

THE PEDAGOGICAL *TIPITAKA*: OER & THE THREE BASKETS OF ANCIENT LANGUAGE INSTRUCTION

Todd B. KRAUSE; Hans C. BOAS; Danny LAW

University of Texas at Austin

bobtodd@utexas.edu; hcb@mail.utexas.edu; dannylaw@austin.utexas.edu

Abstract:

Ancient languages present a unique teaching challenge: for spoken languages, common pedagogy recommends engaging students via dialogue; for ancient languages, no speakers survive with whom to practise. This paper highlights how the Linguistics Research Center at the University of Texas at Austin has approached this challenge by creating the Early Indo-European OnLine (EIEOL) collection, an online educational resource whose lesson series present early languages directly through original, unsimplified ancient texts. Currently accessed by over 20,000 users per month, EIEOL spans 18 languages, from Greek and Latin to Old Church Slavonic, Sanskrit, and other important languages of ancient Asia such as Hittite, Classical Armenian, Avestan, and Tocharian. Each series presents extensively annotated excerpts of original texts in the target language, with accompanying modules explaining grammar and context. The text-centred approach affords learners a direct path to understanding that suits a variety of experience levels and minimises the conceptual grammatical apparatus necessary to begin interpreting original texts. This format fosters theoretical flexibility, adaptable to different approaches and grammatical descriptions of ancient languages. It is also useful for languages whose grammatical structures have shifted dramatically over their history, like Tocharian, or remain hotly debated or under-described by experts. Finally, it facilitates applications to typologically diverse languages and language families, with early Mesoamerican, Semitic, and Sino-Tibetan language series already under development. The EIEOL infrastructure therefore provides a robust platform for free, text-centred, self-paced introductions to ancient languages from a variety of language families.

Keywords: Sanskrit, Tocharian, Chinese, Classical Chinese, Digital Humanities, Language Pedagogy, OER, MVC, Web Design, Historical Linguistics

1. Introduction

Open educational resources (OERs) for ancient language study, particularly in an online context, exhibit both a wealth of innovative pedagogical insights and a simultaneous lack of clarity as to purpose or goals. Online pedagogy for modern languages naturally shares a clarity of purpose, namely communication. The generally larger pool of students and available resources leads to a “more is more” approach – more texts, more audio, more video, more role-playing – that seeks to emulate an immersive linguistic and cultural experience as OERs blend with hybrid learning (cf. Blyth 2012). A similar approach for ancient language instruction, however, quickly runs up against barriers: adding “more” encounters obstacles in the form of limited corpora, and goals of faithfulness to “the original” often disincentivise creation of “new” ancient language material.¹ Moreover, communication no longer represents an obvious goal: some ancient languages enjoy large corpora and have maintained a traditional scholarly or liturgical role even as the spoken language has undergone language death,² while others possess more limited or fragmentary corpora that provide little clear guidance to the “conversational” patterns necessary to support an immersive, communicative approach. In such instances, goals can shift to accurate evaluation and detailed understanding of unaltered original documents. How then should OERs approach ancient languages to foster such detailed understanding?

Of course, the particular difficulties of ancient language pedagogy far predate online OERs. We should therefore ask: should online OERs for ancient languages simply recapitulate their offline forebears? If so, which methodologies should they recapitulate? Is there a one-size-fits-all solution to deciding on online presentation formats? Any answer to such questions must simultaneously address the issue of how to measure the success of any given online methodology; however, this caveat likewise remains valid, though unevenly or infrequently applied, for related offline methodologies.³ Fundamental motivations for any particular online design should probably

¹ Cf., e.g., Köntges *et al.* (2017) for approaches to OERs in the context of Classical Latin and Greek, and Bird *et al.* (2022) for approaches to classical topics more broadly.

² For example, Latin has maintained a continuous, though restricted, spoken tradition over the centuries. With the advent of social media connecting disparate practitioners, this has seen a resurgence in recent years. The same holds true to varying degrees for Classical Greek and Sanskrit (Krause 2019).

³ Cf. e.g. Sato & Loewen (2019) for efforts to improve evidentiary support for a variety of pedagogical techniques.

include some basic criteria, e.g. clarity of exposition, flexibility of deployment, adaptability to different learning styles, inclusivity and ease of access, among others.

The Linguistics Research Center (LRC)⁴ at the University of Texas at Austin has been experimenting with how to present ancient languages online to a general audience since the early days of the Internet. Over the last two decades, the LRC has continued to expand the Early Indo-European OnLine (EIEOL)⁵ collection of lesson series, providing introductions to ancient languages whose speakers spanned most of Europe as well as parts of the Middle East, Central Asia, and the Indian subcontinent. These lesson series chose a format derived from early studies of language acquisition carried out by researchers in linguistics and language pedagogy (Lehmann 2003).

This approach emerged in a period of early use of the Internet as an academic space. As web design frameworks have matured over the years, however, such design decisions have needed to be revisited. While EIEOL lesson series have retained the basic features of the original format, the last several years have seen upgrades to the underlying website production system that allow for more flexible deployment of the underlying content. Now incorporating the popular model-view-controller (MVC) architecture over an underlying relational database, the website architecture supports deployment for a range of site configurations and pedagogically oriented interfaces.⁶ This increased flexibility allows us to revisit the question of what viable options exist for pedagogically oriented presentations of ancient languages in the context of online language resources, and whether those options can be adapted and optimised for particular languages and language learning traditions.

Below we investigate this question by outlining three ‘baskets’ (Pali: *tipitaka*, the term used for the traditional threefold division of early Buddhist liturgy) of pedagogical approaches to teaching ancient languages through written materials. We exemplify these approaches by considering three distinctive methods of

⁴ <https://liberalarts.utexas.edu/lrc/>

⁵ <https://lrc.la.utexas.edu/eieol>

⁶ The change of infrastructure initially served to support lesson creation. Multiple authors can now collaborate simultaneously on a lesson series in a fully online environment similar to, though somewhat simpler than, Google Drive. Moreover, the lesson series can now be edited, updated, published, and maintained without explicit recompilation by the webmaster, ensuring users continually have access to the most up-to-date versions of the online materials. This same infrastructure now allows us to offer responsive design on the user side.

instruction that are common for three ancient languages of Asia, namely Sanskrit, Tocharian, and Chinese.

2. Pedagogical approaches

The three ‘baskets’ of common approaches to ancient language pedagogy that we explore here could be termed *theoretical*, *historical*, and *commentarial*. To make these classifications more concrete, we have chosen three exemplars, languages whose traditional pedagogies exemplify each of these three approaches: Sanskrit, Tocharian, and Classical Chinese. We discuss these exemplars, and the approaches they typify, in greater detail below with respect to 1) the corpus of primary texts available for study and 2) the common pedagogical approaches to the grammar of the language.

Certainly, any individual scholar or student may have a personal preference for the particular pedagogical approach they find most effective for teaching or learning a particular language, or languages in general. Careful studies of L2 acquisition might further support the validity of such preferences. But through consideration of the historical developments and cultural contexts in which certain languages have been studied, we can appreciate the origins of certain pedagogical approaches and their adequacy or appropriateness for the task of transmitting knowledge of particular languages at particular times. The description below aims to illuminate aspects of the origins of the respective pedagogical approaches, motivating a discussion of online resource design that seeks to preserve such approaches where traditional, and expand them where appropriate.

2.1 Sanskrit

2.1.1 Corpus

Sanskrit, an ancient language of India, comprises a vast textual corpus. This corpus includes numerous manuscripts which, due to the deleterious effects of harsh heat and humidity upon perishable palm-leaf and birch-bark pages, frequently date only to the middle of the second millennium. However, their contents often exhibit linguistic features that clearly predate the manuscripts by several centuries or even millennia. In the case of the Vedic literature, scholars

believe that portions of the text, preserved through a rigorous process of oral recitation and memorisation, date back to the latter part of the second millennium BCE (Jamison 2008a). During the period between this dawn of identifiable Vedic Sanskrit composition and the earliest Vedic manuscripts of roughly the 1100s CE (Witzel & Wu 2018), Sanskrit remained in continuous use, even though at some point, perhaps as early as the middle of the first millennium BCE with the compilation of Pāṇini's exhaustive Sanskrit grammar, the language itself underwent language death: the Vedic language ceased to be spoken as a native tongue and persisted primarily as a liturgical language, while the version of Sanskrit outlined by Pāṇini remained, not as a mother tongue, but as a scholarly *lingua franca*.

Nevertheless, this phase of scholarly and literary use of Sanskrit gave rise to a vast literature that spans a wide range of genres: epic poetry; dramatic plays; fables; mathematical, astrological, and astronomical treatises; grammatical explications; philosophical explorations; legal codes; and medical texts, to name but a few. While many of these genres serve a more academic or scholarly purpose, some aim to treat everyday topics. This is especially true of the works of drama and of the fables (cf., e.g. Johnson 1847; Kāle 1961). Though interspersed with poetic elements and artistic flourishes, these works on occasion also depict characters speaking plainly and in simple terms. That is, the Sanskrit corpus over the centuries shows a variety of registers that illustrate a language used by kings, courtiers, priests, and scholars that could also accommodate more colloquial needs. The dramas provide a particularly interesting venue for observing the sociolinguistic dynamics of the period: attendants of, and servants to, characters of a higher social stratum would themselves often not speak in Sanskrit, but in the so-called dramatic Prakrits, their own contemporaneous regional Middle Indic languages, which themselves developed into literary languages (Kāle 1961; Woolner 1917). Commentaries on these works then rendered these utterances *back* into Sanskrit for those unfamiliar with that particular popular language (Kāle 1961).

While we must not overlook the fact that these dramatic Prakrits represent “common speech” only as rendered through the lens of a courtly composer, the mix of languages in a given dramatic scene, and the re-rendering of common speech into the courtly medium of Sanskrit, testify to the wide range of everyday functions that Sanskrit aimed to serve. At the same time, such dramatic works

illustrate the linguistic milieu of Sanskrit within India during the Middle Indic period (Woolner 1917). Though these local popular languages can often trace their linguistic heritage to the language of the Vedic period and common parent of this period's scholarly Sanskrit, they had undergone such transformation over the intervening centuries that they are often practically unintelligible to a speaker only familiar with Sanskrit. Conversely, such a variety of languages across regions simultaneously supported the utility of Sanskrit as a courtly, administrative, and literary medium of communication for ministers, bureaucrats, and artists from different regions and backgrounds.

2.1.2 Pedagogy

Scholars occasionally argue that Pāṇini's codification of the grammatical system of the Vedic language effectively "froze" Sanskrit because, roughly speaking, the importance of the language in its religious function tended to support a perspective by which Pāṇini's originally descriptive grammar came to be viewed as prescriptive. While perhaps only a simplification of the actual processes involved, the subsequent period occasioned a divergence between a refined stratum of educated and liturgical Sanskrit and the more practical language by which speakers conducted matters of everyday life (Woolner 1917; Jamison 2008b). In this bifurcated or diglossic state, the scholarly stratum of Sanskrit nevertheless evolved, and its speakers remained attuned to its grammatical structure at each stage. Sanskrit maintains a long tradition of incisive commentary ranging over a variety of grammatical topics, though with particular acuity in matters of phonology and morphology. Such commentary dates back to the immediately post-Vedic period, contemporaneous and in conversation with a parallel scholarly and religious spoken tradition (Staal 1972).

As the understanding of Sanskrit made its way into the educated circles of Europe, at the end of the 1700s and the early decades of the 1800s, the native tradition of a top-down, theoretical approach to the language influenced the structure of instructional treatises on the language in Europe itself, where Latin and sometimes Greek had come to fill a similar role among the most formally educated stratum of the continent's population (Staal 1972). We see in initial works by Monier-Williams and Benfey in the early decades of the 1800s (over several editions of their treatises) the beginnings of the later trend to write a scientific descriptive grammar of the Sanskrit language, and then accompany this

with a chrestomathy or reader containing text selections by which to practise the principles learned and familiarise oneself with the range of genres encountered in Sanskrit literature (Monier-Williams 1864; Benfey 1852, 1853).

These works displayed an arrangement typical of the period: they begin with a treatment of the sound system, discuss sound rules in the context of internal and external sandhi, and from there move on to treating nominal and adjectival morphology, pronominal morphology, verbal morphology, and so on. That is, they break down the morphology of the language by word class and discuss one class at a time, in relative isolation from other word classes. These works relegate syntax to a fairly short chapter toward the end. Such placement and the general sparseness of the syntactic discussions were influenced on the one hand by Sanskrit's relatively ornate morphology and its expression of grammatical relationships, and on the other hand by the underlying assumption that Sanskrit syntax was less ornate, or more straightforward, than that of the Latin and Greek with which the authors themselves and their assumed readers were already familiar (cf., e.g. Delbrück 1976; Speijer 1886).

This same division of content also made its way across the Atlantic, finding perhaps its most canonical expression in Whitney's *Sanskrit Grammar* (Whitney 1889). This too served as what we might call a reference grammar: a thorough description of the overall grammar of the language, with discussion broken down into somewhat logically self-contained sections based on grammatical categories. This text, however, made no pretence of including reading selections, a task which ultimately fell to Charles Rockwell Lanman, whose Sanskrit reader forms the necessary pedagogical companion to Whitney's grammar (Lanman 1884). Lanman's reader begins with numerous text selections from a range of genres, parallel to the chrestomathies of Europe (e.g. Lassen 1865). It follows these selections with an erudite glossary, similar to Benfey's own, listing not only the individual words encountered in the readings, but frequently the associated etymological cognates from classical languages like Greek and Latin.

Shortly thereafter a different pedagogical approach emerged. Based on notes by Georg Bühler already in circulation in Europe by the early 1880s, Edward Delavan Perry issued his famous *A Sanskrit Primer*, which provides a helpful hand in walking students through Whitney's comprehensive grammar (Perry 1885). Rather than throwing students "into the deep end" with original texts and simple references to relevant paragraphs of Whitney's grammar, Perry adopts a

more supportive approach: he divides the grammar into lessons, each lesson centred on certain essential points of grammar (always referring back to Whitney) and accompanied by practice sentences or simplified passages with attendant vocabulary to gradually build up familiarity with grammatical constructs and foundational terminology. Such a trend was already underway in Europe by the 1870s, where terse summaries of the essentials of Sanskrit grammar, like Stenzler's, already included "lessons" or "practice examples" (*Übungsbeispiele*) at the end of the books; these consisted of a proposed list of paragraphs to consult in the grammar, coupled with some simple sentences for illustration (Stenzler & Pischel 2016). However, works like Perry's finally inverted this structure, making the lessons and exercises the essential backbone of the book. With this restructuring of pedagogical approaches, Perry's work positioned itself as a precursor to reading unsimplified texts like those in Lanman's reader, and to a more detailed study of the grammar through references like Whitney.

Perry's remains the common approach to Sanskrit study in the West to this day. Naturally, the education systems have changed around these works, as has student preparation. Perry could assume his students' familiarity with the grammar of Latin, and as a consequence his explanations of case functions and features like locative absolutes often appear as one or two terse lines calling students' attention to parallel features of Latin grammar. Contemporary treatments, such as Ruppel's, cannot assume the same student preparation and therefore devote more time to introducing and teaching the structures of the language (Ruppel 2017, 2021). Nevertheless, the basic lesson format, with grammar rules followed by vocabulary, exercises, and simplified excerpts to practise the newly acquired features, remains largely intact. Moreover, the aptness of this approach derives less from modern assessments of pedagogical methodology, and more from its clear lineage in a rules-based approach to understanding the language, codified first in an indigenous oral tradition and later transferred to written presentations allowing wider diffusion.

2.2 Tocharian

2.2.1 Corpus

Modern scholarly access to Tocharian differs radically from that of Sanskrit. The extant Tocharian corpus is relatively small (Pinault 1992; Malzahn 2007,

2017; Peyrot 2008). It consists chiefly of written fragments found in caves and monastic retreats scattered throughout the Tarim Basin on the northern arm of the Silk Road in the Xinjiang province of modern China. The documents encountered early in the history of Tocharian decipherment preserved almost exclusively Buddhist religious texts: these now contain both canonical literature, such as texts on monastic discipline and religious philosophy and metaphysics; as well as para-canonical texts including Buddhist styles of poetry, narrative, and drama. Beyond this religious core, technical genres include calendrical material, texts on magic and divination, grammatical treatises and word lists. More personal documents include confessions, donations, blessings, and a love poem. The corpus, however, comprises fewer than 15,000 small fragments often consisting of only a few lines each (Malzahn 2018). This complicates, though by no means precludes, scholars' ability to draw conclusions about the characteristics of extended narrative and dialogue beyond these written styles (Peyrot 2008, Malzahn 2017).

The Tocharian-speaking communities evidently formed a crucial link in the transmission of Buddhism from northern India to Central Asia and farther on to China and the rest of East Asia. This would help explain why the Tocharian documents frequently appear in monastic libraries near other Buddhist texts in a variety of regional languages. Some texts – mostly cave graffiti, monastic records, and receipts – contain remnants of writings composed originally in Tocharian. But the majority comprise Buddhist texts translated into Tocharian from neighbouring languages, most often Sanskrit. The fact that we often have versions of these same Buddhist texts preserved in a language other than Tocharian allows scholars to identify fragments as part of one or another Buddhist treatise, and to begin the painstaking process of ordering them into a patchwork representation of a presumably continuous Tocharian text. The Tocharian corpus is rounded out by a relatively small collection of border passes: documents written on wood and carried by merchants, denoting the wares transported in desert caravans traversing the contemporary regional powers of Central and East Asia.

2.2.2 Pedagogy

The early Tocharian finds harken back to European expeditions to Central Asia in the late 1800s (Pinault 1992). Written in a form of the Brahmi script in common use across a wide swathe of Central Asia, these texts revealed their

importance in the following years, as scholars sifting carefully through the fragments realised they recorded languages that were altogether different from the texts written in Middle Persian, Sogdian, Khotanese, and other Iranian languages.

Sieg and Siegling were among the first to systematise the study of the language itself and the translation of the early fragments (Sieg *et al.* 1931). Their initial grammar falls naturally in line with the method of engagement with the texts up to that point: the authors carefully described the sound system and morphology of the language as they encountered it, dividing the grammatical exposition into logical segments corresponding to the various principal parts of speech. This parallels the style of Monier-Williams, Benfey, and Whitney in describing the grammar of Sanskrit, i.e. a reference grammar. Over the course of this process, scholars came to realise that the fragments preserved two distinct but related languages, termed Tocharian A and B (Poucha 1955, 1956). Sieg and Siegling's early grammatical description and translations focused on Tocharian A, describing Tocharian B only in certain points of contrast. Soon, however, scholars such as Krause focused more closely on the description of specific features of Tocharian B (Krause 1952). After a short treatise on the structure of verbs in Tocharian B, Krause collaborated with Thomas on the production of the *Tocharisches Elementarbuch* (Krause & Thomas 1960, 1964). This two-volume work is reminiscent of Benfey's work on Sanskrit: the first volume comprised a reference grammar, the second a chrestomathy with unsimplified reading selections and a glossary. In one and the same work they treated the two languages simultaneously. But importantly, they also provided a historical derivation of the phonology and morphology, demonstrating the relationships to other languages of the Indo-European family. Inasmuch as the language was still in the process of being understood, so was its grammar, and to justify these new interpretations the scholars relied on comparisons to other Indo-European languages. In this way the mode of explication took on a slightly different character from many of the Sanskrit treatises discussed above, less explaining the language and more deriving it from its historical origins.

The *Elementarbuch* still remains the standard reference for Tocharian A and B. There have been notable pedagogical advancements beyond the historical reference grammar with chrestomathy, in particular with Pinault's *Chrestomathie tokharienne* (Pinault 2008). This work nevertheless begins with a historically

oriented reference grammar to introduce the workings of the language. But rather than following the exposition with text selections and a separate glossary, Pinault's introduction provides unsimplified text selections with word-by-word glosses, grammatical and historical notes, and a continuous translation. This greatly simplifies the student's work in correlating Tocharian structures with their meaning in translation, but it still physically and logically separates reading texts from learning the associated grammar.

Only recently has Tocharian benefitted from an introduction in the style of Perry's *Sanskrit Primer*: Michael Weiss's *Kuśiññe Kantwo: Elementary Lessons in Tocharian B, with Exercises, Vocabulary, and Notes on Historical Grammar* (2022). As the title suggests, the work builds up the grammar bit by bit for the student, providing examples and practice readings in each chapter. Thus the trajectory of introductions to the Tocharian languages seems to parallel that of Sanskrit, though it remains at a very different stage along the path. Moreover, Sanskrit grammatical exposition can trace a direct line back to a native tradition with which it has maintained contact. Tocharian does not draw on such a tradition, and the scholarly framework of viewing the language through its historical development and relation to other Indo-European languages remains accessible everywhere within Weiss's introduction. This historical mode of presentation is by no means necessary; but its use seems highly appropriate for teaching a language whose grammar scholars continue to elucidate and whose texts have traditionally been deciphered with reference to other languages in its historical and social environs.

2.3 Chinese

2.3.1 *Corpus*

As with Sanskrit, the Chinese corpus is quite extensive and encompasses a roughly similar span of time, from a millennium or more BCE to roughly the 14th or 15th century CE by a conservative estimate (Hartman 1998; Peyraube 1999). The texts exhibit a wide range of genres, including histories, poetry and song, military treatises, philosophical discourses, didactic materials, diplomatic and administrative documents, dramatic and fictional works, among others. The earliest remnants of writing appear as inscriptions on bones or bronze, but early

writing also appears in artwork, craftwork, and in literary and administrative contexts on varied and refined paper products.

As in India and elsewhere, the language tied to this extensive collection of texts largely served a highly educated stratum of society involved in, closely allied to, or supported by the functioning of government over successive dynasties (Hartman 1998). The writing system itself encapsulates the profound level of education needed to engage with this literature over its long history. Specifically, the system stands on a roughly logographic foundation that has expanded and adapted itself to new needs and contexts over centuries. Though it originally depicted words with small pictures, it later expanded by adapting existing pictures to represent homonyms, or near-homonyms, even of dramatically different meanings. Over time, as curves shifted to angled corners, the characters slowly became divorced from the original visual depictions of worldly objects to a more abstract, linguistically-attuned symbology. Mastery of this system became a major focus of education in and of itself, reinforcing the role of the classical language in serving a highly educated stratum of society (Wieger 1965; Qiu 2000; Dong 2020).

2.3.2 Pedagogy

While the Chinese corpus exhibits a vast concurrent tradition of linguistic commentary, its character differs quite substantially from the Indic tradition, largely by virtue of its focus on the specifics of the writing system (Qiu 2000). A central strain of this linguistic commentary lies in broadly lexicographical works compiled by numerous scholars as the Chinese corpus expanded and evolved (Yong & Peng 2008). Some early references suggest that dictionaries had already begun to appear by roughly 800 BCE (Hartman 1998; Mair 1998). These lexicographical works fall into three main categories.

Early dictionaries fall into the *xùngǔ* (訓詁, “exegesis” or “philology”) category (Mair 1998). They principally ordered elements semantically, e.g. collecting terminology for kinship, architecture, geography, etc. Within categories, words were often grouped by synonyms. For example in the *Ēryǎ* (爾雅 “Approaching Elegance” or “Ready Guide”) of the 3rd century BCE, initial sections contain commentary or exegesis on verbal phrases and particles in earlier classical texts, while later sections group further terms into 19 semantic categories (Mair 1998). This system requires the user to know or guess the rough meaning of a character before being able to look it up.

The *Shuōwèn Jiězì* (說文解字 “Explanation of Simple and Analysis of Compound Characters”) from the 2nd century CE inaugurated the class of *wénzì* (文字, “script” or “grammatology”) works, distinguishing two major character types: 文 (*wén*, a “simple figure”), in which a single drawing represents an object or idea; or 字 (*zì*, a “compound character”), which combines several symbols to assign meaning (Mair 1998). This work introduced a distinction among 形 (*xíng*, “shape” or “structure”), 音 (*yīn*, “sound”), and 義 (*yì*, “meaning”). The organisation centred on characters’ graphical characteristics, using divisions based on 540 部首 (*bù shǒu*, “section headers”). Translated commonly as ‘radicals’, suggesting minimal or essential elements of character composition, the term was originally meant to be neither elemental nor exhaustive, but rather representative. Only in the early 17th century CE work *Zìhuì* (字彙 “Character Glossary”) did the number of ‘radicals’ reduce to 214. The *Zìhuì* also introduced stroke counting: ordering characters within the radical groupings by the number of strokes added (Mair 1998).

The 7th century CE saw the introduction of rhyming dictionaries with the *Qiēyùn* (切韻 “Cutting [i.e. Writing] Rhymes”). This presented the first phonological ordering of elements, a result of “the enhanced phonological awareness that developed in China after the advent of Buddhism and the elaborate Indian linguistic science that came in its wake” (Mair 1998: 168). Users required a deep familiarity with rhymes to look up elements. Works organised according to this style belong to a grouping of the grammatical tradition known as *yīnyùn* (音韻, literally “sounds and rhymes”, but in the sense of syllable initials and finals, or roughly “phonology”; Mair 1998).

These major and long-lived categories of linguistic writings focused on the lexicon and its interaction with the writing system, dedicating relatively little discussion to the morphological and syntactic aspects of the language. This derived in part from the bureaucratic and political function the language and literature came to fill. Though Chinese had many regional variants, the imperial administration required a common language to carry out its functions. To serve a growing bureaucratic class, education focused on ensuring the ability to produce well-formed and elegant administrative documents (Hartman 1998). Instruction within this system centred on memorizing the classical texts. Since the pronunciation of early texts changed over time and often differed from a student’s own regional dialect, the bureaucratic “Mandarin” rose as a standard. Students learned to read and recite classical texts, also memorizing the relevant sections of commentary. Rules of grammar received little emphasis, since the grammatical

patterns were implicit in the memorised material, and the student merely had to produce compositions along the same lines (Hartman 1998).

In his five-volume *Cursus Litteraturae Sinicae Neo-Missionariis Accommodatus*, published in the late 1800s, the Italian Jesuit missionary P. Angelo Zottoli presents a primarily European audience with a synopsis of this traditional education, beginning with a text in traditional Chinese characters with accompanying transliteration in Roman letters, a word-by-word translation, and a character-by-character commentary in Latin (Zottoli 1879–1882). The commentary explains each character’s pronunciation, meaning, and function, and includes more general commentary on literature, culture, and grammar. But the reader does not easily find broad statements on how certain classes of words function in the language. Rather, more general patterns arise inductively as the reader learns the details of a specific text.

Common contemporary introductions to the classical language follow an approach largely retaining this commentarial format. Moreover, many draw implicitly on a working knowledge of some modern Chinese language, like Mandarin or Cantonese, as a first introduction to both the language structures and writing of Classical Chinese. Only a minority of scholars espouse introductions without such prior knowledge (Van Norden 2019; Mair 2018). Textbooks by Shadick and by Dawson start, like Zottoli’s, with original texts in traditional characters (Shadick & Ch’iao 1968; Dawson 1984). But a different section holds the notes to accompany each text, and a yet separate section contains a list of the vocabulary encountered. Other recent approaches, e.g. by Rouzer or by Lock and Linebarger, likewise begin with text excerpts, but they hew closer to Zottoli’s presentation by listing vocabulary and commentary close to the individual readings (Rouzer 2007; Lock & Linebarger 2018; cf. also Fuller 2004; Van Norden 2019).

This pedagogical approach also provides a practical mode of addressing a writing system that omits many specifics of Chinese morphophonemics: e.g. 王 can represent either the noun *wáng* ‘king’ or the verb *wàng* ‘be king’, which differ in the modern standard language by tone, but (according to some reconstructions) were distinguished by derivational morphology – *wáng* < **wan* ‘king’ vs. *wàng* < **wanh* ‘is king’ – in earlier stages of the language (Vogelsang 2021; cf. also Pulleyblank 2010, Baxter & Sagart 2014). The writing system leaves unclear what pronunciation to impute to the characters of early texts. Moreover, since many such distinctions have been lost in the modern language, the practice of pronouncing classical texts with the characters’ modern pronunciations (in *any*

modern idiom, including Mandarin, Cantonese, or others) glosses over the problems posed by the script for earlier periods.

Given that many learners approach older periods of Chinese via a modern Chinese language, it seems unsurprising that many popular introductions to the modern language – e.g. Chao’s or DeFrancis’s as two linguistically incisive efforts, or even Giles’s sparse, early textbook – likewise adhere to a similar commentarial approach: the learner memorises conversations, and the remainder of each lesson lists notes on the characters or words of the text (Giles 1922; Chao [1948] 2013; DeFrancis 1976). Finally, Mark Edward Lewis’s course *Chinese Philosophical Texts* uses a roughly comparable format, conveniently illustrating the utility of a commentarial approach even within an online setting (Lewis 2014). Though nothing about the Classical Chinese language or writing system necessitates a commentarial approach, an understanding of the corpus and its historical context suggests why such an approach may have developed naturally within that setting and may remain useful today.

3. Considerations

Among the pedagogical approaches adopted for important early Asian languages, we find three major types: theoretical, as typified by introductions to Sanskrit; historical, as exhibited in approaches to Tocharian; and commentarial, as found in the long lineage of introductions to Chinese. In embarking on a new introduction to an early language, what pedagogical approach is preferable?

Naturally, this depends on the audience for the introduction. In particular, we must consider the learners’ motivations. What are their relevant interests? They might wish to study comparative linguistics, or perhaps archaeology or anthropology, or even literature and history, to name but a few. In the context of Tocharian, the fact that the corpus consists almost exclusively of Buddhist texts serves as a bonanza to the student of Pali interested in religious transmission to Central and East Asia, yet it appears to be a tragedy to the Indo-Europeanist trying to reconstruct the earliest remnants of prehistoric Indo-European society and religious beliefs. Different students can come to early languages with different aims, and some pedagogical approaches may serve the interests of certain students better than others.

At the same time, we must consider the learners' background. Sanskrit, to speakers of modern Indo-European languages, can appear to be a straightforward expansion of features already present in what they know: a possessive *-s* ending, verbs marking person and number of the subject, nominal distinctions based on roles of subject or object of verbs, etc. By contrast, when introducing Tocharian, the grammatical explanations generally assume the students' familiarity with other Indo-European languages and their grammatical categories; but the readings tend to assume a high level of familiarity with Buddhist literature and terminology. The student who comes to Tocharian after studying the common Indo-European religious heritage through the lens of Latin, Greek, and Sanskrit texts may be familiar with the former and not the latter; the student coming to Tocharian from studies of Japanese, Chinese, and Tibetan Buddhism might thrive in the latter and have no knowledge of the former. Moreover, the student coming to classical Chinese could be at ease if already familiar with the script and language structures from a modern Chinese language; but a student coming from outside that tradition might find the unmarked shifts from noun to verb off-putting, and the commentarial approach so skeletal that the grammar appears to be a patchwork of holes.

In addition, we must consider the constraints imposed by each context. Does the content itself determine the proper pedagogical approach? Perhaps a commentarial approach works best with analytic languages or logographic writing systems, whereas a theoretical approach to such material might quickly become too abstract for students to apply consciously in reading or speaking. Several other factors could influence the choice of a particular pedagogical approach: e.g. the literary or linguistic structure of the texts themselves, the processes of textual transmission, the cultural context and the intended audience, to name a few.

We should point out that the search for viable language pedagogy all plays out against the backdrop of pedagogy employed for learning spoken languages. But only rare instances present instructors with the support necessary to employ such pedagogy with early languages. For example, Sanskrit, like Latin in Europe, couples a phonetically attuned writing system with a long spoken tradition long after the language's grammar was "frozen", which thus provides a model for extending the early linguistic corpus to modern situations not originally contained in that corpus. There has in fact been no break in the spoken history of the

language, and this provides a continuity that fosters adaptation to new and current contexts. As a result, Sanskrit, like Latin, enjoys contexts in which a spoken language pedagogy can be and is fruitfully employed.⁷

Tocharian, by contrast, lacks a sufficiently coherent corpus for robust support of such pedagogy. Even the Classical Chinese corpus, despite its longevity and unbroken history, would have difficulty supporting such efforts in an analogous sense, because it has long since lost contact with the actual phonology of earlier eras. What one might hope to achieve would be the speaking of a modern Chinese language, but using a formalised and stylised grammatical apparatus modeled on the constructions still evidenced in the Classical Chinese corpus. In fact, this is akin to the approach often taken in learning the classical language.

4. OERs in the Online Context

In the early stages of academic forays into the Internet, online OERs generally sought to emulate print media. This followed naturally from a conception of online resources as little more, conceptually, than yet another print medium, a conception made clear in terminology as basic as *web page*. As books and other printed media could incorporate images, a web page was not essentially different in kind from the same material on a printed page, though perhaps the images could now move in the form of videos.

The subsequent evolution of the Internet and of frameworks for simplifying and standardizing web development has made the analogy less perfect, and websites now are less confined by the conventions of printed pages than they once were. More concretely, a book must choose a single, particular format (content layout) before it can be printed, and once printed, this format remains invariant. But this need not hold true for a web page in many modern development environments: the same content can be redeployed – in the context of another round of development, or even on the fly as users interact with it – many times over within the confines of a single web page.

This opens up new ways of thinking about creating online open resources, not only for modern languages (cf. Blyth & Thoms 2021), but for early languages as

⁷ Cf., e.g., Avitus (2018) for a perspective on Latin. Cf. Mair (2016) for a personal perspective on spoken Sanskrit. Hastings (2003; 2008) add further context on the motivations and politics surrounding some aspects of modern movements for spoken Sanskrit.

well. As the discussion above has highlighted, the pedagogy of different ancient languages from across Asia has employed quite divergent approaches adapted to the varying contexts and histories of the languages and cultures involved. While many web pages dedicated to introducing readers to these languages employ similar styles of presentation, this need not be the case.

We might consider as an example one particular web development framework: Model-View-Controller (MVC; cf. de la Guardia 2016; Pinkham 2016). After roughly a decade of popular use, this framework has become mature and flexible. The MVC framework conceives of a website in three parts.

The Model refers to the actual content, the data, to be “served up” (given or transmitted) to the user. The nomenclature derives from the fact that the content must be stored somewhere, somehow: in a database. The “shape” or “structure” of this database encapsulates how the site designers conceive of, or “model”, the site content: conceiving of the database as a big spreadsheet, the model describes what data columns the spreadsheet will have, whether that data might be spread across different “tabs” or “sheets” (data tables), and how data in one table can be cross-referenced with data in another.

The View, by contrast, refers to how the user “sees” the content. Showing the user a big spreadsheet might not provide the most engaging or understandable representation of the site’s data. The designers might decide that revealing all the details of the data to users at once might be inefficient or unhelpful: for example, they might only show names of the items in the data, but not the numerical identifiers that the database actually uses for purposes of cross-referencing; or they might only show what they consider to be essential data, leaving many columns hidden. They might not display the data as spreadsheet data at all: what the user sees might be paragraphs, where only the programmer knows that certain names or details have been automatically inserted from the data stored in the database.

Finally, the Controller refers to the automated system that connects the View to the Model. Most importantly, this system is bidirectional. Not only does it automatically decide how to take the data from the Model and display this to the user in the View, but also, depending on the options permitted to the user, it can process user commands through the View and transmit those back to the Model. For example, the Controller might initially take from the Model the data from just a few spreadsheet columns and display those to the user on the initial page-load

of the View. But a well-designed View might let the user opt to display data from more columns: selecting further columns, the Controller will communicate these new selections back to the Model, gather the relevant data, and update the display in the View to include the data requested.

The MVC architecture thus represents a particular organizational scheme employed by standardised web frameworks to facilitate a dynamic user experience for web sites. In particular, it lets developers specify not just a particular display of particular data, but rather a range of ways in which the developers wish to allow users to experience, explore, and even update or correct the data on which the website itself is based. A straightforward and ubiquitous example of this type of framework appears in commercial websites including user reviews, such as Amazon, where the user searches for a product, sees the specifications, adds a review, and now this review becomes part of the data associated with the product and can be displayed to other users, changing the database and hence the website itself.

In the context of online OERs, this provides a novel reframing of the design problem. Whereas early online resources, like printed material, needed to choose a single format and stick to it, web frameworks like MVC loosen some of the strictures of language presentation. In particular, a website displayed initially to support a commentarial pedagogical style might be reformatted, depending on user input, to support a historical reference grammar style. Computationally speaking, the web development framework remains indifferent to the style of presentation and even to the need to readjust how information is presented.

Instead, the work lies with the OER developers themselves. Rather than making a decision on the most effective pedagogical style for presenting learning materials for an early language – or, just as often, choosing not to decide and merely falling back on how the developers learned it themselves – developers must instead imagine a range of different presentation styles and incorporate these into the range of the website’s capabilities. For example, a commentarial presentation of excerpts of Confucius’ *Analects* could be reconfigured into a presentation of points of historical phonology and morphology *illustrated* by snippets of Confucius’ *Analects* (cf., e.g., Van Norden 2019). But what would this entail? On the View side, web developers would need to give the users adjustable parameters or selectors to decide which presentation they would prefer and how they would like aspects of it to display. On the Model side, developers

would need to have a database containing not only excerpts of the *Analects* with a standard commentary, but they would have to tag elements of that commentary as containing, or augment that commentary with, information describing what tidbits of historical linguistic information are on display in that particular excerpt. The Controller would encapsulate how developers envision the website as responding to users' wishes about what information to display and how to serve that information up from the database in an understandable and coherent fashion.

In essence, the online resource can only reproduce different pedagogies so long as the database contains sufficient information to support it, and so long as the user interface offers the option. As scholars develop such resources, this imposes quite a novel charge: as data is entered into the database, scholars must begin to conceive of various different methodologies through which the data could be accessed and make sure that the data is sufficiently granular and appropriately annotated to be able to support a variety of Views through which to access it. The Model should, in essence, contain a number of sub-Models which the Controller could potentially serve up to the View. And the Controller itself must envision how to communicate between the two.

The LRC, for its part, has embarked on just such a reimagining of its resources. The original EIEOL website chose a particular format that had proven effective in print: a series of lessons, with each lesson containing an introduction, a glossed text followed by a continuous translation, and a following discussion of several points of grammar. The originator of the collection, Winfred P. Lehmann, had already edited a short-lived book series on early Indo-European languages adhering to a similar format for the Modern Language Association (Lehmann & Lehmann 1975, Lehmann 2003). The LRC created a computational infrastructure to reproduce a similar format in a set of static web pages with great success. But with the advent of new web frameworks, the LRC has begun the process of reworking the infrastructure behind the scenes. While the visible user interface remains roughly the same as before, the underlying architecture has shifted to the MVC paradigm. What remains now is to re-envision the *range* of interface parameters and *amplify* the underlying database in such a way as to support a more dynamic and interactive variety of pedagogically sensitive user experiences.

This can be tricky and tedious work: it can involve steps as simple as tagging grammar sections according to the part of speech they focus on, or as fine-tuned as updating glosses character by character to decide which parts of a long string

exhibiting sandhi belong to one word and which to another. But as such work progresses, the LRC will eventually be able to offer users the ability to refine the EIEOL presentations in ways that more appropriately support their individual approaches to language learning or teaching. They might switch between theoretical, historical, and commentarial presentations; or they might reorder text excerpts in a lesson series by the number of words they have in common to maximise learners' gains from the vocabulary they've acquired; or they might isolate sentences from the glossed texts, order them by a measure of vocabulary frequency while minimizing the number of new words introduced from one sentence to the next, and then export these to common flash card programmes for learning by spaced repetition. We should no longer conceive of websites introducing early languages as a glorified book, but rather as a collection of books on a particular language, or something much more expansive altogether.

5. Conclusions

Over the preceding centuries instructors have employed a range of pedagogical frameworks through which to teach early languages. In the particular case of three selected ancient languages of Asia, three principal pedagogical 'baskets' have come to the fore: a theoretical approach, typified by Sanskrit pedagogy, whereby instruction centers on the accumulation of grammatical rules in sequence and accompanying text excerpts facilitate their practice; a historical approach, typified by Tocharian pedagogy, where instruction presents a compartmentalised discussion of different grammatical categories, each viewed in the context of its historical evolution from earlier stages of the language, and then passing from grammatical discussion to unsimplified text excerpts drawn straight from the corpus; and a commentarial approach, as typified by Chinese pedagogy, where instruction dives immediately into original texts, with notes commenting on the form and use of individual characters as they appear, but relegating the discussion of general grammatical tendencies to a background role. These pedagogical approaches have arisen rather naturally and make sense given the details of the corpora involved, the specific cultural contexts, and the aims of the particular educational traditions in which they function. But as language learning crosses borders, contexts and educational systems change, and student

interests, preparation, and goals shift, early language instruction may have to adopt a new pedagogy or mixture of pedagogical approaches.

In the print-centred legacy of traditional instruction on these languages, a given instructional text has had to choose one pedagogy and maintain it. As Open Educational Resources have moved online, they have likewise tended to adopt a similar, single-pedagogy approach. This too made sense in earlier epochs of academic forays into the Internet, where web development typically centred on creating similarly static resources, focusing more on a shift of accessibility and reach than on a rethinking of modes of presentation. But modern web frameworks have begun to redefine the way online users can and wish to interact with online resources, and online OERs must find ways to respond to, engage with, and stimulate these new conceptions of interaction and learning.

The Model-View-Controller paradigm encapsulates one way in which web developers can logically structure the possibilities inherent in online resources responsive to user input. As early language resources adapt to the online environment, they can use MVC or similar frameworks to plan user experiences that conceive of something beyond a straightforward print-like presentation or mere access to video tutorials. These frameworks can help developers create resources that can potentially shift between pedagogies, so that theoretical, historical, and commentarial approaches can be conceived simply as different views on the same data. To support this, however, developers and scholars must carefully craft the ways in which they gather and store data, and likewise go beyond their own personal or “traditional” trajectories of having learned a given language to provide users with a range of instructional interfaces attuned to a variety of possible learning styles and informational foci. With such attention to the careful construction of data and interface in early language online learning resources, we might finally begin to achieve

OER + MVC ≥ print(‘medium’).

References

- Avitus, A. G. (2018). Spoken Latin: Learning, Teaching, Lecturing and Research. *Journal of Classics Teaching*, 19(37), 46–52.
- Baxter, W. H., & Sagart, L. (2014). *Old Chinese: A New Reconstruction*. Oxford: Oxford University Press.
- Benfey, T. (1852). *Handbuch der Sanskritsprache: Band I. Grammatik*. Leipzig: F. A. Brockhaus.
- Benfey, T. (1853). *Handbuch der Sanskritsprache: Band II. Chrestomathie*. Leipzig: F. A. Brockhaus.
- Bird, J., Blagburn, B., & Osborne, M. (2022). Reclaiming the Classics for a Diverse and Global World Through OER. *KULA*, 6(1), 1–11.
- Blyth, C. S. (2012). Opening Up Foreign Language Education with Open Educational Resources: The Case of Français interactif. In *Hybrid Language Teaching and Learning: Exploring Theoretical, Pedagogical and Curricular Issues*, 196–218.
- Blyth, C. S., & Thoms, J. J. (eds). (2021). *Open Education and Second Language Learning and Teaching: The Rise of a New Knowledge Ecology*. Bristol: Multilingual Matters.
- Chao, Y. R. ([1948] 2013). *Mandarin Primer: An Intensive Course in Spoken Chinese*. Cambridge: Harvard University Press.
- Dawson, R. S. (1984). *A New Introduction to Classical Chinese*. Oxford: Oxford University Press.
- de la Guardia, C. (2016). *Python Web Frameworks*. Sebastopol, CA: O'Reilly Media.
- DeFrancis, J. (1976). *Beginning Chinese*. New Haven: Yale University Press.
- Delbrück, B. (1976). *Altindische Syntax*. Darmstadt: Wissenschaftliche Buchgesellschaft [Max Niemeyer Verlag].
- Dong, H. (2020). *A History of the Chinese Language*. London: Routledge.
- Fuller, M. A. (2004). *An Introduction to Literary Chinese*. Cambridge, MA: Harvard University Press.

- Giles, H. A. (1922). *Chinese Without a Teacher: Being a Collection of Easy and Useful Sentences in the Mandarin Dialect, with a Vocabulary*. Shanghai: Kelly & Walsh.
- Hartman, C. (1998). Literary Chinese. In Nienhauser *et al.*, *The Indiana Companion to Traditional Chinese Literature: Volume 2*. Indiana University Press, 92–97.
- Hastings, A. M. (2003). Simplifying Sanskrit. *Pragmatics*, 13(4), 499–513.
- Hastings, A. M. (2008). Licked by the Mother Tongue: Imagining Everyday Sanskrit at Home and in the World. *Journal of Linguistic Anthropology*, 18(1), 24–45.
- Jamison, S. W. (2008a). Middle Indic. In Woodard, R. D. (ed.), *The Ancient Languages of Asia and the America*. New York: Cambridge University Press, 33–49.
- Jamison, S. W. (2008b). Sanskrit. In Woodard, R. D. (ed.), *The Ancient Languages of Asia and the Americas*. New York: Cambridge University Press, 6–32.
- Johnson, F. (1847). *Hitopadeśa: The Sanskrit Text, with a Grammatical Analysis, Alphabetically Arranged*. London: Wm. H. Allen and Co.
- Kāle, M. R. (1961). *The Abhijñānaśākuntalam of Kālidāsa: with the Commentary of Rāghavabhaṭṭa, Various Readings, Introduction, Literal Translation, Exhaustive Notes, and Appendices*. Delhi: Motilal Banarsidass.
- Köntges, T., Lesage, R., Robertson, B., Sellick, J., & Stylianopoulos, L. W. (2017). Open Greek and Latin: Digital Humanities in an Open Collaboration with Pedagogy. Presented at: IFLA WLIC 2019 – Athens, Greece – Libraries: dialogue for change. August 27.
- Krause, T. B. (2019). IE5.com. Historical Linguistics pro populo. *Folia Linguistica Historica*, 40(2), 621–632.
- Krause, W., & Thomas, W. (1960). *Tocharisches Elementarbuch. Band I: Grammatik*. Heidelberg: Universitätsverlag Winter.
- Krause, W., & Thomas, W. (1964). *Tocharisches Elementarbuch. Band II: Texte und Glossar*. Heidelberg: Universitätsverlag Winter.
- Lanman, C. R. (1884). *A Sanskrit Reader*. Boston: Ginn and Company.

- Lassen, C., & Gildemeister, J. (1865). *Anthologia Sanscritica Glossario Instructa in Usum Scholarum*. Bonn: Adolphus Marcus.
- Lehmann, R. P. M., & Lehmann, W. P. (1975). *An Introduction to Old Irish*. New York: Modern Language Association of America.
- Lehmann, W. P. (2003). General Principles of the Series. Personal communication.
- Lewis, M. E. (2014). *Chinese Philosophical Texts*. Stanford: Chinese Texts. Retrieved November 21, 2023, from <https://history.stanford.edu/publications/chinese-philosophical-texts>
- Lock, G., & Linebarger, G. S. (2018). *Chinese Buddhist Texts: An Introductory Reader*. London: Routledge.
- Mair, V. H. (1998). Tzu-shu 字書 or tzu-tien 字典 (dictionaries). In Nienhauser *et al.*, *The Indiana Companion to Traditional Chinese Literature: Volume 2*. Indiana University Press, 165–172.
- Mair, V. H. (2016). Spoken Sanskrit. *Language Log*. Retrieved August 30, 2022, from <https://languagelog ldc.upenn.edu/nll/?p=23412>
- Mair, V. H. (2018). How to teach Literary Sinitic / Classical Chinese. *Language Log*. Retrieved August 30, 2022, from <https://languagelog ldc.upenn.edu/nll/?p=39905>
- Malzahn, M. (ed.). (2007). *Instrumenta Tocharica*. Heidelberg: Universitätsverlag Winter.
- Malzahn, M. (2017). Tocharian and Historical Sociolinguistics: Evidence from a Fragmentary Corpus. *Open Linguistics*, 3(1), 157–177.
- Malzahn, M. (2018). A Contrastive Survey of Genres of Sanskrit and Tocharian Buddhist Texts. *Written Monuments of the Orient*, 4(1), 3–24.
- Monier-Williams, M. (1864). *A Practical Grammar of the Sanskrit Language: Arranged with Reference to the Classical Languages of Europe, for the Use of English Students*. Oxford: Clarendon Press.
- Nienhauser Jr., W. H., Hartman, C., Ma, Y. W., & West, S. H. (eds). (1986). *The Indiana Companion to Traditional Chinese Literature: Volume 1*. Indiana University Press.

- Nienhauser Jr., W. H., Hartman, C., & Galer, S. W. (eds). (1998). *The Indiana Companion to Traditional Chinese Literature: Volume 2*. Indiana University Press.
- Perry, E. D. (1885). *A Sanskrit Primer*. Boston: Ginn and Company.
- Peyraube, A. (1999). Historical change in Chinese grammar. *Cahiers de Linguistique – Asie Orientale*, 177–226.
- Peyrot, M. (2008). Variation and change in Tocharian B. Amsterdam: Rodopi.
- Pinault, G.-J. (1992). Introduction au Tokharien. *LALIES: Actes des sessions de linguistique et de littérature*, 11, 5–224.
- Pinault, G.-J. (2008). *Chrestomathie Tokharienne*. Leuven: Peeters.
- Pinkham, A. (2016). *Django Unleashed*. Indianapolis, Indiana: Pearson Education.
- Poucha, P. (1955). *Thesaurus Linguae Tocharicae Dialecti A*. Prague: Czechoslovak Oriental Institute.
- Poucha, P. (1956). *Chrestomathia Tocharica*. Prague: Czechoslovak Oriental Institute.
- Pulleyblank, E. G. (2010). *Outline of Classical Chinese Grammar*. Vancouver: University of British Columbia Press.
- Qiu, X. (2000). *Chinese Writing*. Berkeley: University of California, Berkeley.
- Rouzer, P. F. (2007). *A New Practical Primer of Literary Chinese*. Cambridge, MA: Harvard University Press.
- Ruppel, A. M. (2017). *The Cambridge Introduction to Sanskrit*. Cambridge: Cambridge University Press.
- Ruppel, A. M. (2021). *An Introductory Sanskrit Reader: Improving Reading Fluency*. Leiden: Brill.
- Sato, M., & Loewen, S. (eds). (2019). *Evidence-Based Second Language Pedagogy: A Collection of Instructed Second Language Acquisition Studies*. London: Routledge.
- Sieg, E., Siegling, W., & Schulze, W. (1931). *Tocharische Grammatik*. Göttingen: Vandenhoeck & Ruprecht.

- Shadick, H., & Ch'iao, C. 喬健. (1968). *A First Course in Literary Chinese*. Ithaca: Cornell University Press.
- Speijer, J. S. (1886). *Sanskrit Syntax*. Leiden: E. J. Brill.
- Staal, J. F. (ed.). (1972). *A Reader on the Sanskrit Grammarians*. Boston: MIT Press.
- Stenzler, A. F., & Pischel, R. (2016). *Elementarbuch der Sanskrit-Sprache: Grammatik, Texte, Wörterbuch*. Berlin: Walter de Gruyter GmbH & Co KG.
- Van Norden, B. W. (2019). *Classical Chinese for Everyone: A Guide for Absolute Beginners*. Indianapolis: Hackett.
- Vogelsang, K. (2021). *Introduction to Classical Chinese*. Oxford: Oxford University Press.
- Weiss, M. (2022). *Kuśiññe Kantwo: Elementary Lessons in Tocharian B, with Exercises, Vocabulary, and Notes on Historical Grammar*. Ann Arbor: Beech Stave Press.
- Whitney, W. D. (1889). *Sanskrit Grammar: Including both the Classical Language, and the Older Dialects, of Veda and Brahmana*. Cambridge, MA: Harvard University Press.
- Wieger, L. (1965). *Chinese Characters: Their Origin, Etymology, History, Classification and Signification. A Thorough Study from Chinese Documents*. New York: Paragon Book Reprint Corporation; Dover Publications.
- Witzel, M., & Wu, Q. (2018). *The Two Oldest Veda Manuscripts: Facsimile Edition of Vājasaneyi-Saṃhitā 1–20 (Saṃhitā- and Padapāṭha) from Nepal and Western Tibet (c. 1150 CE)*. Cambridge, MA: Harvard University Press.
- Woodard, R. D. (ed.). (2008). *The Ancient Languages of Asia and the Americas*. New York: Cambridge University Press.
- Woolner, A. C. (1917). *Introduction to Prakrit*. Calcutta: Baptist Mission Press.
- Yong, H., & Peng, J. (2008). *Chinese Lexicography: A History from 1046 BC to AD 1911*. Oxford: Oxford University Press.
- Zottoli, P. A. (1879–1882). *Cursus Litteraturae Sinicae Neo-Missionariis Accommodatus*. Chang-Hai: Typographia Missionis Catholicæ in orphanotrophio Tou-sè-wè (Tou-chan-wan).