PLANNING CHINESE CHARACTERS
The last half century has witnessed an explosive shift in language diversity not unlike the Biblical story of the Tower of Babel, but involving now a rapid spread of global languages and an associated threat to small languages. The diffusion of global languages, the stampede towards English, the counter-pressure in the form of ethnic efforts to reverse or slow process, the continued determination of nation-states to assert national identity through language, and, in an opposite direction, the greater tolerance shown to multilingualism and the increasing concern for language rights, all these are working to make the study of the nature and possibilities of language policy and planning a field of swift growth.

The series will publish empirical studies of general language policy or of language education policy, or monographs dealing with the theory and general nature of the field. We welcome detailed accounts of language policy-making – who is involved, what is done, how it develops, why it is attempted. We will publish research dealing with the development of policy under different conditions and the effect of implementation. We will be interested in accounts of policy development by governments and governmental agencies, by large international companies, foundations, and organizations, as well as the efforts of groups attempting to resist or modify governmental policies. We will also consider empirical studies that are relevant to policy of a general nature, e.g. the local effects of the developing European policy of starting language teaching earlier, the numbers of hours of instruction needed to achieve competence, selection and training of language teachers, the language effects of the Internet. Other possible topics include the legal basis for language policy, the role of social identity in policy development, the influence of political ideology on language policy, the role of economic factors, policy as a reflection of social change.

The series is intended for scholars in the field of language policy and others interested in the topic, including sociolinguists, educational and applied linguists, language planners, language educators, sociologists, political scientists, and comparative educationalists.
PLANNING CHINESE CHARACTERS
Reaction, Evolution or Revolution?

by

SHOUHUI ZHAO
Nanyang Technological University, Singapore

RICHARD B. BALDAUF Jr.
University of Queensland, Australia
DEDICATION

This book is dedicated to
our mentors and our colleagues
who are working to understand the
limits of language planning
# Contents

Dedication.......................................................................................................................... v

Authors ..................................................................................................................................... ix

Abbreviations ........................................................................................................................ xi

Preface .................................................................................................................................... xiii

Foreword ............................................................................................................................... xv

Prologue ............................................................................................................................... 1

Chapter 1 Making Hanzi Accessible................................................................................... 23

Chapter 2 Reflections on New Perspectives ..................................................................... 71

Chapter 3 New Challenges for a Digital Society ............................................................ 111

Chapter 4 Standardization as a Solution ......................................................................... 137

Chapter 5 Influencing Outcomes ..................................................................................... 191

Chapter 6 Envisioning the Future ..................................................................................... 249

Chapter 7 Some Critical Issues ......................................................................................... 287
## Contents

- **Acknowledgments** ................................................................................... 329
- **Appendix** ............................................................................................... 331
- **Notes** ........................................................................................................ 361
- **References** .............................................................................................. 377
- **Content Index** .......................................................................................... 407
- **Author Index** ........................................................................................... 415
Authors

Zhao Shouhui (PhD, University of Sydney) is Research Fellow at the Centre for Research in Pedagogy and Practice (CRPP), NIE at Nanyang Technological University, Singapore. A professional language teacher by training, Zhao has been teaching and researching in Chinese applied linguistics for the past 19 years at 7 universities in 5 countries beginning in 1988, when he completed his MA in Chinese Applied Linguistics at Renmin University of China (Beijing). Zhao is a member of All-China Society of Teaching Chinese as a Foreign Language and the Singapore Association of Applied Linguistics. He has published in the areas of linguistics and education. Zhao is the primary co-author of three dual-authored Chinese language textbooks and the co-complier of two dictionaries. He has also contributed a number of ephemeral pieces (e.g., prose, short stories, film reviews, etc.) to a range of literary publications, websites and broadcasts.

Richard B. Baldauf Jr. (PhD, University of Hawai‘i) is Associate Professor of TESOL in the School of Education at the University of Queensland and is on the Executive of the International Association of Applied Linguistics (AILA). He has published numerous articles in refereed journals and books. He is co-editor of *Language Planning and Education in Australasia and the South Pacific* (Multilingual Matters, 1990), principal researcher and editor for the *Viability of Low Candidature LOTE Courses in Universities* (DEET, 1995) and co-author with Robert B. Kaplan of *Language Planning from Practice to Theory* (Multilingual Matters, 1997) and *Language and Language-in-Education Planning in the
Pacific Basin (Kluwer, 2003). He is Executive Editor of the journal Current Issues in Language Planning and co-editor of six volumes in the Multilingual Matters Series, Language Planning and Policy.
Abbreviations

AI – Artificial Intelligence
GB 2312-80 – Basic Set of Standard Chinese Characters for Information Interchange-GB 2312-80
CCCF – Chinese Character Cultural Faction
CCP – Chinese Communist Party
CCSR – Commission of Chinese Script Reform
CIP – Chinese Information Processing
CTSC – Comprehensive Table of Standardized Characters
CWCC – Corpus of Whole Chinese Characters
FSS – First Simplification Scheme
FTVVF – First Table of Verified Variant Forms
GB – Guojia Biaozhan: National Standard
GLPFCC – General List of Print Fonts of Chinese Characters
GLSC – General List of Simplified Characters
IR – Information Retrieval
IT – Information Technology
LPers – Language Planners
LPP – Language Policy and Planning
NCLW – National Conference on Language Work
OCR – Optical Character Recognition
PRC – People’s Republic of China
RC – Rare Characters
RIAL – Research Institute of Applied Linguistics
SCLW – State Commission of Language Work
SR – Speech Recognition
SSS – Second Simplification Scheme
TSC – Table of Simplified Characters
Preface

One of the remarkable things about Chinese language policy over the millennia has been the power of the writing system to unite what are disparate if related spoken varieties. We have already published one book on PRC language policy in this series, a collection edited by Minglang Zhou that covered the full range of topics including the development of Putonghua (common speech), the status of minority languages, and some interesting chapters on the reform of the writing system. This last subject is of such complexity and importance to justify adding to the series a book going into more detail on developments in efforts to manage and cultivate the writing system in the last 50 years, taking into account the rapid growth of the Chinese economy and the technological developments associated with computers and the web.

It is important, we feel, to back up the common generalizations about language policy with detailed studies of specific topics, where one can observe at close hand the plans and activities of language managers, the problems they set out to solve, and their successes and failures. This is precisely what Zhao and Baldauf set out to do in this full account and analysis of the challenges met by recent efforts to adjust the Chinese writing system to new demands.

We would like to express our thanks to the authors and to the others who contributed to the production of this volume. In particular, we want to
thank the specialist reviewers who carefully read the manuscript, raising interesting points in a controversial field, and proposing some modifications that have made it a better book.

*Bernard Spolsky* and *Elana Shohamy*
Series Editors
March 2007
Foreword

The People’s Republic of China (PRC) is a growing, developing, and dynamic society that is moving to take its place as a major world power. Having the world’s largest population – i.e., human resource base – it is estimated that sometime in the next 20 years, China will move from having the third to having the world’s largest economy. Everyone knows that the world’s most populous country has been experiencing radical change over the past two decades, yet few people are aware that at the same time the Chinese writing system has also undergone a less prominent, but no less significant silent revolution. This development has led a number of language experts to predict that as the power and influence of China grows, so too will the spread of the Chinese language. Some commentators have even suggested Chinese will surpass English as the world’s most common *lingua franca*.

However, for the Chinese language to grow in power and influence, a number of hurdles first must be overcome, in particular the challenges of technology. The barriers to the growth and spread of the Chinese language are very different for spoken and written Chinese. This monograph focuses on the written variety as it traces the language policy and planning-related developments for Chinese characters, with particular emphasis on the post-1950 period of the PRC and the more recent challenges that technology, and particularly the use of Chinese on the world wide web have posed for the language. This volume examines the linguistic, cultural, political, and economic debate, the outcome of which could determine whether Chinese reaches its international language potential, and explores some possible language policy and planning directions for the future.
Historically, while Chinese language policy and planning developments can generally be characterized as evolutionary, during the last century, and particularly after the establishment of the People’s Republic of China in the 1950s, there have been forces, both reactionary (i.e., elements of traditional language and cultural purism) and revolutionary (i.e., ‘alphabetic’ or radical simplification schemes) at work in language policy and planning (LPP).

Over recent decades, this conflict of forces about the direction of character (hanzi) development has become intense, particularly since the milestone National Conference on Language Work held in 1986, where future possible simplification or radical treatment of the physical shape of hanzi and full-fledged Romanization were ruled out by the official working agenda of the conference. At the same time, to address the new challenges brought on by the advances that have occurred in communication technology, a more comprehensive and larger scale LPP program with the goal of ultimately overhauling the hanzi repertoire has been undertaken. The focus of this major operation has been the so called ‘Four Fixations’ or standardizations, aimed at settling the four most unstable attributes of hanzi, namely the number of hanzi, their ordering, shape and pronunciation. It is widely believed that once these four features are fixed or standardized and implemented through legally binding governmental measures like the Comprehensive Table of Standardized Characters (CTSC) and the Corpus of Whole Chinese Character (CWCC), some of China’s major language and communication problems will be on the way to being resolved. More importantly, and of more direct consequence, once they are required for new information technology (IT) products, it is argued there will be far-reaching effects and a significant impact on the linguistic life of the whole population. This may, eventually, in a silent and programmatic way, come to create a revolutionary impact on the visual communication life of hanzi users across the globe.

Given the importance of China and the potential of the Chinese language, an understanding of these issues is critical for not only our understanding of China, but also for second-language policies and their implementation as they relate to the teaching of Chinese around the world. This book summarizes struggle to modernize the Chinese writing system, with a particular focus on the emerging conflict between the world’s oldest surviving writing system and the latest advances that have occurred in technology.

The structure of the book has been developed around the problem-solution paradigm characteristic of much of classical language planning. The historical context and some discussion of basic features of hanzi are briefly reviewed in the Prologue for the benefit of readers not familiar with
this background. In Chapter 1 we provide a relatively detailed presentation of modern simplification history, starting with a summary of dynastic simplifications, followed by the three simplification processes that occurred in the period spanning the 1930s to the 1970s. These were the most important clinical treatments that ever have been performed on the hanzi in their historical development.

Chapter 2 briefly reviews past LPP experiences with a focus on the undesirable outcomes of the previous reforms from the perspective of the new historical context. The authors examine the chaotic situation that existed in the public’s use of hanzi in 1980s, and the prevailing confusion in academia about the evaluation of hanzi reform and hanzi’s future as it was affected by simplification and social transformation processes. The analysis of this transformation focuses on the impact of the cultural and socio-political dimensions of LPP activities. Chapter 3 sets out the problem. It concentrates on the conflicts between hanzi as a visual communication system and the needs of modern information technology in a digital society. A description of the enormous efforts Chinese scientists have undertaken to computerize hanzi is provided, illustrating for non-specialist readers the problems that Chinese IT experts and LPP professionals confront in tackling hanzi’s mechanical deficiencies. In particular, a substantial part of this section is devoted to the two key concepts: chongma and luanma. The former can be defined as accidental homophonic occurrence in hanzi computer input typically found when using phonetic input methods, the dominant scheme for entering hanzi into a computer. The latter refers to a string of unintelligible gibberish that occurs when hanzi are decoded by different encoding platforms or applications or during the transmission of Chinese encoded information over international computer networks. Chapters 2 and 3 provide the rationale for the argument that further LPP programs need to be carried out, thus forming the central focus for the subsequent chapters.

Chapter 4 discusses hanzi’s standardization, or the so-called four fixations (hanzi siding), as the means to overcome hanzi’s inherent limitations in the technological applications described in Chapter 2. Since the mid-1980s, standardization has been pinpointed as the societal solution to address a situation where hanzi have become a retarding factor in upgrading Chinese communication technology. The standardization issue has come to take on greater significance as it is increasingly essential for the enhancement of communication in Chinese. It may not be an exaggeration to say that it has now come to the stage where no major advances in Chinese computerization can be expected until the hanzi infrastructure is improved. In other words, if this issue is not resolved actively and quickly,
the question of hanzi’s survival as a mass writing form may be at stake. The major portion of this chapter provides a comprehensive examination of the contents of the four fixations and the various difficulties related to the standardization movement. Then, in the last section, an introduction is provided to the Comprehensive Table of Standardized Characters and the overhauling of the repertoire of all Chinese characters, the two major undertakings of official IT-oriented LPP agendas in the current context.

In Chapter 5, we set out the external enabling factors that are bound to influence and determine hanzi standardization work and future trends in a more general fashion. Seven influential factors or sociolinguistic dimensions are identified and discussed. The discussions of these seven dimensions are presented according to a hierarchy of the influence or power that individual factors can exert on hanzi reform development, aiming to identify possible cause-and-effect patterns that would influence the direction of future reform. Each dimension is critically discussed through either: a) generalizations from historical experience and inquiry, b) an examination of the current trends and future directions using analytic approaches, or c) international comparisons.

In Chapter 6, which summarizes the past and looks to the future, the external conditions for LPP are further examined from a sociolinguistic perspective. The first section focuses on the conflict between social development and technological advancement by analyzing the compatibility between standardization and societal conditions. To accommodate the new context, a pluricentric model of hanzi standardization is proposed. As standardization involves the choice between more than one alternative variant, the changes in linguistic behavior can only come about if there is a more open consensus building approach to LPP. Therefore, as part of an examination of the reform process, the last section of this chapter argues that it is also necessary to explore the possibility of establishing reliable and trustworthy LPP mechanisms that guarantee democracy and transparency for future hanzi reforms.

The concluding chapter reflects critically on some key issues concerning hanzi planning that do not conveniently fit into the previous chapters. These issues are arranged under three rubrics: 1) the renewed interest in Romanization, 2) how to deal with two kinds of problems arising from the two simplification schemes implemented in the 1950s and 1970s, i.e., a) a strong survival of some simplified characters from the Second Scheme (1977) among the public, and b) the failure to develop an automatic conversion between simplified and traditional characters as the result of oversimplification in the Table of Simplified Characters (1956, 1964), and 3) a section on a common script that considers the possibility of a cross-
national unifying standard among hanzi-using regions. This final issue is examined at two levels: the use of a set of real life standardized hanzi shared by all concerned hanzi-using polities and a unifying set of IT standards for encoding their use in cyberspace.

By tracing the developments and challenges for China through an examination of the Chinese writing system, in this book we contemplate some possible futures for the country. Put another way, given China’s phenomenal economic growth, and the beginnings of the accompanying radical political and social development, the question may well be asked whether the evolutionary process that has occurred over the last half century with Chinese characters will lead to reaction or revolution. Because of the world’s dependence on the communication of information, these LPP developments may well help to determine the place of Chinese language in the world. Thus, while the discussion of these issues could be seen as just another scholarly debate, we believe the broader implications of these questions deserve a wider audience.

Shouhui Zhao

21 April 2007

Richard B. Baldauf Jr.
1. THE ORIGINS OF HANZI

Due to the lack of authentic records, there is very little certainty about the earliest development of hanzi and only rough conclusions can be drawn. Archaeological findings have shown that the earliest embryonic forms of Chinese writing existed as far back as 4000 BCE on Neolithic pottery vessels. These earliest pictorial signs, referred to as Early Proto-Characters (e.g., Hook and Twitchett 1991), are believed to be closely linked to signs inscribed on animal bones and tortoise shells (jiaguwen) dating from the 13th century BCE. Although they were just the forerunners of Chinese ideograms, it shows that hanzi had developed for thousands of years before they were fully established as a writing system, adding more weight to the claim that hanzi are the world’s oldest continuing writing regimen. Jiaguwen are undisputedly regarded as the earliest relatively stabilized and systematically constructed form of Chinese writing. During the Shang dynasty (BCE 1600-1100), the major function of these inscriptions was to keep records of divination, rather than being a communication system for people’s everyday use. Some time after it was developed, the jiaguwen disappeared from history until they were accidentally rediscovered in 1899, and the complications surrounding this discovery have added an additional mysterious element to hanzi’s history. Before that, they went unrecognized for what they were and were called dragon-bone, an expensive ingredient in traditional medicine. About 150,000 pieces of jiaguwen
from the Shang and Zhou dynasties have been found, with the best example bearing 128 recognizable inscriptions.

Archaeological evidence of hanzi’s early development are so fragmentary that all theories about their origin are considered to be conjectural and speculative before jiaguwen’s time. However, even today there are at least four factors that keep Chinese characters shrouded in the mist of legend. First, there is hanzi’s long enduring history and imaginary explanations about its origin. Second, there is hanzi’s pictographic/ideographic structure. Third, the purposes of the early forms of Chinese characters are unclear (i.e., whether the oracle bone inscriptions were for divination purposes, or bronze inscriptions were for ancestor worship, the diviners were the only people who could give explanations of their meaning). Finally, there is the beauty of recondite Chinese calligraphy, which is essentially a very personal faculty that epitomizes spiritual expression and completely denies analysis.

It is difficult to say who was the first person to concoct these stories, but traditionally, there are four theories about how hanzi were formed. Cooper (1989: 129), who finds a number of languages were imbued with legendary stories about the creation of their writing system, concludes that “supernatural assistance in the invention of writing systems is a common claim”. The most unique and best known legend ascribed the invention of hanzi to Cangjie, the official recorder of the mythical figure Huang Di (Yellow Emperor). It is said that Cangjie created hanzi wholesale out of the inspiration provided by a bird’s foot print. The other three theories are that hanzi: a) originated in incising notches on wood; b) were related to the invention of the notched stick which is dated at in the 27th century BCE, and c) evolved from tying knots in strings. The latter explanation is a universal visible communication method, found in many other cultures across the world, e.g., Quipus (Quechua for ‘knot’) used by tribes living in South America prior to the Spanish conquest, which consists of cords of different colors or shapes knotted in a very complicated way (Cobarrubias 1983). The Bagua (Eight Trigram) hypothesis provides another possible fascinating explanation. In the Yinyang Bagua system, which was said to have been invented by the legendary Emperor Fu Xi, all natural phenomena in the universe can be represented through a notational system using only two primary mystic signs ‘–’ and ‘– –’, or yang and yin³, in an oracular arrangement. This very sophisticated method can be practically understood using the analogy of the binary digit system (Zheng 1988: 316), although whether there is any relationship between Yinyang Bagua and the modern computer is debatable.
2. **STRUCTURE EVOLUTION – FROM PICTURES TO SYMBOLS**

It is generally agreed that there have been five major changes in the physical shape of hanzi: i.e., Oracle (including turtle shell and animal) bone inscription (jiaguwen, 甲骨文), bronze inscription (jinwen, 金文), Seal Script (zhuanshu, 篆书), Clerical Script (lishu, 隶书) and Square Script (kaishu, 楷书) (see Figure P-1). In the following sections, we briefly trace the five iconic stages that have marked the periodization of hanzi’s physical changes during its historical development, illustrating the general trend of how hanzi have evolved from their original prototype to what we see today, and examining the conditions under which these changes have occurred.

![Figure P-1. The evolution of Chinese hanzi (Ager 2005)](image-url)
2.1 Oracle Bone Inscription (jiaguwen, 甲骨文).

Hanzi’s traceable history began from pottery inscriptions, which were found to have existed 6,000-7,000 years ago, and have been archaeologically verified as the earliest signs for which the genetic link with jiaguwen can be established. Out of 4,672 character types found and identified on jiaguwen, only 1,723 can be deciphered without dispute among jiaguwen scholars.

Because these characters are engraved by knife on turtle shells or big animal bones – written characters are pretty rare – jiaguwen are structured in thin and rugged lines and feature a long shape characterized by sharp shoulders, giving a strong impression of cutting. In comparison with its forerunners found on pottery, jiaguwen developed into a fuller linear form, but is still conspicuously pictorial and most are varied in a wide range of inconsistent shapes and unstable numbers of strokes. Constrained by the hard materials (knife and bone), only the outline of the object can be conveniently delineated. The elaborate drawing that often was found in pottery inscriptions was inevitably simplified and replaced by a few summary lines. Another striking feature of jiaguwen is that it is sometimes hard to tell if it is one or more than one character from the way it is written (see Figure P-2).

The underlying principles for later hanzi formation also took shape at this time, as jiaguwen were sufficient in number for those needed for basic expression, and were a movement away from drawing actual objects. For these reasons, jiaguwen are considered to represent the coming to maturity of hanzi as a writing system.

Figure P-2. Inscribed Oracle Bone (Crystal 2005). (The concise and readable inscription is a record of deciphering the natural shell texture or crack patterns after being toasted over fire.)
2.2 Bronze Inscription (jinwen, 金文)

Bronze inscription can be said to have directly evolved from jiaguwen. The biggest difference from jiaguwen is its rounded shape with thicker lines, as these characters were cast or incised on bronze ware including ritual utensils and containers, and sometimes on stone monuments. Despite bearing some similarities to jiaguwen, and occasionally using picture-like signs in the text, a further step was taken in moving away from drawing pictures. Furthermore, not only was the size of each individual character standardized, but the rugged, uneven strokes that had characterized jiaguwen were also transformed into smooth, regular ones. The shape of these hanzi was more graceful and symmetrical (see Figure P-3).

While jiaguwen were mainly used for divination purposes, bronze inscriptions were used in important ceremonies and the service of ancestor worship, so most contents (the longest extant text has 497 characters) functioned as a vehicle to perpetuated the patriarchal system, serving as a kind of materialization of social power and the status quo from a political perspective (Wang and Zou 1999). For the same reason, as Chiang Yee (1973: 43) observes, all characters are elaborately made to “be dignified and sublime, and designed to endure for many generations”, so bronze inscription has become the model for practicing calligraphy. A new academic subject – epigraphy (Jinshixue, the study of ancient inscriptions on cast bronze and carved in stone) – became a thriving academic study from the Song period (960-1279) lasting until the Qing period, when the scholars’ interest in hanzi switched to Jiaguxue as more oracle bone inscriptions were excavated.

![Figure P-3. Bronze inscription (Galambos 2005)](image)
During the later developmental phase of bronze inscription, a recorder at the court called Zhou (籀) synthesized the writing styles of bronze inscriptions, unifying to some extent the structure of these characters, making them easier to recognize and write. These redefined characters, typically represented by the so-called Stone Drum Inscription, depicting hunting expeditions, later came to be known as the Great Seal Script or Zhouwen. The great seal script Zhouwen was the major written communication means during the Zhou Dynasty (1100-221 BCE) and was used by all of the feudal states up to the Qin unification.

2.3 Small Seal Script\(^4\) (xiaozhuan, 小篆)

Although it continued from the previous period, and was used into the late Zhou period, particularly during the time of the Warring States (475-221 BCE), the great seal script was no longer maintained as a universal standard in a disunified country. It was very common for the same character to have numerous variants in different competing states. Therefore, when the Qin Empire completed unification, a standardized form of great seal script was adopted as the standard character set to overcome the chaotic orthographic situation caused by half a millennium of feudal wars. It was called small seal because it was simpler and had a smaller number of strokes (although this is not true for all characters). Furthermore, each character was made to occupy a smaller imaginary square. Physically, the descriptive pictorial and representational ideographic rationality is still a dominant feature of small seal characters, albeit being further weakened, i.e., the pursuit of representing real things began to give way to the abstract symbols. Through this human intervention, all the irregularities of the earlier forms were dropped, and the writing style was marked by a preference for symmetry, balanced structure and a gently-curving stroke.

Small seal script can be seen as a stepping-stone, connecting the earlier structure with the later evolution; it “break[s] away from pictographic forms, moving more towards linear and symbolic forms” (Yin and Rohsenow 1997: 35). But, by the small seal period, people were finding it increasingly impossible to express complex reality through near-real-life imitation, and a more productive method had to be found to represent oral expression. This development led to the combination of the existing ideographical signs and oral words, increasing semantic-phonetic characters to 80 percent of the total. This was a significant step towards creating a symbolic communication system. In this sense, although it gained official status by political coercion, the appearance and structure of small seal
Prologue

Clerical script has been generally regarded as being in accordance with the natural direction of script development.

2.4 Clerical Script (*lishu*, 隶书)

Clerical script was a parallel style that prevailed at the same time as small seal script, and it came into being out of the need for a script with a higher speed of execution, so clerical script is the pragmatic shorthand form of seal script. As its name suggests, it was first used by clerks, military book-keepers and other civilian personnel serving in the huge administrative system of the highly centralized Qin Empire.

As the state apparatuses grew more complex, the convenience and operational efficiency of communication became a major concern. Finding that small seal characters were still slow and laborious to write, Qin clerks simply ignored those unnecessarily strict formal pictorial elements needed to create an actual resemblance. Thus, for writing convenience, the circle was squared and rounded drawing lines were replaced by dots and straight lines, i.e., strokes. Through extensive use of strokes and square elements, the solid angles of *jiaguwen* were much softened and a great deal of movement and ease for writing was added. Hence the remarkable feature of this transformation is that the pictorial nature of *hanzi* was largely lost. This great change in developmental history has been referred to as ‘Li (cleric)-Change’ by Chinese philologists.

Clerical script was not only popular among public servants in the short-lived Qin Dynasty (221–207 BCE), but was also strongly promoted during the next dynasty – the Han (206 BCE-220 CE). One of the major factors for promoting clerical script might have been that the majority of the members of the ruling class during the initial period of the Han Dynasty were from a lower social stratum, therefore, they probably found themselves more comfortable with these pragmatic forms. The clerical script actually developed so quickly that it usurped the position of the small seal as the most popular style and achieved recognition among both ruling elites and plebeians, and by the Eastern Han (25-220 CE) period, the small seal script had lost favor even among the scholars. It was during this latter period that Xu Shen compiled the dictionary *Shuowen jiezi* (《说文解字》), in an attempt to resume the orthodox status of small seal script (Wang and Zou 1999). From this point of view, this dictionary can be seen as a result of the struggle between the pragmatists and the conservatives in the ruling elites.
2.5 Square Script (*kaishu*, 楷书)

As a further simplification reform of the clerical script, the square script took root during the Eastern Han period and became the mainstream style during the Northern and Southern dynasties (420-581), but it only was finalized during the Sui (581-618) and Tang (618-960) periods. While writing speed had been greatly increased by the change from small seal script to clerical script, it still was constrained by some structural conventions, such as the worm-head-like dot and swallow-tail-like ending stroke that posed obstacles to further efficiency. Square script reduced the remaining factors related to concrete objects to a few simple abstract lines. Its square structure and smooth level lines make it easier for common people to master, and partly due to the wide support lent by prestigious calligraphers, it quickly became the model style for learners to practice, hence its Chinese name, which literally means ‘Model Script’. *Hanzi*, at this stage, had deviated so much from its original forms that there would seem to be not much to associate it with the photographic image of its ancient origin. This last refinement created stability. For a thousand years after the introduction of square script, *hanzi* did not change significantly in its shape or style until the 1950s, when comprehensive simplification was first sanctioned by the government\(^5\).

Today, these four types of archaic characters that came before square script are not a practical medium for daily use except as graphic art, i.e., Chinese calligraphy, or more rarely, personal communication among idiosyncratic eccentric individuals (e.g., *Language Reform* 1974). However, with nationalism on the rise in the PRC since the 1990s, there has been a lot of interest from both LP professionals and academics from other disciplines in preventing these culturally charged characters from disappearing, and for integrating them into modern communication medium.

The graphical changes discussed in the previous sections are basically confined to the stroke and composition. However, *hanzi*’s shape also saw radical changes at the major developmental phases of the four typical styles, which in general can be illustrated by the shapes in Figure P-4\(^6\).

---

*Figure P-4. Historical changes in Hanzi shape*
2.6 Summary of Historical Development

In summary, the development of Chinese hanzi has followed a circle of standardization and simplification, and standardization characterized by evolution and reaction, i.e., the constant conflict between the desire to maintain the standard and the demand for operational efficiency which was facilitated by new writing instruments. For example, clerical script transformed the decorative elements of circular, curved and rounded lines of the small seal into the square, the polyangular and straight ones. This occurred to meet the requirement of writing quickly, as well as being the result of a change of the writing medium from hard and sharp implements to soft and pliable brushes. The clerical script transformation (Li-change) was a critical step in the process of developing from the imitation of primitive real things to an abstract and symbolic system. This development conformed to historical tendency, and the square script saw another sudden leap forward in construction and shape of hanzi that has remained basically unchanged up to the present. This extremely long period of total stabilization of the square script was no doubt due to the improvements made to the writing materials and means of writing, namely, the invention of the brush-pen, paper, and most critically, printing technology. This testifies to the proposition that technique innovation is the material foundation for any significant change in the Chinese writing system (see Table 7-2, Section 2.3, Chapter 7).

3. CHARACTERISTICS OF CHINESE HANZI

Hanzi are a culturally rich script system and have many unique traits. The following discussion focuses on four major factors that are believed to be relevant to the discussion in the subsequent chapters of this book. These are: abstraction, structural complexity, open-ended number and artistic features.

3.1 Abstraction: From Pictogram and Ideogram to Phonogram

Hanzi are generally known as a prime example of an ideographic writing system, but as Tang Lan (1965: 81) points out, an important tendency found in its general evolution “is the transformation from the original elaborate picture characters into purely symbolic signs which were
graphically composed of a narrow range of conventional strokes”. But the process, by which Chinese script evolved from primitive pictures to characters, has taken a long time. Even with the emergence of greater abstraction, the pictographic components have remained the basic building blocks that form hanzi. As hanzi have progressively become more abstract, the potential to create new characters has greatly increased.

As the script matured, it became simpler and progressively began to lose some of its pictographic qualities; devices other than pictures were eventually formulated in order to represent concepts and abstract terms which could not otherwise be represented graphically (Gao 2000: 75).

Consequently, parallel to this process, phoneticization began to emerge. Developing from an ideographic structure to a phonetic compound is generally seen as another marked tendency. Hanzi’s phonetic tendency was so evident as to lead some researchers to believe that Chinese script “is phonetic, not an ideological system of writing” (DeFrancis 1984b: 133). However, instead of outright phoneticization (graphic representation of units of sounds in a language), phonetic factors in hanzi, as a writing system, were largely manifested in two ways. On the one hand, there was the development of a phonetic compound in a semantic-phonetic character. The semantic-phonetic method was so robust that it is believed that by the time the Shuowen jiezi dictionary was compiled, nearly 80 percent of hanzi were composed by semantic-phonetic characters. On the other hand, there was an increase in the occurrence of homophony, i.e., a number of very frequently used characters served as an inventory of possible syllables to represent characters with the same pronunciation but different semantic meanings.

This method is in essence a homophonous substitution, or the extension of the traditional ‘Jiajie’ (phonetic loan) method. The Jiajie method allows people to employ a small number of the most used simple hanzi as basic syllables to represent other concepts or words with the same or similar sounds, disregarding these characters’ original meaning. These relatively small numbers of the commonest characters are used purely to vocally annotate the pronunciation for another unknown character and are called syllabic hanzi. The increased use of syllabic characters, only for their phonetic value in the text, is seen as a big step in the direction of a phonographic script (Li 1934). As a consequence, the number of homophonous characters or words have greatly increased, which in turn has become a driving force leading to word di-syllabication and poly-syllabication. Wang Fengyang (1989) points out, a developing trend in di-syllabization and poly-syllabication of Chinese vocabulary is the manifestation of a
phonographic tendency, and it provides the material basis for hanzi
developing from pictograms and ideograms to phonograms, because only
disyllabic and polysyllabic words or compound words can tolerate homophone
ous substitution without causing semantic misunderstanding.

A number of linguists insist on viewing the development of writing in
terms of two stages, “language tends to undergo a linear process of evolu-
tion from the ideographical writing to the phonetic system as man becomes
more sophisticated in his knowledge about the use of language as code for
communication” (Jackson and T’sou 1979: 81). In the first stage, the picto-
graphic symbols, which Gelb (1963: 59) called “the forerunners of writing”,
are descriptive and representative. At the second stage, a kind of device
like the phonetic elements was needed, as the previous pictographic and
ideographic methods were insufficient to suitably and competently repre-
sent increasingly complex speech. Pictographic forms, most of which were
carried over from the first stage, broke away from the real-life objects and
began to orally represent the concept. Woon (1987: 270) observes, “[a]lth-
ough many things happened in the structural evolution of Chinese charac-
ters, the most important development was that the writing system began to

![Figure P-5. The structural complexity of Chinese characters (Ager 2005)](attachment:image)
tend towards phoneticization and that this led to the development of the phonetic compound.” The unique aspect of this change process is that it suddenly stopped, instead of developing into a full-fledged phonetic orthography, a phenomenon that still puzzles modern hanzi specialists and linguists.

3.2 Structural Complexity

Ostensibly, complex strokes make hanzi a time-consuming writing system (see Figure P-5). For 6,763 characters (39 are non-character graphic signs) in the Basic Set of Standard Chinese Characters for Information Interchange (henceforth, GB 2312-80; GB stands for Guojia Biaozhun: National Standard), the average number of strokes is 10.665; characters with more than nine strokes account for nearly 70 percent of the total (Su 2001b). More problematically, although strokes in a character are written according to fixed rules, the shapes of components are diverse and notorious for their inconsistency. As a consequence of these factors, and others, it is generally agreed that acquisition of writing skills requires a minimum of five years’ formal schooling, and as many as ten years to achieve durable literacy. Furthermore, there are many uncertainties when describing characters, such as the definition of the component, the number of radicals, what to call radicals and what the differences are between radicals and components, creating a high degree of complexity. (see Section 2.2, Chapter 4) Hanzi’s complexity in composition can be understood through three basic terms – strokes, radicals, components – discussed in the following sections.

3.2.1 Strokes (bihua, 笔画)

Strokes are the smallest possible structural units used to form a character. When writing Chinese characters, each individual stroke can be seen as the continuous contact of one’s pen with the paper. Bihua is actually a term that came into being after lishu and kaishu, because, as previously mentioned, before Li-change, a character’s stroke was not clear-cut as most of time the writing tools were not the pen (bi). A character may consist of between 1 and 64 strokes. The stroke is the universal basis for any type of writing system. An alphabetic script is also comprised of letters that are composed of strokes, but hanzi are written using a much more diversified and recurrent array of strokes in which the
shape varies for different types of strokes. Out of the approximately 30 strokes, only 8 are considered basic ones and all others are their variants. For example, a hook stroke can be a left hook or a right hook, a straight hook or a bent hook. Although the stroke’s thickness and length do not impinge on the meaning of a hanzi, a number of factors are taken into account in forming a character correctly: stroke shape, the relationship among the strokes, stroke ordering, and stroke number.

3.2.2 Radicals (bushou, 部首) and Compounds (pianpang, 偏旁)

It is the radical⁷, however, not the strokes, that is basic to the hanzi writing system. The radical, the head under which characters have been classified in dictionaries, is the minimum meaningful composing element. Characters may be classified into two categories in terms of structural complexity: simple characters and compound characters. Simple characters, accounting for about four percent of the total, are arranged as a compact integral and are not further divisible into distinct components. Many radicals are themselves simple single-unit characters, thus they can only be seen as a radical when used as a composing element in compound characters.

The radicals that are used in the dictionary as a classifying index to group the characters are bushou, which are just graphic parts of characters used as headings in dictionaries without heuristic values. Characters containing the same radical are arranged under the same bushou in ascending order by number of residual strokes. Overall, radicals can range in number from 200 to 600, according to the approach taken in different dictionaries. One of the biggest problems with a radical is that there is no standardized way to verbally describe it, and this is increasingly becoming a daily problem as computers rapidly become more common (see Section 2.4, Chapter 4).

Another term related to radical is compound (Pianpang), which is another structural element that makes up a character. There are two kinds of basic compounds in characters: a semantic compound (yipang – 义旁), which more often than not overlaps with radicals, that specifies the meaning category of the whole character, and the phonetic compound (sheng pang – 声旁), which signals something about the pronunciation. However, not only semantic compound, but a number of phonetic compounds also have separate lexical status. Semantic-phonetic compounding is the most dominant method of character formation.