Foreword

I was born in Shanghai in 1946, the youngest son to parents who were natives of Zhenjiang 鎮江, a city to the north of the Yangtze River. I grew up in Hong Kong speaking a variety of dialects, vastly different dialects. Seemingly an impossible task, but, to a kid, it was a simple matter of matching different sounds to different faces. My mother spoke to us only in the Zhenjiang dialect, a member of the southern Mandarin family, which was truly my mother tongue. My siblings spoke with each other mostly in Shanghainese, a Wu 吳 dialect. I was never conversant with that Shanghai accent, and it wasn't until I came to the San Francisco Bay Area in the late 1960s that I began to realize those buzzing apical sounds, which I used to find so conceitedly unnatural, had always been in my language repertoire; when I hung around friends who had come from the old Shanghai, I spoke almost like a (semi-)native. And, of course, as I went to school in Hong Kong, from kindergarten to graduate school, Cantonese was by default my second native, and most fluent, language. But, for whatever inexplicable reasons, however proficient I may seem to have mastered these dialects I grew up with, I often speak with a little accent, which natives can easily detect—and I am not one of them.

I started Mandarin in school when I was seven or eight years old. Thanks to the teachers I had since then, especially those who came from Beiping 北平 in the 1940s, I acquired a rather competent command of this northern dialect. I began teaching Mandarin at Berkeley as a teaching assistant in 1970, and I continued my work in language pedagogy for the next thirty years. People were always surprised that I was from Hong

Kong. "Where did you learn to roll your tongue and acquire that beautiful Peking accent?" A flattering remark that never failed to boost my pride as a language virtuoso. That was until one day when I went into a store where the storeowner was an old Beijing gentleman. Perhaps out of curiosity, he asked, "Where are you from?" And, with a mischievous grin I answered, "Beijing." He paused, and then curtly remarked, "You must have been away for quite some years." I got similar feedback on my Shanghainese and even Cantonese—my tones always seem a teeny tiny bit off key. When I visited Yangzhou 揚州, a neighboring city of Zhenjiang, the people in town refused to talk to me in the dialect I felt most intimate with. Either my command of words was strangely deficient, or my intonation was perceptibly odd. One of the first linguistic terms I learned when I came to Berkeley to do linguistics was "SWONL"—an acronym for "a speaker without a native language." In simple terms, I am a good example of SWONL competence. English came to me as a foreign language. But, growing up and going to school in Hong Kong and having been living in the United States for decades, I have had more than plentiful chances to work on and polish my English. When I was in England in the late 1970s, I was on a train going to Portsmouth. I struck up a short conversation with the ticket conductor. The first response I got from him was, "Are you from America?" Apparently, I was speaking with a Californian accent. It took me years of hard effort to switch from the British rounded *o* [ɒ] to the American unrounded *o* [ɑ] as in the name Bob, or the vowel sound of *ey* to *e* as in the word "again." But, has anyone ever mistaken me as an America-born Chinese? No. Another proof of my SWONL identity.

Did I ever feel discouraged by my less-than-competent performance in language? Perhaps in the early years when I realized that my faulty command was hopelessly beyond repair. Nonetheless, when I was in college, a professor offered me these words of encouragement. "A great pianist," he said, "aspires to sounds that could be very different from the kind of notes that engages the mind and imagination of a musicologist." In other words, he concluded, "A polyglot does not necessarily make a good scholar of language." So, what's wrong if I am not blessed with the gift from the Babel Tower? I could still pursue my interest in language by exploring its immense domain in words and sounds.

I took two linguistics classes as an undergraduate, one in general linguistics and one in historical Chinese phonology. The classes were vastly different in their coverage and research methodology. But what I learned from one class bore significantly on what I took in the other. Language-specific examinations helped me understand how Bloomfield and Chomsky would arrive at their general characterizations of language, and the theoretical input made me wonder how I could approach Chinese with a different perspective. In 1967, I decided to pursue further training under the tutelage of Professor Chou Fa-kao 周法高 in the postgraduate program at The Chinese University of Hong Kong (CUHK). Aside from attending his lectures, I was entrusted with two additional assignments: to translate B. Karlgren's *Compendium of Phonetics in Ancient and Archaic Chinese* (1954), and to assist in another translation project of Y. R. Chao's 趙元任 *A Grammar of Spoken Chinese* (1968). Both projects, demanding as can be imagined, gave me the extra advantage not only to keep abreast with the most current developments in the fields, but also to plough through the writings by two great minds, page by page and word by word. I began to appreciate how their thoughts came together from data, and how their arguments were constructed with implications far beyond the scope that the pages covered. At the end of my first year in the MA program, I wrote an article on the use of aspectual markers in Cantonese, which Professor Chou recommended for publication.[1] It was indeed a huge vote of confidence on what I was trying to explore, and with his further encouragement, I began to work on my Master's thesis, a study of Cantonese grammar. I was able to gain access to a rather extensive corpus of raw data, including 15 hours of recording of a radio soap opera, which I transcribed and examined with painstaking care. Following Chao's theoretical model, I made an attempt to analyze the way how Cantonese

---

1 "Some Common Predicative Suffixes in Cantonese" (in Chinese), *Journal of The Institute of Chinese Studies of The Chinese University of Hong Kong*, 3.2 (1970): 459–487.

behaved in a structuralist framework. The thesis was eventually published in 1972. In the same year, my translation of Karlgren's phonology came out in Taiwan.

When I embarked on my doctoral studies at the University of California, Berkeley, I had the great privilege to work under Professor Kun Chang 張琨 on Chinese phonology, dialectology, and historical grammar. I decided to choose grammar as the focus of my research and took on a corpus of medieval manuscripts, generally known as the *bianwen* 變文, as my data for looking into early Chinese grammar. The dissertation was filed in 1974. For the next decade or so, during my early years of teaching at Berkeley, I was primarily engaged in language pedagogy and literary studies, and it was not until the 1980s that I began to gradually return to my initial interest in Cantonese. With a lapse in time of many years, and physically living in an English-speaking world, my understanding of Cantonese was essentially limited to what I knew of it in the late 1960s. To make up for what was lacking, I concentrated on what I could find in printed materials and what I could gather from informants, admittedly rather small in number, in the Bay Area, looking for language phenomena that might have escaped my attention in the past. The more I ventured out, the more piqued was my curiosity. One of the first articles I published during that period was based on an investigation of more than three hundred *xiehouyu* 歇後語 expressions, witty and often enigmatic folk sayings, which I examined with a specific focus on the structural principles and the intricate linguistic mechanisms involved in the construction of these colloquial idioms. Another project just as challenging was one on kinship terms; I collected a rich stock of familial terms in Cantonese, both designative and vocative, across a span of five generations, and I reported on the various linguistic operations responsible for the making and (the) use of these monikers, including phonological variation, morphological modification, semantic shift, and social borrowing. My articles on the measure words and the pretransitive in Cantonese allowed me to compare Cantonese with Mandarin in their use of grammatical patterns which they seem to share, but actually with many subtle and intrinsic typological differences between the two dialects. These research projects, tricky and time-consuming during the process, proved not only rewarding in what I had to offer; the efforts were particularly gratifying to me on a personal level—I felt like coming home as a prodigal son.

It was during the 1990s when I took a radical turn, almost by accident, from where my work had been headed all along. From my fascination with the contemporary language, I turned to diachronic investigation. One day, on one of my regular visits to the East Asian Library at Berkeley, I came across an old volume, a bit dusty, hiding between hard-bound publications, a book on Cantonese published in the late 19th century. I took a quick look under the dim light at the stack, and I quickly checked it out—for the rest of the semester. It was a language manual, in both Chinese characters and romanization, designed for teaching Westerners a language often considered almost impossible to decipher, let alone to acquire. Lessons after lessons, the texts were full of surprises for a 20th-century reader. There were sounds and words that I failed to recognize, and strangely constructed sentences on almost every other page. Like the spell that opened the cave in the Arabian story, this mysterious volume unlocked a gate for me to venture into an unchartered land; it gave me a privileged vantage point to look at the past of a language that had not seemed to have much of a history of its own. The article I published in 1997, "Completing the Completive," marked the beginning of my next phase of linguistic pursuit, a historical project to reconstruct early Cantonese grammar.

I returned to Hong Kong at the end of the 1990s to take up new teaching appointments, initially at the Hong Kong University of Science and Technology (HKUST) and then at CUHK, my alma mater, until my retirement in 2010. For almost two decades, I focused my efforts primarily on old Cantonese. Through generous supports of both universities and also of the Research Grants Council in Hong Kong, I was able to conduct intensive search for language materials of the 19th and early 20th centuries. Also through the help of many friends who shared my interests, I put together a good collection of early writings of various kinds, including language manuals, Bible translations, and maps of Hong

Kong. Some of the titles were included in a database I built, entitled *Early Cantonese Colloquial Texts: A Database*, now available online.[2] A number of new databases have been introduced by other scholars in recent years, constituting an enviably rich pool of raw materials awaiting exploration.

When faced with piles of pages from old documents and thousands of entries online, I often wonder: where do we begin our search and research? Many of the words we find in the data are strangely familiar—yes, we seem to recognize the characters and/or the spelling, but the meanings and usages seem to evade our comprehension. Many a time, I would dismiss what I could not decipher as textual errors—and, in fact, there are quite a few graphic and spelling mishaps in the texts. For example, one of the earliest texts, Morrison's *Vocabulary of the Canton Dialect* (1828), did not include tone marking in its romanization. Could it be an indication that the language was atonal in the early 19th century, or was it simply a deliberate omission so as to facilitate the transcription? Consultation with other texts around the same period readily relieved our concerns—tones were left out by design. Oftentimes when I was hesitant with a quandary, I would recall what Professor Chang used to say about the use of data from ancient documents: let the data speak to you before you speak on behalf of the data. His words rang a bell and I would then studiously double check the suspicious items against what other sources might have before I would pronounce what they were. Were they recorded by mistake? Or, they could represent features of an early language that were no longer active in current usage. A good example of this baffling phenomenon is the use of *hiu* 曉 as a marker for the perfective aspect in early Cantonese. The first time I spotted the form, I did not recognize the character, and I was appalled by its given pronunciation, which sounded to me more like a curse word with obscene innuendo. Upon further search, however, I found *hiu* popping up in almost every old text, and it was not until the 1930s that the marking was eventually replaced by a new suffix, *jo* 咗, a grammatical form that we still use today. Or, an example from phonology. The deictic marker, *go* 個 or 嗰 ("that") was marked with a

2 See https://database.shss.hkust.edu.hk/Candbase/.

level tone in all early materials, making it completely identical with the classifier *go* 個, in writing as well as in pronunciation. The first time *gogo* 個個, meaning "that item," appeared in my data; the combination struck me as a clear typo where the deictic was wrongly marked in tone. No, again, it wasn't. The deictic *go* retained its level tone till much later when it underwent a tonal modification, a process that was convoluted but clearly documented in the texts.

From those early blundering trials, I learned that whenever I came across a feature that looked unusual, I would pause before hurriedly declaring it fake or wrong, and my mind would start to look for a possible and reasonable explanation. Sleuthing has indeed been such a rewarding challenge for me, a task that has enabled me to read between lines and to trace the changes across timelines. I have written a number of articles on early Cantonese, covering topics ranging from phonology to grammar and to lexicology. The texts inform us about the many subtleties in the old language that have disappeared over the years, and what we have learned from them has in turn helped us understand how language evolves often in paradigms that can be mapped out in clear terms. When I look at an old map of the 19th century, I will no longer be disturbed by the unfamiliar spelling of place names; they are clues to changes in the sound system, and, in addition, they could also be demographic markers for speakers of different dialects who settled in Hong Kong during the early years.

Not much was recorded in official documents or gazetteers about the early history of Hong Kong. Our knowledge of Cantonese is likewise quite limited except for occasional mentions of its culture and customs in writings here and there. For a long time, Cantonese was deemed a local dialect enjoying little prestige among the intellectuals. Its language and its origin remained much of a mystery until the mid-20th century when scholars started to accord it with increasing attention. The efforts were, however, primarily focused on contemporary Cantonese, its language system of sounds and grammar, its typological affiliation with other Yue 粵 dialects in South China, the role it played as the medium of instruction, and the political status it enjoyed vis-à-vis English and Mandarin during the colonial reign. What do we know about the past of the Cantonese language? Knowledge of the past requires access to bona-fide materials from the past. Language manuals of the 19th century have supplied the missing link.

Systematic preparation of language materials began with the missionaries and scholars who came to the Pearl River Delta for various reasons and who saw the need to learn the local language so as to facilitate their activities in official, religious, and commercial businesses. Instructional books were compiled, dictionaries were published. And, to cater to a Western audience, romanization was a necessary means to transliterate and record the speech of the locals. Grammatical notes were drawn up to explain how the language worked, and vocabulary projects were launched to collect and translate words of common usage. Pedagogical manuals also showed up overseas and, subsequently, Cantonese textbooks were prepared for dialect speakers around the Delta. Language learning became a vibrant enterprise at the turn of the 20th century, as a result of which some manuals went through several printings so as to meet the demand. *Cantonese Made Easy*, for example, saw its first publication in 1883 and was subsequently reissued in 1887, 1907, and 1924.

Thanks to the dedicated efforts of these early pedagogues for faithfully recording the language of the time, we are now blessed with a rich pool of firsthand materials that make it possible for a modern reader to go back in time and experience vicariously what it was like speaking a language that gave rise to modern Cantonese. If we were to eavesdrop on a chitchat at a street corner in the 19th-century Hong Kong, would we be able to follow their conversation? If possible, would we be able to partake in their discussion? According to what we could gather from the century-old materials, the language then was essentially the same as modern Cantonese. There are differences in sounds, in words and in grammar, but the overall ingredients and workings are characteristically Cantonese. Granted that general identification, how do we tell that the language is of an early period? What are the obvious features that would readily distinguish the old from our speech today?

As a first approximation, the following is a list of such distinctive features that I have gathered from my examination. By no means exhaustive or in great detail, the list provides a useful guide to distinguish the new from the old, with features that highlight not only some of the major developments of the language in recent times, but also the paths and pace that the language took in the process of becoming what it is today.

In terms of sounds:

(1) There was a set of apical affricates and fricative, *ts*, *ts’*, and *s*, that coexisted with their alveolar counterparts *tʃ*, *tsʃ’*, and *ʃ* in early Cantonese. Words such as 資秋先 (of the first set) and words as 之醜扇 (of the second set) now share the same initial consonants.
(2) There was an apical vowel to accompany the apical consonants. A character like 資 was spelled as *tsz*, in contrast with 之 *chi*.
(3) There were two finals *om* and *op*, which eventually became *am* and *ap* in later development (e.g., 庵 *om* → *am*, and 合 *hop* → *hap*).
(4) The 陰平 tone was high-falling in pitch contour, with a possible high-level reading as a variant form. In modern Cantonese, high-level is the norm.

In terms of grammar:

(1) The *yes/no* question was formed with the construction V-NP+*mh*-V (e.g., 食飯唔食？) in early Cantonese, but the pattern was eventually replaced by V *mh*-VP (e.g., 食飯唔食？ → 食唔食飯？). The direction of deleting identical constituents has changed from forward to backward in application. A disyllabic compound XY such as 歡喜 formed its *yes/no* question with XY *mh*-XY, i.e., 歡喜唔歡喜, which has now followed the same pattern of backward deletion in modern Cantonese, namely, X *mh*-XY, to yield 歡唔歡喜.
(2) The perfective marker is *hiu* 曉 instead of *jo* 咗. For example, 食咗飯 would be 食曉飯 in early Cantonese.
(3) The deictic marker for *that* was *go* 個 or 嗰 with a mid-level tone, which later became high-rising.
(4) The adverbial marker 咁 was *gom* in pronunciation with a mid-level tone, which later became *gam* with a high-rising tone. The tonal change was accompanied by a change of the vowel from *o* to *a*.
(5) The negative imperative was marked by 莫個 *mogo*, a marker that eventually disappeared in later years.

In terms of vocabulary, words such as *gaoguan* 交關 (very much), a degree modifier, *shatshou* 實首 (truly), an adverbial modifier, and *fenggau* 風颶 (typhoon), were common terms in old Cantonese but have become obsolete in the 20th century. There is indeed a huge group of words, nouns and verbs, substantive and functional, in the data that deserves close lexicographic attention.

By early Cantonese, we refer to the language primarily of the 19th century, a timeline that may be extended to the 1930s, when changes began to crop up in our data. By the mid-20th century, with almost all the early phonological and grammatical features slipping into obscurity, Cantonese greeted its speakers with new sounds and revised grammar, and of course with many novel terms in circulation to capture the new culture of a modern era. If we were to be shown a text without a clear date of composition or publication, we could resort to its language and judge whether it was a work of which period, say, roughly before or after the turn of the 20th century. Or, as I once noted in another article, if we were to produce a movie about the revolutionary history in early Hong Kong or Canton, we would have to rely not only on the movie set or the costume design to create a credible period drama; a good script, with language devoid of modern sounds and word usages, would be just as crucial.

As I have worked on the Cantonese of the 1960s, I am quite aware of some of the changes that the language has undergone since then. Choice of diction of course accounts for the biggest change; loan words such as *shido* 士多 (store), *shidaam* 士擔 (postage stamp), or common terms at least in my speech such as *baakfo gungsi* 百貨公司 (department store), are no longer prevalent among the younger generation of speakers. A time expression such as the use of *gwat* 骨, a borrowing from "quarter (of an hour)," has dropped out in usage. The grammatical use of adjectival reduplication, such as *honghong* 紅紅 with a change tone (high-rising) on the first syllable "very, very red," has become a morphological process of the bygone years. Phonologically, aside from denasalizing *n-* to *l-* and *ng-* to zero initial, the tendency to palatalize dental obstruents (e.g., *ts* → *tɕ*) and the dropping of the final *-k* and *-ng* (e.g., *dang* → *dan*; *dak* → *dat*) have often been red-flagged as indications of a corrupted form of "lazy

Cantonese." Laziness or not, these pronunciations have come to define a new form of speech habits, a new setup that has in fact reconfigured the entire Cantonese phonological system, a paradigmatic change that I report in my article on the language of the 21st century.

In recent years, I have produced a number of articles on old Cantonese, with findings that I feel will be useful for our future efforts to trace the development of a language that has gone through many rounds of incredible and, at times dramatic, changes during the last two hundred years. Now that as I am in my seventies, I feel it is perhaps time for me to put together some of these findings into a collection for general reference. As the articles appear some in Chinese and some in English, the collection is to be in two volumes.[3] The present volume, entitled *Cantonese: Since the 19th Century*, contains a total of eight articles in English, four each on Old Cantonese and on Contemporary Cantonese. The articles appeared in various journals, each here properly acknowledged in the volume. Aside from some stylistic editing and amendments, little changes have been made to the content or the use of data.

I began this foreword by saying that I was a SWONL with no real native competence in any of the languages I speak or have studied. Having worked on Cantonese for more than fifty years, however, I have always known that Cantonese is the language most intimate to my mind and soul. I might have spoken with a slight accent, but I am sensitive to any change, however subtle, that may have happened to its sounds, words, or grammar. When I was working in Hong Kong, I took the public transport to and from work, and I often eavesdropped on what others were talking, either on the phone or to those next to them. Of course, it was the speech rather than its actual content that would catch my fancy. Ha, ha—this person speaks perfect Cantonese, and that person unabashedly is showing off his "lazy" accent. Inadvertent use of "incorrect" grammar? Or, a slip of tongue with some never-heard-of cursing slangs. Or, a random mix of English and Cantonese, a code-switching phenomenon quite common in

---

3 The Chinese volume is titled 香港粵語：二百年滄桑探索 (*Cantonese in Hong Kong: An Exploration of the Past 200 Years*) (Hong Kong: The Chinese University of Hong Kong Press, 2021) and consists of 12 articles with 3 appendices.

Hong Kong. I took notes and wrote brief remarks about my observations. Many of these fortuitous observations would become anecdotal details in my lectures in class, exciting matters to discuss with colleagues over coffee or in workshops. I am hopeful that, someday, someone could take on a new project to decipher the language of today and determine the distinctive features of the 21st-century Cantonese. As my professor claims, knowing a language and knowing of a language are pursuits just as demanding and meaningful as each other. A linguist may not be talented in speaking a foreign tongue, and, likewise, a polyglot does not necessarily know the distinction between a *phone* and an *allophone*. Yes, that distinction may appear as a comforting reassurance to a SWONL; yet, I am also of the opinion that the fun of studying a foreign tongue is such an enjoyable and fulfilling experience that I would never forego. I strongly believe that the more languages you know, the more you will learn about your own language, and the better you will appreciate those around you who speak different tongues or with different accents. When I was working on the pretransitive in Cantonese, I relied on my Mandarin to redefine its use in Cantonese. When I was working on place names in an 1866 map of Hong Kong, I had to consult several other dialects to come to terms with names that were spelled drastically different from their Cantonese pronunciations. And, when I was trying to pick up a new dialect, I tried to figure out its phonological relationships to the dialects I already knew, so as to help me acquire a better and more efficient command of its sounds and tones. I truly believe that, while new findings in linguistics have much to bear on language pedagogy, learning to speak in a foreign tongue has just as much to contribute to our scientific study, empirical or theoretical, of the greatest gift that humans are ever blessed with, namely, *speech*.

I should take this chance to thank my professors, especially Professor Chou Fa-kao at CUHK and Professor Kun Chang, my mentor at UC Berkeley. I owe all of what I know about language and linguistics to their teaching and support. To my colleagues and many, many young friends whose passion for language has made it possible for us to cross paths and to work together in a fertile land for linguistic cultivation. What I have reaped is far more than I can ever imagine or deserve. Special thanks to The Chinese University of Hong Kong Press for its generous offer to publish the collection, taking care of every detail in the process, including obtaining approval from individual journals where my articles originally appeared. In particular, I wish to thank Mr. Brian Yu of the Press for his meticulous efforts in editing both Chinese and English volumes, and also Ms. Pauline Pang for her willingness and patience in helping me with reformatting and retyping some of the articles at the early stage.

My wife, Adaline, has been my greatest support throughout the years. She is my muse and my best companion, especially during the pandemic crisis when I was always in an absolutely lockdown mode working on the manuscript. Like me, Adaline grew up in Hong Kong; but unlike me, she was born in Canton to parents who are natives of the Canton region. For decades, she is my informant, and her Cantonese has always been my most appreciated source of information. If the Zhenjiang dialect is my mother tongue, then, without doubt, Cantonese is my wifey speech. To her, therefore, I dedicate this collection on Cantonese.

Hung-nin Samuel Cheung