistic approach was highlighted in the analysis of speech in order to unveil the articulatory and acoustic features of speech in Putonghua, attention was also given to needs of speech engineering for data and knowledge of language.

In addition to Putonghua, experimental method was also utilized in the analysis of Chinese dialects in a very limited way. For example, in the early 1980s, Lin Tao and students from the Chinese department at Peking University conducted a survey on the Beijing dialect spoken in the city of Beijing, its surrounding suburbs, and a few nearby regions. They transcribed and recorded the speech data.

Several important monographs and edited volumes emerged in this period, mostly by researchers in the Institute of Linguistics. *An Outline of Experimental Phonetics*, coedited by Wu Zongji and Lin Maocan, provided a comprehensive introduction of main areas in phonetics and an overview of experimental phonetic study of Putonghua (Wu and Lin 1989). A revision was published with Bao and Lin as coeditors (Bao and Lin 2016). *An Acoustic-Phonetics Oriented Speech Synthesis Technology in Putonghua* by Yang Shun'an (1994), despite of its short length, detailed rules and parameters used in the highly acclaimed formant-based speech synthesizer of Putonghua designed by the author. Unfortunately, the monograph was published posthumously.

Across the city at Beijing University, a volume of experimental work on the Beijing dialect was compiled by Lin Tao and WANG Lijia 王理嘉 (1985). It contained acoustic studies of several important topics in the Beijing dialect, namely, neutral tone, r-suffix, tone and intonation, and duration distribution in diphthongs, and is still an important reference in experimental phonetics.

Period 3: 1990s–2010

Phonetic research continued to thrive in the 1990s. One significant shift was evident in its increasingly closer tie with speech engineering. The object of inquiry was moving from isolated speech units to naturally occurring continuous utterances, and from segments alone to the inclusion of suprasegmentals. As a result, much effort was put into study of coarticulation, sound change, stress, rhyme, intonation, and prosodic structure. As computers became more powerful and advanced tools of speech analysis more available, phonetic sciences were developing at an accelerated pace into the twenty-first century.

Period 4: 2011–Present

The recurring theme in this period was innovation and solving previously unsolvable problems. More phonetic research projects were funded by the national program on key basic research, with new development in interdisciplinary work. At the same time, the scope of phonetic research continued to expand as interdisciplinary research in phonetics began to gain momentum. Thanks to the strong support from the national research and talent programs, new phonetics labs were founded in research institutes and universities, with quantity and quality of research immensely enhanced. All the above laid a solid foundation for more interdisciplinary and cross-disciplinary work in the future.

Specifically, different stages in speech communication as schematically represented in the speech chain were being intensively investigated, from speech production and articulation mechanism to speech perception and speech and cognitive development. Other areas of research included speech pathology, phonetic study of Chinese dialects, and development for speech and language research platforms. Equipment and tools used in research were more diversified. In addition to tools for acoustics analysis, equipment was also available for the psychological and physiological analyses, such as electroencephalography (ERP), eye-tracking, electromagnetic articulography (EMA), NDI wave speech research system, high-speed glottal photography, and ultrasound system.

In sum, the past seven decades from 1950 to the present have witnessed the development and thriving of modern phonetic research in China. Despite different priorities in the four periods, phonetic research in China has been moving at an accelerating pace largely thanks to the establishment of key research institutes, strong leadership of senior scholars, funding support from the government, and increasing collaborations with other disciplines.

Overview of Phonetic Research in China

Phonetic research in China has covered all three aspects of phonetics, namely, articulatory, acoustic and auditory. Both segmental and suprasegmental properties have been investigated from these three perspectives. Tones, in particular, have been extensively studied. Earlier researches are mostly about Putonghua, the standard variety of Mandarin, while studies of other Chinese dialects and ethnic minority languages in China are gaining momentum lately. This section offers an overview of phonetic research in China in three parts, concentrating on speech production, acoustics of speech, and speech perception, respectively.

Speech Production Mechanism

Research in the kinematics of speech production focused on two processes, phonation and articulation. Phonation is the process by which airflow from the lung causes vibration of the vocal folds to generate the quasi-periodic glottal waveform. Articulation is the process by which the glottal waveform is modulated by the response characteristics of the vocal tract, resulting in different speech sounds.

The glottal vibration can be described by characteristics in the time and frequency domains. Characteristics in the time domain reflect the speed at which the vocal folds vibrate, corresponding acoustically to the fundamental frequency (F0). Characteristics in the frequency domain reflects the manner in which the vocal folds vibrate and can be described by two parameters, open quotient (OQ) and speed quotient (SQ). BAO Huaiqiao 鲍怀翘 is the first researcher to show a serious interest in the phonation types. He and ZHOU Zhizhi 周植志 conducted the first study of the phonation types of consonants in the Wa language in 1990 (Bao and Zhou 1990). Following Bao, KONG Jiangping 孔江平 carried out extensive study on phonation (Kong 2003). He identified the most common phonation types and offered detailed acoustic analyses of seven different types. Like Bao, he also analyzed the phonation types of consonants and vowels (and their acoustic properties) in different ethnic minority languages, including Hani, Miao, Liangshan Yi, Axi Yi, Zaiwa, and Jingpo (Kong 2001). ZHU Xiaonong 朱晓农 focused more on the linguistic functions that phonation types could be recruited to perform by pointing out that phonation types can be contrastive in defining "pitch register" in phonology (Zhu 2009).

The process of articulation is concerned with the vocal tract resonance. There are three main areas of research.

The first area is the study in the kinematics of articulation. Zhou Dianfu and Wu Zongji worked with X-ray photography and palatography devices to record the shape of the vocal tract, contact of the tongue with the palate, and the shape of the lips for each of the 22 consonants, 10 simple vowels, and 2 nasals (used as syllable endings) in Putonghua (Fig. 6). The results were presented in *The Atlas of Sound Spectrograms in Putonghua* (Zhou and Wu 1963). Two decades later, the publication of *Characteristics of Articulatory Gestures in Putonghua* (X-ray videotapes) by Bao Huaiqiao and YANG Lili 杨力立 (1985) represented an important step forward, in

Fig. 6 Professor WU Zongji was operating his palatography designed in 1950s

its comparison with Zhou and Wu (1963), in that it contains the kinematic trajectories of more complex articulatory configurations in monosyllables, disyllables, and syllables with r-suffix. Compilation of the articulatory data received a major boost when the first dynamic electropalatography (EPG) database – containing dynamic EPG data collected in the production of monosyllables, disyllables, syllables in neutral tone, syllables with r-suffix, ancient Chinese poems, sentences, and dialogues – was created at Institute of Ethnology and Anthropology (Bao and Zheng 2011). Soon after, similar databases were set up at Institute of Linguistics, Peking University and Fudan University.

In recent years, EMA has been used more often than before. For example, HU Fang 胡方 from Institute of Linguistics at Chinese Academy of Social Sciences was able to study with EMA the articulatory characteristics and tongue movement patterns of vowels in the Ningbo dialect (Hu 2014). LI Aijun 李爱军 used EMA to investigate the vowel articulatory spaces of Chinese and Japanese speakers in their speech under different emotions by examining the patterns of tongue movement (Li 2015).

The second area is the study of coarticulation, which refers to the effect of articulatory overlapping of neighboring segments. XU Yi 许毅 distinguished four levels of segmental junctures in Putonghua in his description of how closely neighboring segments are related: close juncture between segments within a syllable, syllabic juncture between syllables, rhythmic juncture between rhythmic units and pausal juncture in and between utterances (Xu 1986). Research in coarticulation started at the segmental level first. Wu Zongji and Sun Guohua 孙国华 (1990) was the earliest study of segmental coordination and they examined the coarticulatory effects of unaspirated stops and fricatives within and between syllables in disyllabic words (C1V1#C2V2). Lin Maocan and YAN Jingzhu 颜景助 (1994) followed up with an experiment on the coarticulation between vowels and nasals in zero-initial syllables with nasal endings (i.e., with the syllable structure VN). Development of the EPG technique and creation of the dynamic EPG database have made available a wealth of EPG data for the study of coordination. For example, consonant production in different segmental contexts was examined on the basis of EPG data by Bao Huaiqiao and ZHENG Yuling 郑玉玲 (2011).

Most studies on coarticulation zeroed in on the articulatory overlapping and acoustic consequences of neighboring segments in the production of consonants and vowels in many languages. Chinese researchers turned their attention to the coarticulation of suprasegmental features such as tones. Chao discussed interaction of tones in his most influential book *A Grammar of Spoken Chinese* (Chao 1968/1979). He noted that in the combination of two high falling tones in Mandarin, the preceding falling tone did not fall to a lower pitch level whereas the following falling tone did not start at a higher pitch level due to the influence of the high pitch onset in the second tone on the low pitch offset in the first tone, and vice versa. Wu Zongji described the basic pitch contour patterns for the 16 disyllabic tonal combinations out of the four lexical tones in Putonghua, and based on the disyllabic patterns, he further proposed tone sandhi rules for the trisyllabic and quadrisyllabic units (Wu 1982). More studies on tonal coordination followed. In this line of research, it has been found that anticipatory and carryover effects are both present but differ in magnitude and in nature in that the former are relatively small and mostly dissimilatory while the latter are large and mostly assimilatory (Xu 1997).

The third area is modeling of the speech production. In terms of articulatory models, the finite element method was employed to simulate coordinated activities of articulators and muscles based on the MRI data and model the human speech production process (Dang and Honda 2004).

Acoustic Phonetic Study

Unlike Indo-European languages, Chinese language in general uses pitch variations to distinguish lexical meanings, resulting in complex interactions of tones. In addition to extensive research on tones, acoustic phonetic studies were also carried out on consonants and vowels in a wide variety of Chinese dialects, which revealed some very interesting facts. An overview on the acoustic phonetic studies of vowels, consonants, tones (including neutral tone), and intonation is offered in this section.

The study of vowels mainly focused on two areas. The first area was concerned with the analysis of acoustic parameters of vowels, including formant frequencies, formant movement trajectories, and duration. As early as the 1960s, Wu Zongji proposed formulas to calculate vowel formant frequencies, bandwidths, and amplitudes and used them in his calculations of formants for the 10 simple vowels in three groups of speakers – male, female adults, and children – in Putonghua (Wu 1964). He plotted the vowel diagram based on the calculated vowel formants. In addition to monophthongs, complex vowels (diphthongs and triphthongs) were investigated at length on their dynamic formant movement, properties in time and frequency domains, and spectral features (Yang and Cao 1984). While vigorous studies on Putonghua continue, interest in the acoustic analysis of vowels in Chinese dialects has grown significantly lately. Examples include devoicing of vowels in the

Shanghai dialect (Xu 1990), and the acoustic analysis of diphthongs in the Ningbo dialects (Hu 2014).

The second area involved the normalization of formants. Normalization was necessary in acoustic-phonetic analysis of speech to account for between-speaker and cross-dialectal variations in order to derive meaningful linguistic information from the acoustic signals. For example, LING Feng 凌锋 suggested that the normalization method used by Gunnar Fant worked better for gender-induced variations and also proposed a modified normalization method based on standard score to smooth out cross-dialectal variations (Ling 2008).

Consonants in Chinese have also been studied from different perspectives. In a series of early studies of Putonghua, Wu Zongji analyzed formants of voiced consonants, concentration areas of acoustic energy in voiceless consonants, and acoustic properties of zero-initials in onsetless syllables; he also conducted studies on the aspirated and unaspirated obstruents in terms of their differences in airflow, air pressure, and acoustic properties such as voice onset time (VOT), sound intensity, and spectrums (Wu 1992). Marilyn Chen investigated acoustic properties of nasal endings ([n]and [ŋ]) in Putonghua in search of acoustic cues of nasalized vowels and nasals (V-N) in the acoustic signals (Chen 2000). Chen proposed the A1-P1 and A1-P0 parameters as a measurement of nasalization, while Fang (2004) exploited the zero-pole model. Approaches other than acoustic analysis has also emerged in the study of consonants. Hu Fang analyzed EPG, EMA, and acoustic data in his detailed study of the three sibilants (both aspirated and unaspirated) in Putonghua (Hu 2008).

Consonants in the Wu dialects have received a special attention, due to the three-way contrast in stops shared by many Wu dialects. Cao and Maddieson (1992) conducted acoustic and aerodynamic analyses of voicing in four Wu dialects and observed that voicing in Wu dialects is not cued by the vibration of vocal folds during the closure period of the voiced stops; instead, it was manifested in the breathy voice of the following vowel.

Chinese is well-known as a tone language. The four tonal categories were referred to as "Ping," "Shang," "Qu," and "Ru" in Middle Chinese. The nature of the four tones remained murky and oftentimes mysterious – whose descriptions were solely based on perceptual impressions of high, low, long, and short – until the 1920s when experimental methods were introduced in the measurement of tones in Chinese by Liu Fu, Wang Li, and Yuen-Ren Chao. In "*An Experimental Study of Chinese Tones*," Liu Fu measured the tones in 12 Chinese dialects using kymograph and Tone Reference Scale. In 1931, he created "*A Chart of Phonetic Symbols for Investigating Chinese Dialects*." BAI Dizhou 白涤洲, the student of Liu Fu, undertook investigations of dialects in the Guanzhong regions and offered a systematic analysis of the acoustic properties of tones and the evolution of the "Ru" tone in a series of articles, including "An Experimental Study of Tone in Guangzhong." Yuen-Ren Chao proposed the "tone letters" used in IPA for the transcription of tones, in which the normal pitch range is divided into four equal parts, marked by five pitch levels, numbered 1, 2, 3, 4, 5, with 1 being the lowest and 5 the highest (Chao 1930/2002).

There are abundant studies of tones both in Mandarin Chinese and other Chinese dialects. Some researchers focused on the analysis of acoustic features such as F0, duration and amplitude of tones in isolation and in a sequence, as in Lin and Yan (1992). An interesting development in recent years is the study of phonation types as a cue of tonal contrast. Evidence was presented to show that tones can also be distinguished using phonation types, against the previous assumption that pitch is the sole cue of tonal contrast for tone languages like Mandarin Chinese (Kong 2001). Another area was more methodology-oriented with a focus on the normalization of F0. Various normalization methods were proposed, as reviewed in detail by Zhu Xiaonong. He made a comprehensive comparison of six methods using data from speakers of the Shanghai dialect and the logarithmic z-score transform has the best performance in terms of normalization index and dispersion coefficient (Zhu 2004).

In Putonghua, a special phenomenon called the neutral tone has got ample attention in two aspects. The first aspect involves the status of neutral tone in the Mandarin phonological system, to wit: whether the neutral tone is a tonal category or part of the metrical system. A syllable is said to be in neutral tone when it does not carry one of the four lexical tones in Mandarin. According to Chao (1979), neutral tone is related to weak stress. When a syllable is in weak stress, its tonal range is almost reduced to zero and its duration significantly shortened. Lu and Wang (2005) distinguished “neutral tone” from “weak stress,” reserving the former as neutralized tone in the tonal system and the latter as unstressed syllable in the metrical system. The second aspect involves studies on the acoustic properties of syllables of neutral tone. In particular, Li considered the effects of prosodic boundary and information structure on the phonetic realization of neutral tone. Pitch and durational properties of neutral tone in different prosodic contexts were examined (Li 2017).

Since both tone and intonation exploit pitch as its primary acoustic cue, the interaction of tone and intonation in Chinese is an intriguing issue in intonational phonology. Since the seminal work of Chao (1932/2002, 1933/2002), a rich body of research, both descriptive and experimental, has been produced to advance the understanding of the linguistic functions and physical properties of tone and intonation in Mandarin Chinese, especially regarding the interaction of tone and intonation. Chao came up with two well-known metaphors in characterizing the ways tone and intonation interact: the “rubber band effect” and the “small ripple and big wave” theory. According to the latter, tone and intonation are related in the form of superimposition – either successive or consecutive – just like small waves sitting on top of big waves (Chao 1933/2002: 198–220).

Wu Zongji, who inherited and expanded Chao’s theory, proposed the “transposition model” of intonation, which accounts for obligatory and optional tone sandhi patterns in Chinese (Wu 1982). In his model, an intonational phrase is composed of one or more tone groups at the phrase level and an additional boundary tone. Between tone groups, the pitch range (in semitone) does not fluctuate much, and within each tone group, the interaction of tones follows relatively fixed patterns. This is the obligatory tone sandhi of the model. Paralinguistic factors such as focus and emotions could cause the pitch range of a tone group to shift to a new pitch key, which is the optional tone sandhi of the model. SHEN Jiong 沈炯 characterized

Mandarin intonation in terms of the upper line and lower line of F0 that define a pitch register and argued that the two lines can be manipulated independently of each other in different intonational patterns (Shen 1992). Xu (2005) proposes the Parallel Encoding and Target Approximation (PENTA) model of speech prosody, which is a framework for conceptually and computationally linking communicative meanings to fine-grained prosodic details, based on an articulatory-functional view of speech. In a monograph, *The experimental study of tone and intonation in Chinese* (Lin 2012), Lin took Chao's insights as a point of departure and explicitly adopted the autosegmental-metrical (AM) model of intonation. In his model, focal prominence and boundary tone are the two key elements in describing intonation in Chinese. For example, the difference between declarative intonation in statements and interrogative intonation in questions without sentence-final question particles resides in the boundary tone, which is realized acoustically as variations in pitch register and slope of the contour. SHI Feng 石锋 looked at intonation from a broader perspective and proposed a systematic method to define an "intonation pattern" with three parameters – F0 contours, pause-lengthening ratio, and sound intensity (Shi 2013). In his study of declarative and interrogative intonations in Putonghua, Cantonese, and Korean, Shi tried to figure out cross-linguistic patterns in intonation in terms of the quantitative measurements of F0, duration, and intensity.

Focus is probably the topic in the study of intonation that has received the most attention. Xu (1999) looks into effects of tone and focus on the alignment of F0 contours and found that focus exerts influence on the pitch range of different components in different ways. The pitch range of the syllables before the focal position remains unmodulated, and pitch range is dramatically expanded for syllables in the focal position and compressed for syllables after focus. He termed the last phenomenon "post-focus compression" (PFC). Other studies by Jia, Li, and Chen (2009) analyzed phonetic realizations of different types of focus and situations in which there is one, two, or multiple foci in the utterance. In CAO Wen 曹文's production and perception study, he approached the prosodic realization of focal accent in short utterances in Chinese by adopting a perception test using synthetic materials (Cao 2010). He concluded that the most crucial cue in conveying focus is the extent of changes in F0 maxima, with duration coming second.

In addition to focus, pause (or break before a prosodic boundary) was another closely examined area in the study of intonation. Main areas of research included acoustic realizations of pause, such as changes of F0 before and after the pause (Shen 1992), the duration of syllable rhyme before the pause and the duration of syllable initial after pause (Wang, Yang and Lü 2004) and also prosodic transcription of pause (i.e., break indices) (Li 2002).

The experimental studies have flourished on the influence of emotions on intonation patters in recent years. Li undertook an extensive study on the role of intonation in conveying emotion in a tone language like Chinese, with focus on F0 levels and pitch contours in what she termed "successive addition boundary tone" (Li 2015). She proposed that the boundary tone is composed of two components – the base tone of the syllable and an addition contour – in expressing basic emotions such as happiness and anger. She also offered a phonological representation with phonetic descriptions.

Step further into higher prosodic levels, intention understanding and generation in human-machine interactions call for greater integration of discourse-level prosodic information in spoken dialogue understanding systems. In a series of studies on the interface of prosody and discourse, Li, Jia, and their collaborators conducted detailed analyses on prosodic features in connection with discourse structure, information structure, and speech acts, which is still an ongoing enterprise.

Speech Perception

One important area of research in speech perception is the mode of speech perception, whether it is categorical or continuous. Categorical perception refers to the phenomenon that speech stimuli along a continuum are perceived as being discrete linguistic categories. Continuous perception is the opposite. Early studies focused on consonants and vowels at the segmental level. The perception of vowels was generally considered to be continuous, while consonants in Chinese are categorical in perception, as indicated in a study by XI Jie 席洁, JIANG Wei 姜薇, ZHANG Linjun 张林军, and SHU Hua 舒华 (2009).

Perception of Chinese tones was approached from two perspectives: modes of perception and perceptual cues. William S-Y. Wang is the first to demonstrate that tone perception in Putonghua was categorical. In Wang (1976), he investigated the perception of tones in Putonghua by manipulating stimuli along a continuum of Yinping Tone (Tone 1, high-level tone) and Yangping (Tone 2, rising tone) on the syllable /i/ in identification and discrimination tasks. Results showed crossovers on the identification curve, which corresponds to the perceptual boundary of the two tones, and prominent peaks on the discrimination curve. These are general characteristics of categorical perception. WANG Yunjia 王韫佳 and QIN Xihang 覃夕航 hypothesized that tones having distinct pitch contours exhibit strong tendency for categorical perception while the clear perceptual boundary does not exist for tones with similar pitch contours (Wang and Qin 2015). Generally speaking, research so far has established that the perception of Yinping and Yangping tones and that of Yinping and Qusheng (Tone 4, falling tone) tones are categorical, but the perception of Yangping and Shangsheng tones is complicated due to the similarity of surface pitch contours. Categorical perception only occurs with availability of specific acoustic cues.

Research on perceptual cues on tonal perception has discovered that fundamental frequency (F0) is the key to the perception of tones. In addition to F0, other acoustic properties such as syllable duration and phonation types (Kong 2001, 2003) also play a role. Contextual factors also contribute to tonal perception. In "*Issues in Tonal Perception*," Lin and Wang's (1984) study addresses these factors. For example, in a disyllablic word the F0 onset of the second tone could influence the listener's judgment of the first tone. The higher the F0 onset of the second tone, the more likely the first tone is to be heard as Yangping (Tone 2, rising tone), or Shangsheng (Tone 3, low dipping tone) if the first syllable itself is with a longer duration. Linguistic background of the subject participating in the listening tasks has also been proved to be an important factor in tonal perception in Putonghua (Peng et al. 2010).

The third tone (Shangsheng, low dipping tone) in Putonghua is considered a special category largely because of its contour shape of a low falling followed by a rise, traditionally described as “214” in Chao’s system. As a result, plentiful studies were done to evaluate effects of the onset, turning point, and offset of the pitch contour on the perception of the third tone. The falling and rising components of the pitch contour have also been investigated. For example, Lin started out with a claim that the onset and offset of the pitch contours did not provide crucial information on the perception of the third tones (Lin 1965). In Shen and Lin (1991), they conducted a perceptual study of Tone 2 and Tone 3 in Putonghua, the two tones said to be most confusable among the four lexical tones due to their similar concave shapes. Their results showed that the distinction between Tone 2 and Tone 3 was cued by the timing of the turning point, which is correlated with the degree of the initial falling F0. In addition to perceptual cues, the third tone sandhi and its underlying phonological form were also actively studied. For example, Cao (1995) discusses various surface patterns after application of the third tone sandhi and its phonological, semantic, and syntactic constraints. In terms of its underlying phonological form, there is strong consensus that the third tone, despite its concave shape in pitch contour, is a low tone. Disagreement centers on whether it was a level tone. Cao (2010) points out, based on results of a perception test of level tones, that one should be cautious to describe the third tone as a “low level” tone.

With the developments in cognitive neuroscience, new experimental methods were adopted to explore the cognitive and neural mechanisms underneath speech perception. In terms of language processing, studies of event-related potential (ERP) revealed that in speech perception segmental (i.e., consonants and vowels) and suprasegmental (e.g., tones) are handled by different processing mechanisms (Li et al. 2010), activating different regions in the brain. Most of the processing of the segmental information happens in the left hemisphere, and no such strong left-laterization has been found for the processing of the tonal information (Liu et al. 2006). Adoption of ERP and fMRI in studies of tonal perception made it possible to examine neural responses to within-category and across-category stimuli. Experimental results showed that within-category deviants and across-category deviants both triggered mismatch negativity (MMN), but the former induced larger electrophysiological responses (Xi et al. 2010). In a recent study on the categorical processing of Chinese lexical tone, Si et al. (2017) conducted cortical surface recordings in surgical patients and revealed a cooperative cortical network along with its dynamics responsible for this categorical perception. Based on an oddball paradigm, they found amplified neural dissimilarity for cross-category tone pairs, over cortical sites covering both the ventral and dorsal streams of speech processing, but no such effect has been found for within category tone pairs.

Phonetic research of the last half century has made significant progress in almost all aspects of speech events that constitute the Speech Chain, from speech production to the acoustics of speech and speech perception. Topics such as tones and intonation in Chinese, perception of Chinese lexical tones, and phonation types have been vigorously pursued. While extensive research has been done on Putonghua or the Mandarin variety of Chinese, other Chinese dialects and minority languages spoken in China were also investigated using experimental methods of phonetics. While designed speech data are still dominant in current studies, increasing effort has been made to exploit naturally occurring spoken dialogues to explore topics such as utterance- and discourse-level prosody.

Interdisciplinary Research in Phonetics

Apart from research of phonetics in itself, phonetics has attracted attention from various disciplines. The interdisciplinary research of phonetics and the application of phonetics have been onverarching recently. The wide-ranging applications of phonetic research have been well attested in a variety of interdisciplinary fields such as speech pathology, forensics, language acquisition, language pedagogy, and artificial intelligence. The natural maturation of linguistic ability and normal speech mechanisms are fundamental to the well-being of human beings. The curiosity into how children acquire their mother tongues leads to the study of language acquisition, while the diseases and physical disorders faced by many patients urge the researchers to study speech pathology. Phonetics can also be employed to assist with forensic evidence. This section covers developments in these areas in China.

Speech Pathology

Pathological speech usually refers to the distorted speech resulting from abnormalcies in voice or in the articulatory mechanisms due to diseases or other physical or biological disorder of the speech production system. Speech pathology originated in the medical field. Content wise, early studies mainly focused on three areas. The first area was concerned with the characteristics of pathological speech. For example, WANG Jianhua 王建华 et al. analyzed characteristics of the Putonghua speech sounds produced by cleft palate patients based on transcriptions of their pronunciations of syllable initials and finals in Putonghua (Wang et al. 2003). These studies provided theoretical and clinical support for speech assessment and therapy for patients with cleft palates. The second area was the assessment and detection of pathological speech. *The Word List for Cleft Palate Speech Intelligibility*, developed for assessing speech disorder of patients with cleft palates, contains minimal pairs of speech sounds and serves as an effective tool for qualitative and quantitative evaluations of speech intelligibility (Jiang et al. 2009). The third area was related to treatment and training, for example, treatment of patients with cleft palates (Wang 2013).

In terms of research methods and techniques, spectrograph was used in the early analysis of pathological speech, but not on a large scale. More frequently used methods of diagnosis included stroboscopic measurement, laryngoscopy, endoscopy, and laryngeal electromyogram to locate abnormalcies for treatment. These procedures have some shortcomings. Most of them are invasive and could cause discomfort. Their execution is dependent upon the patients' health conditions and their willingness of cooperation. Diagnosis results are based on doctor's subjective judgments and experience, which can sometimes be unreliable. Early detection is not likely with these methods, which could delay treatment. As speech pathology became more interdisciplinary, methods and tools in phonetic research were gradually adopted. One approach was to track changes in perceptual judgments by modifying acoustic parameters of the pathological speech using a voice synthesizer. Parametric modifications such as slight reduction of F0 and amplitude, and standardized glottal noise energy turned out to be effective in evaluating voice quality, due to its multidimensional property. An attempt in this direction was made by Huang and Wan (2008), which showed that perceptual judgments of voice quality and the acoustic analysis of speech worked together to yield the most reliable evaluations of voice quality. Based on theories of acoustics and anatomy, early diagnosis could be made by adopting computer technology and modern digital signal processing methods to conduct time-frequency analysis and multi-parameter pattern classification on voice data (Peng 2008). Analysis of the acoustic properties of some vowels and consonants in the pathological speech (atypical speech sounds) allowed patients with congenital velopharyngeal Insufficiency (VPI) to be diagnosed and treated (Zhou et al. 2012).

As researchers learned more about phonetic features, studies on the feature-based identification of pathological speech started to emerge, as exemplified in the identification of vocal cord nodules and polyps and normal speech using support-vector machines (Yuan et al. 2015). Developments in pattern recognition technologies made it possible for nondestructive testing methods to be used in the automatic evaluation of pathological speech disorders so that testing results can be more objective. Pathological speech recognition is also replacing traditionally used subjective evaluation methods and it has become a hot research topic.

Along the same line, studies of pathological speech are also gradually expanding their subject populations from adults to children. Hearing impairment is the focus of pathological speech research in children. In terms of research methods, most studies are comparing the speech output of normal hearing (NH) children of Putonghua with speech development of the hearing impaired children who are wearing hearing aids or have received cochlear implants (CI). Content wise, the majority of research was concerned with speech output. At the segmental level, Yang and Xu (2017) conducted the analysis of formant movement patterns for vowels produced by cochlear implant children. Other studies looked into the speech forming ability of children with hearing loss in different age groups, their speech intelligibility and the development of mouth and tongue movements. At the suprasegmental level, some research showed that currently artificial cochleae are not efficient in decoding pitch information and identifying Chinese tones. Most studies found that children with hearing loss produced the Yingping tone (Tone 1) with the best accuracy, demonstrating good speech forming ability. It is also reported that the most difficult tone for the hearing impaired children is Yangping (Tone 2) (Han et al. 2007).

Another actively researched area is the relationship between age of the hearing impaired children and their speech development. Most studies in this area found that the speech development is better in hearing impaired children who received artificial cochlear implantation before age 3 or 4 than after. Han et al. (2007) investigated the tone production performance of native Mandarin Chinese-speaking children with cochlear implants and evaluated the effects of age at implantation and duration of implant use on tone production in those children. Their results showed that an increased duration of implant use might facilitate tone production, but the age of implantation appears to have a negative effect on tone production in cochlear implant children (the later children receive cochlear plantation, the poorer the performance). Therefore, early implantation might be beneficial to tone production in prelingually deaf children whose native language is a tone language.

Forensic Phonetics

Forensic phonetics is an interdisciplinary research discipline that aims to analyze and compare speech characteristics for the purposes of speaker identification, decoding spoken messages, analysis of emotions in voice, authentication of recordings, and related. In practice, theories and methods in the field of forensic phonetics are often known as voiceprint identification technology. Despite its brief history in China, its development has been fast.

In 1988, Yue Junfa from the Documents Inspection Department at Criminal Investigation Police University acquired the Kay Elemetrics 7800 Sonagraph and started research and teaching in forensic phonetics. He is regarded as the founder of voiceprint identification in China. A year later, the Center of the Material Evidence Identification affiliated with the Ministry of Public Security, known as the Second Institute of the Ministry of Public Security at the time, also commenced forensic phonetic research after acquiring a newer model of the Kay Elemetrics 5500 Sonagraph. In 1992, the Center completed a key research project of the Ministry of Public Security called *Application of the 5500 Sonagraph in Voiceprint Identification*. The project recruited 60 speakers (30 males and 30 females) who shared a lot of similar characteristics such as age, educational background, birthplace, and long-term residence. Researchers examined their pronunciations of Putonghua in normal speaking condition and performed a statistical analysis of the acoustic speech data in order to establish the methods and procedures for voiceprint identification in normal speaking condition. The accuracy rate of identification reached 100%. This is the beginning of the standardized forensic phonetic research in China. Since then, the voiceprint identification technology has started to be employed in more than 70 court labs throughout the country.

In 1996, the Center of the Material Evidence Identification started to undertake studies of atypical speech and other variations in voiceprint identification as part of the research work carried out for the project *Voiceprint Identification Key Technology and Speaker Recognition System*, funded by the 9th Five Years Key Programs for Science and Technology Development of China. In order to improve the methods and procedures used in voiceprint identification, researchers at the Center analyzed the disguised voice and other speech variations caused by the cold, different recording devices, dialect accents, whispering, imitating, and impersonating. A proprietary speech workstation was also designed as part of the above-mentioned project. In addition to analyzing acoustic speech data, the workstation can also perform noise reduction for the noisy speech and digital amplification for the weak speech signal. It outperforms similar products abroad in terms of functions and practicality. The successful deployment of the proprietary speech workstation was seen as one of the most important advancements in forensic phonetics in China.

In 2006, the Center of the Material Evidence Identification with its voiceprint identification lab was accredited by the China National Accreditation Service for Conformity Assessment (CNAS), which is seen as an important milestone in the standardization of methods and procedures in voiceprint identification technology. The publication of one national standard and four industry standards in 2017 proved to be another milestone. Recently, the increasing integration of forensic phonetics and artificial intelligence technology, supported by the Intelligent Speech Technology Key Lab of the Ministry of Public Security, is going to play a key role in the deployment of "Smart Courtrooms."

Speech Acquisition and Cognitive Development in Infants and Young Children

Compared to other areas of phonetic research in China, studies of speech acquisition and cognitive development in Chinese-speaking infants and young children did not appear until recently, with early studies focusing on production research of individuals for the purpose of understanding the development sequence of pronunciation and error types. LI Yuming 李宇明, who conducted a research on speech acquisition of infants of 1 to 4 months old, is one of the earliest studies of children language acquisition in China (Li 1991). LI Wei 李嵬 and others analyzed the acquisition sequence of phonemes in Putonghua and error types in 129 children of 1.5–4.5 years old in a comprehensive study of children's speech acquisition (Li et al. 2000). In recent years, research teams in China have adopted more advanced methodologies and produced substantial results on the children's speech acquisition and cognitive development based on large-scale corpora.

Research on infant and young children's speech included analysis of error rates and types of vowels, consonants, and tones produced by children in different age groups based on big data and speech corpora. The children's speech research team, led by Li Aijun and Gao Jun GAO Jun 高军, at Institute of Linguistics, Chinese Academy of Social Sciences, has been collecting large amount of children's speech data since 2008. They analyzed pronunciation data in picture naming tasks of more than 4000 children of 1.5–6 years with normal hearing in the Beijing area and established pronunciation testing standards and norms for Putonghua-speaking children with normal hearing. Gao and Shi (2019) conducted acoustic analysis of the speech output in the children's speech database they created. They found that in terms of error rates, most errors happened in the initials, some in the rhymes and a few in the tones. In terms of the acquisition sequence, acquisition of the tones seemed to occur early and that of the initials late. They also found that acquisition of the single tones and that of the first tone in a disyllabic (i.e., bitonal) sequence followed different acquisition mechanisms and that by the age of 2, children have already acquired the phonological rule of neutral tone.

Studies of speech perception in infants and young children have been lacking, especially of their ability to discriminate tones. The children's speech research team at Institute of Linguistics has started to undertake studies of tonal perception in infants and toddlers since 2008 and published a series of research reports (Shi et al. 2017). Among them, Cao, Li, and Fang developed a language acquisition model based on the interconnected network model structure and simulated the acquisition mechanisms in the process of establishing phonemic categories in Putonghua by infants and toddlers (Cao et al. 2017). The published results on speech perception so far have raised doubts about whether "perceptual reorganization" happened in the early stages of infants' speech development. Perceptual reorganization refers to the phenomenon that infants are able to discriminate native speech sounds from non-native ones, but after 6 months their ability to perceive non-native speech contrasts starts to decline or disappear and their ability to perceive native speech contrasts continues to grow and strengthen. While some studies argued for the perceptual reorganization hypothesis, citing that reorganization has been found for lexical tones between 6 and 9 months of age (Mattock and Burnham 2006), others failed to reduplicate similar reorganization effects (Shi et al. 2017). Yet other studies contended that infants do not possess the innate ability to discriminate phonetic categories in their native language, especially when the two categories are acoustically similar and would require long time to acquire the contrast (Fan et al. 2018).

Currently there have been few studies on speech perception and development in toddlers and preschool children. Some research examined perceived accuracy of monosyllabic Mandarin tones produced by 3- to 5-year-old children growing up in Taiwan and in the United States and found that none of the four tones produced were adultlike. In terms of the accuracy of the tones, more errors are made in Third Tone than other tones, which appeared to follow the order of articulatory complexity in producing the tones (Wong 2013). In other studies (e.g., Xi et al. 2009), 6- to 7-year-old children were shown to have exhibited adultlike behavior in perceiving Mandarin tones categorically.

One interesting phenomenon in children's speech perception is "perceptual asymmetry," which refers to a situation in which the presentation order of stimuli could impact infants' responses in experiments, to wit, the subject performs better in one order than the opposite order. Perceptual asymmetry could happen in the perception of consonants and vowels and that of suprasegmental features such as tone. In experiments of discriminating Tone 1 and Tone 3 in Mandarin, infants demonstrated better discriminatory ability when trained on Tone 1 before Tone

3. The mechanism behind the perceptual asymmetry remains unclear. Speculatively, it might be related to the statistical distribution of the input. For example, when infants are learning a low-frequency tone category, they will establish a phonological feature for that "atypical" token to encode the contrast between low-frequency and high-frequency categories. However, when the infants had learned the high-frequency tone category first, they would likely have treated the low-frequency tone category they learned later as an "atypical" exception of the high-frequency tone category, resulting in nondiscrimination of the two tone categories. Therefore, the statistical distribution of the phonological features in the input speech could impact the infants' speech perception ability. Some research has demonstrated that infants are very sensitive to the statistical distribution of the input speech, suggesting that they have access to a powerful mechanism for the computation of statistical properties of the language input.

Speech Learning of Chinese as a Second Language

Tone has proved to be the most difficult part of the phonology for learners in the acquisition of Chinese as a second language. Lin remarked that "foreign accent" was not necessarily caused by initials and rhymes, but rather by tones and higher prosodic units (Lin 1996). As a result, tone has been the most intensively examined category in L2 speech learning of Chinese.

Studies in the acquisition of tones have been primarily concerned with errors in production and perception. The former dealt with the acoustic analysis (F0) of L2 tones and the order of tone acquisition. As they adeptly pointed out, when foreigners learn to speak Chinese, they have trouble with pitch register, not pitch contour. Errors in pitch contour happen when the F0 contour of a tone produced by L2 learners deviate from the canonical contour shape of the tone, as in the case of a level tone being said as either a rising or falling tone, or a rising tone being said as either a falling or level tone. Errors in pitch register happen when the pitch contour in L2 speech looks fine, but the pitch register ends up either too high or too low, to wit, a high-level tone becomes a low-level tone, or a falling tone becomes a half falling tone (Wang 1995).

Studies in production errors have been plentiful, targeting L2 learners from L1 backgrounds such as American English speakers (Wang 1995). These studies analyzed L2 productions of the four tones in Putonghua in great details and found that errors in L2 tones could be traced back to both pitch contour and pitch register, with different tones manifesting different error patterns.

Perception errors in L2 speech have been approached in two aspects of tonal perception, largely following studies of speech perception in L1. One aspect was to investigate the perception mode of L2 speakers, using continuous synthetic acoustic stimuli with F0 values modified (Zhang 2010). Zhang (2010) sought to understand how Chinese tones were perceived by L2 learners at different proficiency levels from Thailand, Japan, and Korean. He found that true beginners perceived tones continuously and Thai speakers at the elementary level demonstrated characteristics of categorical perception – not as clear as L1 Chinese speakers, but L2 learners at the intermediate level perceived tones more categorically. Another aspect was to find out the perceptual cues that L2 learners use. Between the two pitch attributes of tones, pitch height, and pitch contour, L2 learners and L1 Chinese speakers exploit different strategies. L2 learners relied more on changes in pitch height whereas L1 speakers on changes in pitch contour.

L2 acquisition of intonation has also been examined vigorously. Some of the topics being explored include declarative intonation produced by L2 speakers, interrogative intonation, acoustic features such as pitch and duration of declarative intonation and interrogative intonation, L2 acquisition of prosodic boundaries, and L2 acquisition of focus in declarative intonation. Specifically, a series of studies (e.g., Shi et al. 2015) working in the framework of intonation pattern developed by Shi Feng and his colleagues investigated error patterns in pitch, duration, and intensity, three important parameters of intonation, in the utterances in interrogative intonation produced by L2 learners. Their results showed L2 learners differ greatly from L1 speakers in terms of the movement of the F0 contours and boundary-related lengthening. LI Zhiqiang 李智强 published a monograph on the acquisition and teaching of Mandarin Chinese phonology, in which he advocated the integration of intonation into pronunciation teaching and proposed a model of phonological acquisition in Putonghua based on speech production and perception (Li 2018). In his analysis, L2 speech teaching and learning in Chinese should focus on two key features of Chinese, tone and syllable. In order to cultivate L2 learners' ability to speak Chinese fluently and naturally, L2 speech teaching would benefit from guided practice of prosodic features of the language such as intonation in larger linguistic units than words and phrases.

In addition to suprasegmental features, acquisition of segments has attracted researchers' attention too. Some studies investigated the error patterns and degree of difficulty in the acquisition of vowels in Putonghua by L2 learners from different L1 backgrounds. They also compared perceptual similarities and acoustic similarities of the two vowel systems in the target language L2 and the native language L1. On the consonants, most studies dealt with stops, affricates, and nasal endings in Putonghua. Wang, 王韫佳, for example, studied perception and production of nasal endings, aspirated and unaspirated consonants by Japanese-speaking L2 learners (Wang 2002). She found that L2 learners' perception of VN syllables in Putonghua is affected by the nuclear vowel: higher accuracy is obtained when the nuclear vowels are acoustically very different. The similar tendency does not materialize in the production data. She also found that the voicing contrast in Japanese influences the Japanese speakers when they process aspirated and unaspirated consonants in Putonghua.

Concluding Remarks

Phonetics in China has undergone enormous advancements in all aspects in the last century, especially the last several decades. It has not only deepened our understanding of Chinese languages in itself but has made remarkable contributions to the understanding of human language. The most noteworthy feature of Chinese is the complex tonal system. Since the initial application of instruments to study the acoustic nature of Chinese tones, the phonetic study of tone has been a representative of Chinese phonetics. The tone letter system proposed by Chao has been extremely handy in transcribe tonal values. The interaction of tone and intonation in Chinese is intriguing and led to a large amount of research. The most recent studies have looked into the representation and processing of tone in the brain.

Phonetic research up until now has become more interdisciplinary and better integrated with other disciplines than ever before. Availability of advanced research instruments has allowed researchers to begin explorations of speech production and perception mechanisms at the neurological level by taking advantage of the most salient features of Chinese. There will be more acoustic studies on Chinese dialects and ethnic minority languages in China. With the advent of new technologies, the research of articulatory phonetics will be more fine-grained, and the online processing of speech will be investigated more deeply in the domain of auditory phonetics. Better integration of phonetics into neuroscience holds great promise for shedding new light on the understanding of the neural mechanism of speech production and perception. Last but not least, phonetic research will find more applications in every corner of our life. It is sure that phonetic research of Chinese and languages in China will continue to provide new insights and perspectives to the general understanding of human languages.

References

Bao, Huaiqiao (鲍怀翘), and Maocan Lin (林茂灿). 2016. *Shiyan yuyinxue gaiyao zengdingban* 实验语音学概要《增订版》 [*An outline of experimental phonetics -revised edition*]. Beijing: Beijing daxue chubanshe 北京:北京大学出版社 [Beijing: Peking University Press].

Bao, Huaiqiao (鲍怀翘), and Lili Yang (杨力立). 1985. *Putonghua fayinqiguan dognzuotexing* (X guang luxiangdai) 普通话发音器官动作特性 (X光录像带) [*Characteristics of articulatory gestures in Putonghua" (X-ray videotapes)*]. Beijing: Beijing Yuyanxueyuan Chubanshe 北京: 北京语言学院出版社 [Beijing: Beijing Language and Culture College Press].

Bao, Huaiqiao (鲍怀翘), and Yuling Zheng (郑玉玲). 2011. Putonghua dongtai yuewei yanjiu 普通话动态腭位研究 [Electropalatography studies on articulation of standard Chinese]. *Nanjing Shifandaxue Wenxueyuan Xuebao*南京师范大学文学院学报 [*Journal of School of Chinese Language and Culture Nanjing Normal University*] 3: 1–11.

Bao, Huaiqiao (鲍怀翘), and Zhizhi Zhou (周植志). 1990. Wayu zhuosongqi shengxuetezheng fenxi 佤语浊送气声学特征分析 [An acoustical analysis on voiced aspiration in Wa language]. *Minzu Yuwen*民族语文 [*Minority Language of China*] 2: 62–70.

Cao, Jianfen (曹剑芬). 1995. Lianxu biandiao yu qingzhong duili 连读变调与轻重对立 [Tone sandhi and stress]. *Zhongguo yuwen*中国语文 [*Studies of the Chinese Language*] 4: 312–320.

Cao, Wen (曹文). 2010. *Hanyu jiaodianzhongyin de yunlv shixian*汉语焦点重音的韵律实现 [*Realization of prosody in focus in Chinese*]. Beijing: Beijing Yuyan Daxue Chubanshe北京: 北京语言大学出版社 [Beijing: Beijing Language and Culture University Press].

Cao, J., and I. Maddieson. 1992. An exploration of phonation types in Wu dialects of Chinese. *Journal of Phonetics* 20: 77–92.

Cao, Mengxue, (曹梦雪), Aijun Li (李爱军), and Fangqiang Fang (方强). 2017. Yingyouer muyu yinweifanchou xidejizhi de jianmo yanjiu婴幼儿母语音位范畴习得机制的建模研究

[A modeling study of infants' categorical phonemic perception mechanism in first language acquisition]. *Zhongguo yuwen*中国语文 [*Studies of the Chinese Language*] 3: 116–127+130.

Chao, Yuenren (赵元任). 1956. *Xiandai Wuyu de yanjiu* 现代吴语的研究 [*Studies in the modern Wu dialects*]. Beijing: kexue chubanshe北京: 科学出版社 [Beijing: Science Press].

——— (赵元任). 1979. *Hanyu kouyu yufa* 汉语口语语法 [*A Grammar of Spoken Chinese*]. Beijing: Shangwu chubanshe 北京:商务出版社 [Beijing: The Commercial Press].

——— (赵元任). 2002. *Zhao yuanren yuyanxue lunwenji*《赵元任语言学论文集》[*Linguistics essay by Yuenren Chao*] Beijing: shangwu yinshu guan 北京: 商务印书馆 [Beijing: The Commercial Press].

Chen, M.Y. 2000. Acoustic analysis of simple vowels preceding a nasal in standard Chinese. *Journal of Phonetics* 28 (1): 43–67.

Dang, J., and K. Honda. 2004. Construction and control of a physiological articulatory model. *Journal of the Acoustical Society of America* 115 (2): 853–870.

Fang, Qiang (方强). 2004. *Putonghua biyin de yanjiu* 普通话鼻音的研究 [*Nasal Sound of Mandarin*]. Zhongguo shehui kexueyuan shuoshi xuewei lunwen 中国社会科学院士学位论文 [CASS Master's Essay].

Fan, S., A. Li, and A. Chen. 2018. Perception of lexical neutral tone among adults and infants. *Frontiers in Psychology* 9: 322.

Gao, J., and R. Shi. 2019. Perceptual development of phonetic categories in early infancy: Consonants, vowels and lexical tones. In *Prosodic studies: Challenges and prospects*, ed. Hongmin Zhang and Youyong Qian, to appear. London: Routledge.

Han, D., N. Zhou, Y. Li, et al. 2007. Tone production of Mandarin Chinese speaking children with cochlear implants. *International Journal of Pediatric Otorhinolaryngology* 71 (6): 875–880.

Hu, F. 2008. The three sibilants in standard Chinese. *The 8th international seminar on speech production*, 105–108.

Hu, Fang. 2014. *A phonetic study on the vowels in Ningbo Chinese*. Beijing: China Social Science Press.

Huang, Shaoming (黄昭鸣), and Ping Wan (万萍). 2008. Sangyin shengxuecanshu yu sangyinyinzhi de xiangguanyanjiu 嗓音声学参数与嗓音音质的相关研究 [Relationship between voice acoustic parameters and vocal quality]. *Linchuang erbiyanhou toujing waike zazhi* 临床耳鼻咽喉头颈外科杂志 [*Journal of Clinical Otorhinolaryngology Head and Neck Surgery*] 6: 251–255.

Jia, Yuan (贾媛), Aijun Li (李爱军), and Yiya Chen (陈轶亚). 2009. Hanyu "shi" he "lian" bianji de jiaodianchengfen yuyin tezheng yanjiu 汉语"是"和"连"标记的焦点成分语音特征研究 [Phonetic properties of "Shi" and "Lian" marked constituents in Chinese]. *Qinghua daxue xuebao (ziran kexue ban)* 清华大学学报 (自然科学版) [*Journal of Tsinghua University (Science and Technology)*] s1: 1294–1301.

Jiang, Chenghui (姜成惠), Xinghui Shi (施星辉), Wanzhong Lin (万林), Hongbing Jiang (江宏兵), and Hua Yuan (袁华). 2009. Elie huanzhe putonghua zuixiao yuyindui tezheng chubuyanjiu 腭裂患者普通话最小语音对特征初步研究 [The characteristics of minimal pairs for cleft palate speech in Chinese Mandarin]. *Tinglixue ji yanyujibing zazhi* 听力学及言语疾病杂志 [*Journal of Audiology and Speech Pathology*] 17(4): 347–350.

Kong, Jiangping (孔江平). 2001. *Lun yuyan fasheng* 论语言发声 [*On language phonation*]. Beijing: Zhongyang minzu daxue chubanshe 北京:中央民族大学出版社 [Beijing: China Minzu University Press].

——— (孔江平). 2003. Sangyin fasheng leixing de shengxue xingzhi ji canshu hecheng 嗓音发声类型的声学性质及参数合成 [A study on the acoustic properties of phonation types and parameter synthesis]. *Diliujie quanguo xiandai yuyinxue xueshu huiyi* 第六届全国现代语音学学术会议 [*The 6th phonetics conference of China*].

Li, Yuming (李宇明). 1991. 1–120 tian yinger fayin yanjiu 1–120 天婴儿发音研究 [A case study of baby's phonological development from 1 to 120 days old]. *Xinli kexue* 心理科学 [*Phychological Science*] 5: 21–25.

Li, A. 2002. Chinese prosody and prosodic labeling of spontaneous speech. *Speech Prosody.*

Li, Aijun. 2015. *Encoding and decoding of emotional speech: A cross-cultural and multimodal study between Chinese and Japanese (prosody, phonology and phonetics)*. 1st ed. Springer.

——— (李爱军). 2017. Putonghua Butong Xinxi Jiegou Zhong qingsheng de yuyin texing 普通话不同信息结构中轻声的语音特性 [Phonetic correlates of neutral tone in different information structures]. *Dangdai yuyanxue* 当代语言学 [*Contemporary Linguistics*] 19(3): 348–378.

Li, Zhiqiang (李智强). 2018. *Yuyin xide yu jiaoxue yanjiu* 语音习得与教学研究 [*Research on Chinese pronunciation: Acquisition and teaching*]. Beijing: Beijing Yuyan Daxue Chubanshe北京:北京语言大学出版社 [Beijing: Beijing Language and Culture University Press].

Li, Wei (李嵬), Hua Zhu (祝华), Barbara Dodd, Tao Jiang (姜涛), Danlin Peng (彭聃龄), and Hua Shu (舒华). 2000. Shuo Putonghua ertong de yuyin xide说普通话儿童的语音习得 [Phonological acquisition of Putonghua speaking children]. *Xinli xuebao* 心理学报 [*Acta Psychologica Sinica*] 32(2): 170–176.

Li, X., J.T. Gandour, T. Talavage, et al. 2010. Hemispheric asymmetries in phonological processing of tones versus segmental units. *Neuroreport* 21 (10): 690.

Lin, Tao (林焘). 1962. Xiandai hanyu qingyin he yufa jiegou de guanxi 现代汉语轻音和句法结构的关系 [Neutral tone and syntactic structures in modern Chinese]. Z*hongguo yuwen* 中国语文 [*Studies of the Chinese Language*] 7.

Lin, Maocan (林茂灿). 1965. Yingao xianshiqi yu putonghua shengdiao yingao texing音高显示器与普通话声调音高特性 [Visible pitch display and the acoustic properties of tones in Putonghua]. *Shengxue xuebao*声学学报 [*Acta Acustica*] 1: 10–17.

Lin, Tao (林焘). 1996. Yuyin yanjiu he duiwai hanyu jiaoxue 语音研究和对外汉语教学 [Phonetics research and teaching Chinese as foreign languages]. *Shijie hanyujiaoxue* 世界汉语教学 [*Chinese Teaching in the World*] 3: 20–23.

Lin, Maocan (林茂灿). 2012. Hanyu yudiao shiyan yanjiu 汉语语调实验研究 [*The experimental study of tone and intonation in Chinese*]. Beijing: Zhongguo shehui kexue chubanshe [Beijing: Chinese Social Sciences Press].

Lin, Tao (林焘), and Lijia Wang (王理嘉). 1985. *Beijing Yuyin Shiyanlu* 北京语音实验录 [*Experimental study on Beijing speech*]. Beijing: Beijing daxue chubanshe 北京:北京大学出版社 [Beijing: Peking University Press].

Lin, Tao (林焘), and S-Y Wang William (王士元). 1984. Shengdiao ganzhi wenti 声调感知问题 [Perception of tone]. *Zhongguo yuyan xuebao*中国语言学报 [*Journal of Chinese Linguistics*] 2: 59–69.

Lin, Maocan (林茂灿), and Jingzhu Yan (颜景助). 1994. Putonghua daibiwei lingshengmu yinjie zhong de xietongfayin普通话带鼻尾零声母音节中的协同发音 [An experiment on the coarticulation between vowels and nasals in zero-initial syllables with nasal endings (VN)]. *Yingyong shengxue*应用声学 [*Journal of Applied Acoustics*] 1: 12–20.

Ling, Feng (凌锋). 2008. Sanzhong yuanyin guizheng fangfa zai suzhouhua yuanyin yanjiu zhong de yingyong yu bijiao三种元音规整方法在苏州话元音研究中的应用与比较 [Application and comparison of three vowel adjustment methods in Suzhou vowel research]. *Zhongguo yuyinxue xueshu huiyi ji qinghe wuzongji xiansheng baisui huadan yuyin kexue qianyan wenti guoji yantaohui* 中国语音学学术会议暨庆贺吴宗济先生百岁华诞语音科学前沿问题国际研讨会 [*China conference on phonetics and the international symposium on the frontiers of speech science in Mr. Wu Zongji's 100th birthday*].

Liu, L., D. Peng, G. Ding, et al. 2006. Dissociation in the neural basis underlying Chinese tone and vowel production. *NeuroImage* 29 (2): 515–523.

Lu, Jilun (路继伦), and Jialing Wang (王嘉龄). 2005. Guanyu qingsheng de jieding 关于轻声的界定 [On defining "Qingsheng"]. *Dangdai yuyanxue* 当代语言学 [*Contemporary Linguistics*] 7(2): 107–112.

Luo, Changpei (罗常培), and Jun Wang (王均). 1957. *Putong yuyinxue gangyao* 普通语音学纲要 [*An outline of general phonetics*]. Beijing: kexue chubanshe 北京:科学出版社 [Beijing: Science Press].

Mattock, K., and D. Burnham. 2006. Chinese and English infants' tone perception: Evidence for perceptual reorganization. *Infancy* 10 (3): 241–265.

Peng, Ce (彭策). 2008. *Jiyu shengxue yu xiaoboshang ji zihuigui moxing de bingtai sangyin zhenduan xinfangfa yanjiu* 基于声学与小波熵及自回归模型的病态嗓音诊断新方法研究 [*Study on novel method of pathological voice diagnosis based on acoustics, wavelet entropy and auto-regressive model*]. Tianjin daxue boshi xuewei lunwen 天津大学博士学位论文 [Tianjin University Dissertation].

Peng, G., H.Y. Zheng, T. Gong, et al. 2010. The influence of language experience on categorical perception of pitch contours. *Journal of Phonetics* 38 (4): 616–624.

Shen, Jiong (沈炯). 1992. Hanyu yudiao moxing chuyi 汉语语调模型刍议 [Model of Chinese intonation]. *Yuwen yanjiu*语文研究 [*Linguistic Researches*] 4: 16–24.

Shen, X.S., and M. Lin. 1991. A perceptual study of Mandarin tones 2 and 3. *Language and Speech* 34 (2): 145–156.

Shi, Feng (石锋). 2013. *Yudiao geju-shiyan yuyanxue de dianjishi*语调格局-实验语言学的奠基石 [*Intonation pattern- foundation of experimental linguistics*]. Shanghai: shangwu yinshu guan 上海: 商务印书馆 Shanghai: The Commercial Press.

Shi, Feng (石锋), Baoying Wen (温宝莹), and Yajuan Han (韩亚娟). 2015. Riben xuesheng hanyu yiwenju yudiao xide shiyan duibi fenxi日本学生汉语疑问句语调习得的实验研究 [Study on the Japanese learner's acquisition of interrogative sentence intonation]. *Dishisanjie Quanguo renji yuyin tongxun xueshu huiyi* 第十三届全国人机语音通讯学术会议 [*The 13th national conference on man- machine speech communication*].

Shi, R., J. Gao, A. Achim, et al. 2017. Perception and representation of lexical tones in native Mandarin-learning infants and toddlers. *Frontiers in Psychology* 8: 1–11.

Si, X., W. Zhou, and B. Hong. 2017. Cooperative cortical network for categorical processing of Chinese lexical tone. *Proceedings of the National Academy of Sciences* 114 (46): 12303–12308.

Wang, Bei (王), Yang, Yufang (杨玉芳) and Lv, Shinan (吕士). 2004. Hanyu yunlv cengji jiegou bianjie de shengxue fenxi汉语韵律级结构界的声学分析 [Acoustic Analysis on Prosodic Hierarchical Boundaries of Chinese]. *Shengxue xuebao*声学学报 [*ACTA ACUSTICA*], 2004.1:29–36.

Wang, Li (王力). 1932. *Bobai fangyan shiyanlu* 博白方言实验录 [*Experimental Study on Bobai Dialect*]. Bali daxue boshi xuewei lunwen 巴黎大学博士学位论文 [Université de Paris Dissertation].

Wang, S.Y. 1976. Language change. *Annals of the New York Academy Sciences* 208: 61–729.

Wang, Yunjia (王韫佳). 1995. Yetan meiguoren xuexi hanyu shengdiao 也谈美国人学习汉语声调 [On American Chinese leaners' study of intonation]. *Yuyanjiaoxue yu yanjiu*语言教学与研究 [*Language Teaching and Linguistic Studies*] 3: 126–140.

——— (王韫佳). 2002. Riben xuexizhe ganzhi he chansheng putonghua biyin yunmu de shiyan yanjiu 日本学习者感知和产生普通话鼻音韵母的实验研究 [An experimental study on the perception and production of nasal codas by Japanese learners of Chinese Putonghua]. *Shijie hanyujiaoxue* 世界汉语教学 [*Chinese Teaching in the World*] 2: 47–60.

Wang, Guoming (王国民). 2013. *Chun'e'lie xiufushu yu yuyin zhiliao* 唇腭裂修复术与语音治疗 [*Surgical repair and speech therapy for clefts*]. Shanghai: shanghai shijie tushu chugan gongsi 上海: 上海世界图书出版公司 [Shanghai: Shanghai World Book].

Wang, Yunjia (王韫佳), and Xihang Qin (覃夕航). 2015. Putonghua danzidiao yangping he shangsheng de bianren ji qufen 普通话单字调阳平和上声的辨认及区分 [Grcognization of Tone2 and Tone3 in Mandarin]. *Yuyan kexue*语言科学 [*Linguistic Sciences*] 4: 337–352.

Wang, Jianhua (王建华), Juan Liu (刘娟), Yong Wang (王勇), Aiqin Lv (吕艾芹), and Renhua Zhang (张仁华). 2003. Elie putonghua yuyin de tedian yu fenxi 腭裂普通话语音的特点与分析 [The feature of cleft palate in Mandarin]. *Linchuang kouqiang yixue zazhi* 临床口腔医学杂志 [*Journal of Clinical Stomatology*] 19(8): 464–466.

Wong, P. 2013. Perceptual evidence for protracted development in monosyllabic Mandarin lexical tone production in preschool children in Taiwan. *Journal of the Acoustical Society of America* 133 (1): 434–443.

Wu, Zongji (吴宗济). 1964. Putonghua yuanyin he fuyin de pinpu fenxi ji gongzhenfeng de cesuan 普通话元音和辅音的频谱分析及共振峰的测算 [The spectrogram analysis of the vowels and consonants in standard colloquial]. *Shengxue xuebao*声学学报 [*Acta Acustica*] 1(1): 33–40.
——— (吴宗济). 1982. Putonghua yujv zhong de shengdiao bianhua 普通话语句中的声调变化 [Pitch contour in Mandarin]. Z*hongguo yuwen*中国语文 [*Studies of the Chinese Language*] 6: 439–450.
——— (吴宗济). 1989. Putonghua danyinjie yutuce 普通话单音节语图册 [*Chinese Putonghua monosyllable spectrograms*]. Beijing: zhongguo shehui kexue chubanshe北京: 中国社会科学出版社 [Beijing: China Social Sciences Press].
——— (吴宗济). 1992. Putonghua lingshengmu yinjie qishiduan de shengxue fenxi普通话零声母音节起始段的声学分析 [Acoustic analysis of the initial segment of zero initial syllables in Mandarin]. *Disanjie quanguo renji yuyin tongxun huiyi* 第三届全国人机语音通讯会议 [*The 3th national conference on man-machine speech communication*].
Wu, Zongji (吴宗济), and Maocan Lin (林茂灿). 1989. *Shiyan yuyinxue gaiyao* 实验语音学概要 [*An outline of experimental phonetics*]. Beijing: Gaodeng jiaoyu chubanshe 北京: 高等教育出版社 [Beijing: Higher Education Press].
Wu, Zongji (吴宗济), and Guohua Sun (孙国华). 1990. Putonghua qingcayin xietongfayin de shengxue moshi普通话清擦音协同发音的声学模式 [Acoustic coarticulation patterns of voiceless fricatives in CVCV in standard Chinese]. *Yuyin yanjiu baogao*语音研究报告 [*Annual phonetic research report*].
Wu, Zongji (吴宗济), Jianfen Cao (曹剑芬), and Lili Yang (杨力立). 1979. Shiyan yuyinxue zhishi jianghua 实验语音学知识讲话 [Notes on experimental phonetics]. Z*hongguo yuwen*中国语文 [*Studies of the Chinese Language*] 1979连载于第1,2,4,5,6 期, 1,2,4,5,6.
Xi, Jie (席洁), Wei Jiang (姜薇), Linjun Zhang (张林军), and Hua Shu (舒华). 2009. Hanyu yuyin fanchouxing zhijue jiqi fazhan汉语语音范畴性知觉及其发展 [Categorical perception of VOT and lexical tones in Chinese and the developmental course]. *Xinli xuebao*心理学报 [*Acta Psychological Sinica*] 41(7): 572–579.
Xi, J., L. Zhang, H. Shu, et al. 2010. Categorical perception of lexical tones in Chinese revealed by mismatch negativity (MMN). *Neuroscience* 170 (1): 223–231.
Xu, Yi (许毅). 1986. Putonghua yinlian de shengxue yuyinxue tezheng普通话音联的声学语音学特性 [Acoustic properties of juncture in Mandarin]. Z*hongguo yuwen*中国语文 [*Studies of the Chinese Language*] 5: 353–360.
Xu, Yunyang (徐云扬). 1990. Shanghaihua yuanyin qinghua de yanjiu上海话元音清化的研究 [Study on unvoicing of vowels in Shanghai dialect]. *Dangdai yuyanxue* 当代语言学 [*Contemporary Linguistics*] 3: 19–34.
Xu, Y. 1997. Contextual tonal variations in Mandarin. *Journal of Phonetics* 25 (1): 1–83.
———. 1999. Effects of tone and focus on the formation and alignment of F0 contours. *Journal of Phonetics* 27 (1): 55–105.
———. 2005. Speech melody as articulatorily implemented communicative functions. *Speech Communication* 46 (3): 220–251.
Yang, Shun'an (杨顺安). 1994. *Mianxiang shengxue yuyinxue de putonghua yuyin hecheng jishu*面向声学语音学的普通话语音合成技术 [*An acoustic-phonetics oriented speech synthesis technology in Putonghua*]. Beijing: zhongguo shehui kexue chubanshe 北京: 中国社会科学出版社 [Beijing: China Social Sciences Press].
Yang, Shun'an (杨顺安) and Cao, Jianfen (曹剑芬). 1984. Putonghua erhe yuanyin de dongtai texing普通话二合元音的动态特性 [Dynamic properties of diphthongs and compound vowels]. *Yuyan yanjiu* 语言研究 [*Studies in Language and Linguistics*], 1984.1:15–22.
Yang, J., and L. Xu. 2017. Mandarin compound vowels produced by prelingually deafened children with cochlear implants. *International Journal of Pediatric Otorhinolaryngology* 97: 143–149.
Yuan, Yue (袁悦), Lingling Gu (顾玲玲), Jingya Chang (常静雅), Xiaojun Zhang (张晓俊), and Zhi Tao (陶智). 2015. Jingyan motai fenjiefa shibie shengdai xirou heshengdai nangzhong de yanjiu 经验模态分解法识别声带息肉和声带囊肿的研究 [Empirical Mpdel decomposition

in the research of vocal polyp and vocal cyst]. *Xinxihua yanjiu* 信息化研究 [*Informatization Resaerch*] 2: 27–32.

Zhang, Linjun (张林军). 2010. Muyu jingyan dui liuxuesheng hanyu shengdiao fanchouhua zhiju de yingxiang 母语经验对留学生汉语声调范畴化知觉的影响 [On the influence of native language on the categorical perception of Chinese tones]. *Huawen jiaoxue yu yanjiu* 华文教学与研究 [*TCSOL Studies*] 2: 15–20.

Zhou, Dianfu (周殿福), and Zongji Wu (吴宗济). 1963. *Putonghua fayin tupu* 普通话发音图谱 [*The atlas of sound spectrograms in Putonghua*]. Beijing: shangwu yinshu guan 北京: 商务印书馆 [Beijing: The Commercial Press].

Zhou, Haiyan (周海燕), Xuguang Xu (徐旭光), Guomin Wang (王国民), Zhuang Si (姒壮), and Zhiyong Lv (吕智勇). 2012. E yan bihe buquan huanzhe yichang shengxue tedian de fenxi yanjiu 腭咽闭合不全患者异常声学特点的分析研究 [The study analysis of Velopharyneal incompetent' abnormal speech characteristics]. *Kouqiang yixue yanjiu* 口腔医学研究 [*Journal of Oral Science Research*] 28(5): 457–463.

Zhu, Xiaonong (朱晓农). 2004. Jipin guiyihua – ruhe chuli shengdiao de suiji chayi 基频归一化——如何处理声调的随机差异? [F0 normalization: How to Deal with between-speaker tonal variations?]. *Yuyan kexue*语言科学 [*Linguistic Sciences*] 2: 3–19.

——— (朱晓农). 2009. Fashengtai de yuyanxue gongneng 发声态的语言学功能 [Phonation types: Phonological categories and linguistic functions]. *Yuyan yanjiu* 语言研究 [*Studies in Language and Linguistics*] 3: 1–19.

[This paper was published at The Palgrave Handbook of Chinese Language Studies 2021]