

Entry

Polyglot Lexicons and Encyclopedic Works in Late Imperial China

Ling-Wei Kung 

Institute of History and Philology, Academia Sinica, Taipei City 115, Taiwan; lingwei@ihp.sinica.edu.tw

Definition

The present article reinterprets the history of polyglot lexicography and encyclopedic language projects in late imperial China from the Yuan and the Ming through the Qing periods by tracing a three-stage transformation. The Yuan period inaugurated a foundational regime of phonetic transcription anchored in the 'Phags-pa script (Ch. Basiba zi 八思巴字) while already experimenting with semantic pairing in the early Sino–Mongol glossary conventionally known as *Zhiyuan yiyu* (至元譯語). The Ming consolidated that legacy into a state curriculum centered on the *Huayi yiyu* (華夷譯語) corpus, together with frontier manuals such as *Beilu yiyu* (北虜譯語), which systematized domain-based vocabulary and coupled it with documentary templates for tribute, diplomacy, and administration. The Qing, finally, reconceived multilingual lexicography as a project of imperial integration, recentring Manchu as the pivot language in the *Qing wen jian* (清文鑒) series and culminating in the five-language *Wuti Qing wen jian* (五體清文鑒). Specialized compendia such as *Xiyu tongwen zhi* (西域同文志) normalized toponyms across scripts in newly incorporated territories. Complementing official compilations, market-facing handbooks—including *Menggu zazi* (蒙古雜字)—and the dialogic textbooks *Nogeoldae* (Ch. *Lao qida* 老乞大) and *Bak Tongsa* (Ch. *Piao tongshi* 朴通事) produced within Joseon's translator-training institutions reveal a multi-sited ecosystem in which court, frontier, marketplace, and foreign language schools co-produced the infrastructure of interlingual governance. By following the shift from “how to read” (phonetic) to “what it means” (semantic) and ultimately to “what it governs” (administrative integration), this article argues that polyglot lexicons were not merely repositories of words but instruments that made a multilingual empire legible, speakable, and governable.

Keywords: polyglot lexicons; *Huayi yiyu*; *Wuti Qing wen jian*; *Xiyu tongwen zhi*; multilingualism; Sino–Inner Asian Communication



Received: 7 October 2025

Revised: 5 November 2025

Accepted: 22 December 2025

Published: 25 December 2025

Copyright: © 2025 by the author.

Licensee MDPI, Basel, Switzerland.

This article is an open access article distributed under the terms and conditions of the [Creative Commons Attribution \(CC BY\)](https://creativecommons.org/licenses/by/4.0/) license.

1. Introduction

Late imperial China governed vast, multilingual spaces of Inner Asia in which merchants bargained, envoys exchanged gifts and documents, troops patrolled frontiers, and monasteries mediated ritual diplomacy. Under such conditions, dictionaries were never just scholarly crutches; they were components of administrative machinery. They trained personnel, standardized terms and formats, and stabilized correspondences among languages that did not share scripts or orthographic conventions. Read in this light, the history of polyglot lexicography is inseparable from the history of state formation, for it exposes the techniques through which officials transformed the babel of borderlands into files, tables, and procedural speech.

This study develops a narrative in three movements—Yuan, Ming, and Qing—while also seeking to trace the development of multilingual encyclopedic works beyond dynastic periodization. In addition to the celebrated 'Phags-pa script (Ch. Basiba zi 八思巴字) experiments, the Mongol-Yuan empire contributed the Sino-Mongol glossary known as *Zhiyuan yiyu* (至元譯語), which organized the polyglot world into conceptual drawers [1,2]. The Ming installed that conceptual habit into a durable curriculum, producing the *Huayi yiyu* (華夷譯語) corpus together with frontier manuals such as *Beilu yiyu* (北虜譯語), which paired semantic fields with documentary templates [2–5]. The Qing reanchored the stack, the descriptive metaphor employed by the present article, in Manchu and multiplied it across languages in the *Qing wen jian* (清文鑒) series, culminating in the five-language *Wuti Qing wen jian* (五體清文鑒), while specialized works, including *Xiyu tongwen zhi* (西域同文志), standardized toponyms [6–8]. Alongside these official projects, civilian and market manuals such as *Menggu zazi* (蒙古雜字) and the *Qingyu lao qida* (清語老乞大) dialogic textbooks [9,10], as well as the Joseon's *Nogeoldae* (Ch. *Lao qida* 老乞大) and *Bak Tongsa* (Ch. *Piao tongshi* 朴通事) [11,12], supplied scripts for trade, travel, and interpreter training. Together, these materials illuminate the double life of polyglot books: they were aids to cross-lingual communication and, at the same time, blueprints for bureaucratic procedure.

The argument proceeds by reconstructing the functional design of key corpora rather than cataloging titles for their own sake. The questions are simple but revealing: Which language sits at the center of a given work, and what does that centrality do? What kinds of entries—phonetic, semantic, and/or dialogic—dominate a compilation, and how do they map onto the tasks of the users for whom the book was made? How do classification schemes—heaven, earth, and human affairs; flora and fauna; weapons and tools; as well as offices and laws—encode assumptions about what a multilingual state must know? Taken together, the answers realign our periodization and recalibrate our sense of where lexicographic innovation occurred.

2. Yuan Foundations

The Mongol conquest reconfigured linguistic relations throughout Inner Asia and North China. In the court's early decades, the most visible innovation was the 'Phags-pa script, an imperial alphabet designed to write Mongolian, Chinese, Tibetan, and Sanskrit in a single orthographic key [13]. Much of what survives from this experiment belongs to the sphere of phonetic discipline: rhyme-books that spelt Chinese syllables in 'Phags-pa letters and marked tone categories; appendices in encyclopedic miscellanies that juxtaposed Chinese surnames with their Mongolian readings; and tables that coaxed officials from different speech communities toward mutual intelligibility [14]. This was the age of “how to read.” It came with the urgent problems of everyday practice: how to recite names, to pronounce place-names, to render oral statements from envoys, and to call out commands on campaign. The state's investment in a universal script thus produced a phonetic technology serving a world in motion.

Yet the Yuan was not confined to phonetics. The Sino-Mongol glossary that later tradition knows as *Zhiyuan yiyu* marks a decisive turn. In it, Chinese headwords were paired with Mongolian equivalents and grouped by domains such as astronomy, geography, human relations, garments, colors, horses and saddles, tools, birds and beasts, and cereal grains. The organization converted a loose list into a conceptual map. Looking up words no longer depended on syllabic proximity or purely alphabetical arrangement but on a semantic economy of use. Interpreters could navigate the lexicon by topic, and officials could locate clusters of terms relevant to predictable tasks: receiving caravans, provisioning troops, and assessing tribute lists. The glossary, transmitted in multiple fourteenth- and fifteenth-century reprints, was no antiquarian curiosity; its very reproducibility be-

trays practical value [1]. In that reproducibility one can glimpse a nascent standardization of subject matter across languages, a habit that would become central in the centuries to come. For example, in the *Zhiyuan yiyu*, the entry for “horse” (馬) is paired with its Mongolian equivalent “mori(n)” in the Chinese transliteration of “muli” (木里) within the “equestrian” (鞍馬) domain, illustrating how topical grouping was adopted instead of purely phonetic or alphabetical retrieval. A similar arrangement appears for cereals such as “millet” (budağa, 不歹) and “barley” (arbai, 牙立百), signaling a lexicon designed for provisioning, tribute, and routine assessment rather than literary citation. These topical drawers made the glossary a working tool for interpreters who needed to locate families of terms quickly in predictable administrative tasks.

The institutional context helps to explain this consolidation. Inheriting the institutional precedent of the Liao and Jin dynasties, the Yuan established the Huitong guan (會同館) as a central bureau for training interpreters and overseeing translation services, an institution that was subsequently taken over by the Ming dynasty afterwards [15,16]. Although the surviving paperwork is thin, the office’s existence reveals an administrative demand for the kinds of tools represented by ‘Phags-pa transcriptions and the Sino-Mongol topical glossary. Practical needs—tribute tallies, tax registers, and military dispatches—generated specific genres. The earliest Sino-Mongol materials, then, should be read not as random survivals but as pieces of a newly regularized workflow in which speech from different communities had to be rendered legible.

3. Ming Institutionalization

The Ming state inherited both the problems and the prototypes of its predecessor and transformed them into a program. Two bureaus—commonly known as the Siyi guan (四夷館) and the Huitong guan—organized the recruitment and training of translators with a division of labor that paired written translation (the purview of the Siyi guan) and oral performance (that of the Huitong guan), a separation reflecting the bureaucracy’s distinct needs for documentary precision and diplomatic immediacy [15,17]. Out of these institutions emerged the corpus known collectively as *Huayi yiyu*, a suite of bilingual dictionaries and handbooks assigned to language-specific offices [3]. Over the fifteenth and sixteenth centuries, these offices multiplied to represent a wide arc of neighboring communities, including Mongolian, Jurchen, Tibetan, Sanskrit, Persian, Turkic, Tai languages of the southwest, Burmese, and Siamese. The books were produced not as showpieces but as working tools for bureaucrats who managed annual missions, gifts, and petitions.

At the heart of the corpus was a triadic structure that bound vocabulary to procedure. The *yiyu* (譯語) volumes presented domain-organized lists that mapped Chinese lemmas to equivalents in other languages. Alongside *yiyu* stood “zazi” (雜字) collections, elementary lists that emphasized practical items of daily exchange, and “laiwen” (來文), documentary templates for petitions and replies [18].

Several features deserve emphasis. First, the base language of the series was Chinese, and the moral economy of the books was Chinese-centric. Second, the organization of the *Huayi yiyu* volumes preserved and expanded the Yuan impulse toward semantic classification. One can see this most clearly in the Mongolian volume—often referred to as the *Dada yiyu* (韃靼譯語)—which assimilated earlier Sino-Mongol materials into the Ming’s standardized taxonomy [4]. A representative entry from the *Dada yiyu* makes this design visible: within the “birds and animals” (鳥獸) domain, a Chinese lemma such as “horse” is matched to the Mongolian equivalent “morin” (抹鄰). The fact that the same domain houses “sheep” (qonin > 豁綯), “goat” (imağan > 亦馬安), “cattle” (hüker > 忽格兒), and “camel” (temeğen > 帖蔑延)—together widely known as “the five snouts” (tabun qosıǵu mal) in Mongolian nomadic culture—shows that the series anticipated task-bundles: users

moved not by alphabetical proximity but by workflows that required clusters of semantically adjacent words.

The corpus had neighbors beyond the court. The late sixteenth century saw the creation and circulation of *Beilu yiyu*, a compact Sino–Mongolian lexicon embedded within a military manual known as *Dengtan bijiu* (登壇必究) [5]. Here the lexicon turned outward to the field, curating the world of horses, scouts, formations, commands, and provisioning of garrisons. What *Huayi yiyu* performed for the chancellery, *Beilu yiyu* performed for border troops: it inventoried the words needed to recognize equipment and behavior, to grasp orders, and to negotiate supply [19]. This programmatic Ming regime was what gave the corpus its encyclopedic quality. Encyclopedism here denotes a method of carving the world into retrievable domains and tethering the resulting wordlists to standardized genres of text. A Ming interpreter, therefore, operated within a console of coupled modules: pick a domain, match an equivalent, paste it into a template, and produce a document that both parties would recognize.

4. Qing Integration

The Qing consolidated the Ming inheritance but reconfigured its language politics. The editorial line that began with *Qing wen jian* of 1708, a single-language Manchu classified word book [20], culminated near the end of the eighteenth century in the five-language *Wuti Qing wen jian* (commonly termed the *Pentaglot*), a monumental, state-funded project whose sheer scale of production and controlled distribution to key administrative nodes underscored its role as an instrument of imperial authority [7]. The path from the first to the last is not a simple matter of addition. It signals a shift in the very organization of imperial knowledge. At each stage—two-language Manchu–Chinese, three-language Manchu–Mongolian–Chinese, and four-language Manchu–Tibetan–Mongolian–Chinese—the series retained Manchu as the base and imposed upon the other languages a discipline of mutual equivalence and, crucially, of transliteration into Manchu [21].

The technical consequences of this decision can be measured at the level of a single line on the printed page. Tibetan and Turkic entries came supplied with Manchu transcriptions that regularize how non-alphabetic or differently alphabetized words should be sounded by a Manchu-literate clerk [7]. In practice, this means that someone trained in Manchu orthography could pronounce a Tibetan place-name, read a Mongolian term aloud, and map a Turkic word into a Manchu syllabic grid. The capacity to speak across scripts was not simply a convenience; it was a way of bringing disparate textual communities within a shared acoustic discipline. The page became an interface that allowed a scribe in the capital and a banner officer in a frontier post to converge on a single sound for a single referent.

If *Huayi yiyu* presented a Chinese set of drawers with foreign equivalents tucked into them, the *Pentaglot* and its sister volumes presented a Manchu chest in which Tibetan and Turkic drawers slid on Manchu rails. This was as much an administrative decision as an ethnic one. The Manchu script suited the task of transliterating other languages with an efficiency that Chinese characters could not easily match. Orthographic convenience was not an apolitical rationale—choices about scripts carried symbolic weight—but the Qing multilingualism cannot be explained by symbolism alone. The physical structure of the page made its semantic organization echoed Ming practice in ways that assured continuity of function.

The printed lineation of the *Wuti Qing wen jian* makes this audible. After the Manchu head entry, Tibetan and Turkic forms are given and then normalized by a Manchu transliteration line, so that its Manchu readers could pronounce—and therefore dictate or copy—non-alphabetic or differently alphabetized items with minimal ambiguity. This layout reduces the chance that a Tibetan toponym or Turkic ethnonym will drift phonetically across

documents as it circulates through offices. In practical terms, the transliteration line functioned like a pronunciation checksum for multilingual paperwork.

The wider integrative ambition of the Qing lexicographic project becomes explicit in special-purpose compendia. *Xiyu tongwen zhi* standardized toponyms across Manchu, Chinese, Mongolian (including Oirat forms), Tibetan, and Turkic for the newly conquered spaces to the northwest [8]. The function of the book is obvious from the preface to its tables. A courier reading a Tibetan docket had to be able to align a place-name with its Mongolian form; a Chinese clerk tasked with marking a map had to transcribe a Turkic name in a way that banner officers would recognize; Manchu transliteration mediated among them. The result was a grid of equivalence that reduced the risk of misidentification, duplication, and confusion. The governance of space required a governance of names, and the Qing's toponymic discipline extended lexicographic technique into cartography and logistics.

The same habit of polygraphy permeated ritual life. Bilingual and quadrilingual inscriptions in Beijing and Chengde reproduced imperial pronouncements in Manchu, Chinese, Mongolian, and Tibetan, staging the empire's linguistic unity in stone [22]. The bureaucratic archives of the Lifan yuan (理藩院) multiplied parallel versions and interlinear glosses for internal use [23]. These practices overlapped with the logic of the *Pentaglot*. They are all artifacts of an administrative preference for repeatable correspondences and transliterable forms. In their aggregate they also reveal something more general, namely that the empire used language to define its internal boundaries as well as to talk across them. This administrative logic extended beyond lexicographic practice. The evolving Manchu terminology for “tongshi” (通事, interpreter)—“tungse” and later “hafumbukū”—likewise reflects the effort to discipline multilingual communication [24]. Interpreters, no less than scribes, operated within transliteration regimes and classificatory norms that paralleled those embedded in the *Qing wen jian* series and the *Pentaglot*, binding human intermediaries to the same linguistic order that structured the page.

5. Civilian and Market Handbooks

No account of late imperial polyglossia is complete if it remains confined to the palace workshop and the ministries. The presses of commercial cities and the needs of traveling communities produced their own language books. Among these, the *zazi* tradition is emblematic. A title such as *Man-Han tongwen xin chu duixiang Menggu zazi* (滿漢同文新出對像蒙古雜字) announced a user-oriented design [9]. Its pages arranged words for goods, weights and measures, prices, means of transport, lodging, greetings, and bargaining into parallel columns, typically Mongolian–Chinese with simple phonetic prompts. The illustrated layout was performative. It anticipated how a merchant or broker would move through a market day, providing formulae for opening and closing conversations as well as a set of lexical resources for specifying quantities, qualities, and obligations. The selection of vocabulary, densely focused on commodities and services, made visible a world curiously absent from the court's documentary templates: the world of haggling, hospitality, and credit.

The Manchu *Qingyu lao qida* line advanced this pragmatic codification in dialogic form [10]. Here the characters were set up. For instance, “host” and “guest” as well as “shopkeeper” and “broker” play out conventional exchanges—greeting, requesting, making an offer, demurring, conceding, tendering payment, presenting a gift, or inviting to a banquet or lodgings. The decision to build the book out of scenes produced a template that learners could memorize and repurpose. If *Huayi yiyu* was designed to place words into documents, the *Lao qida* line was designed to place words into social encounters. In so doing, it captured the etiquette and the taxonomies of a market in which the relationship mattered as much as the exchange. It is for this reason that the books paid attention to

forms of address and to the polite management of refusal and concession. Language was not merely a vehicle of information here; it was a choreography of face and obligation.

Seen from the vantage of the polyglot state, these market handbooks represented a diffusion of techniques downward. Manchu transliteration standards filtered into usage guides for clerks in shops; the taxonomies of *yiyu* reappeared as simplified lists of cloth types and tea grades. The cross-traffic was two-way: as new commodities and measures entered circulation, the *zazi*'s granular lexicon often supplied the first stable labels that fiscal and legal offices later codified. But the diffusion ran both ways. The granular lexicon of goods and measures that *zazi* accumulated found its way back into official usage as new commodities and measures required standardized names [25]. The line between courtly and civilian knowledge was porous, and polyglot lexicography was the membrane that connected them.

6. The Joseon Pathway

The circulation of dialogic language templates extended beyond the Ming and Qing polities. In Joseon Korea, the Sayeokwon (Ch. Siyi yuan 司譯院) crafted a formal curriculum for learning spoken Chinese, and two textbooks, *Nogeoldae* and *Bak Tongsa*, dominated the program [11,12]. Their influence grew not because they compiled rare words but because they stabilized speech events. The books distilled the etiquette and formulae of greeting, visit, negotiation, gift exchange, banqueting, and departure into teachable units, presenting them as sequences that a student could internalize. What gave them additional force was the decision to anchor the Chinese language of instruction in a recognizable standard: the Beijing koiné of the late Yuan and early Ming [26].

Over time the dialogic method embedded in these textbooks migrated to other target languages. There were Manchu and Mongolian versions that translated the scripts of conversation while retaining the Chinese role labels as mnemonic anchors [10,27,28]. The persistence of labels such as *Lao qida*/*Nogeoldae* and *Bak Tongsa* across languages is not a trivial stylistic point. It shows that the architecture of roles and speech turns could be detached from any one language and assigned to others, a form of portability that mirrored, at the oral level, the cross-lingual equivalence tables of the *Pentaglot*. A learner trained in Joseon could therefore move with a modicum of confidence from Chinese to Manchu or Mongolian because the scenes were familiar even when the words changed. The textbooks thus formed a transregional conduit through which the Ming and Qing regimes' ways of organizing speech entered a foreign bureaucracy and continued to shape how interpreters worked.

The relation between the Joseon materials and the civilian handbooks produced within the Ming–Qing orbit is mutually illuminating. Both insisted that a script of conversation, with predictable turns and polite formulae, was a necessary supplement to the static pairing of lemmas [29]. Both registered a domain of practice that was underrepresented in the courtly lexicons: the practical arts of hosting, buying, transporting, and closing accounts. But whereas *zazi* assumed a self-motivated user who grabbed phrases as needed, *Nogeoldae* and *Bak Tongsa* insisted on graded instruction and the achievement of fluency in sequences. In short, the civilian handbooks disclosed a market ethic; the Joseon textbooks disclosed a classroom discipline. Together with the court dictionaries they completed the ecology of multilingual infrastructure. Having outlined the institutional and regional ecology of late imperial polyglossia, we can now ask what these design choices collectively reveal about the logic of multilingual governance.

7. Horses, Documents, and Names

Three case portraits may crystallize the preceding analysis. The first concerns the vocabulary of horses and saddlery. In the topical Sino–Mongol lists of the Yuan, one already finds an attention to the objects and operations of mounted life [1]. The Ming elevated that vocabulary, distributing it between *Dada yiyu* and *Beilu yiyu*, the former standardizing the terms for interpretive writing and the latter curating what soldiers needed to know in the field [5]. The Qing *Pentaglot* harmonized the same domain across five languages and embedded Manchu transliteration rules that allowed Tibetan and Turkic to be sounded correctly [7]. The semantic field thus traveled across centuries and forms: from list to curriculum and to an integrated, multi-script dictionary. What persisted was the equestrian empire’s need to name what it rode.

The second portrait compares document templates and conversational scripts. In *Huayi yiyu*, a letter of credence or a petition obeyed a propriety of order in which titles preceded names, seals and dates settled at the foot, and appositional phrases enacted reciprocal obligations [3]. The effect was to routinize the ceremonial page. In the *Lao qida/Nogeoldae* tradition, the plot was different: forms of address secured rapport; a request was hedged with honorifics; refusals were softened before concessions arrived; and a price was adjusted through a standard choreography. The effect was to routinize the ceremonial encounter. Both routinizations were lexicographic achievements. They stabilized sequences into learnable modules and ensured that words could be retrieved and inserted into a correct position in the dance.

The third portrait concerns toponyms. *Xiyu tongwen zhi* seems, at first glance, an arid table of names [8]. But read alongside the *Pentaglot*, it reveals how the empire formalized equivalence not only among words but among places. Once a name could be written in Manchu, Chinese, Mongolian, Tibetan, and Turkic, and once those written forms were aligned by transliteration and meaning, a courier and a clerk could coordinate across scripts without losing the referent. Spatial administration was never purely cartographic; it depended on the ability to utter and transcribe names, to anchor them in official lists, and to keep them stable in correspondence. In this sense, the toponymic compendium was the spatial analog of the lexical dictionary.

8. From “How to Read” to “What It Means” and to “What It Governs”

The story told so far exhibits a stack rather than a simple succession. The Yuan’s phonetic tools did not disappear when semantic dictionaries matured; the Ming’s domain classifications did not dissolve when Manchu became the pivot of the Qing. Each layer enabled the next. ‘Phags-pa transcriptions and rhyme-books built a foundation on which glossaries could specify meanings and on which late imperial users could pronounce foreign words aloud [14]. The topical Sino–Mongol lists of the Yuan, preserved in *Zhiyuan yiyu*, rehearsed the Ming habit of carving vocabulary into recoverable domains [1,30]. The Ming habit, once installed as a state curriculum in *Huayi yiyu*, furnished the bones that Qing editors clothed in additional languages [3]. The Manchu transliteration rules of the *Pentaglot*, in turn, extended the reach of the earlier layers by providing a shared acoustic discipline [7].

This layering allows us to revisit three common assumptions. First, the Yuan was not merely an age of phonetic play. *Zhiyuan yiyu* makes it plain that semantic pairing and domain classification had begun [1]. Second, the Ming was not only outward-looking. Its dictionaries, with their documentary templates and standardized categories, trained a cadre of domestic clerks and interpreters to manage routine writing [3]. Third, the Qing did not install Manchu at the lexical center only to perform ethnic primacy. The technical virtues of the Manchu script—its capacity to render non-alphabetic and abjad-based

languages legible for Manchu-literate readers—made it a practical choice for interscript mediation [7]. Symbolic and technical considerations were aligned, not opposed.

A related question is why neither the Mongols nor the Manchus attempted to install their mother tongue as a universal administrative language. The potential answer is technical and political. Both regimes treated linguistic diversity as an administrative asset: preserving established local scripts and speech communities sustained elite cooperation and lowered the costs of routine communication, while centralized translation infrastructures—glossaries, templates, transliteration rules—supplied interoperability. In short, they optimized for governable equivalence rather than imperial monolingualism.

A further implication concerns the distribution of authority. When the Ming made Chinese the base language of *Huayi yiyu*, it projected a centripetal order in which foreigners were received and translated into the capital's categories [3]. When the Qing made Manchu the pivot of its polyglot dictionaries, it articulated an imperial order in which Mongolian, Tibetan, and Turkic speech could be mapped to a single orthographic key [7]. In both cases, the lexicon was a mirror of rule. And in both cases, the market and classroom materials—*zazi*, *Lao qida*, *Nogeoldae*, and *Bak Tongsa*—ensured that these mirrors did not reflect the court alone; they reflected the texture of everyday exchange [9–12].

9. Reading the Books as Statecraft

To treat polyglot lexicons as instruments of governance is to pay attention to their design. Entry structures matter. A Ming entry that offered a Chinese lemma, a foreign equivalent, and a brief phrase is different from a Qing entry that added a transliteration line, and both are different from a dialogic exchange that framed a phrase within a scenario. The instruments were calibrated to different users: the chancellor clerk, the frontier soldier, the merchant, and the interpreter student. The Qing translation examination cultivated a multiethnic banner literati with advanced Manchu literacy, sustaining the linguistic competencies that these instruments presupposed [31].

Domain coverage matters as well. The relative density of items in categories such as flora and fauna [32], offices and laws, weapons and tools, commodities and measures, or hospitality and transport is an index of the world the book imagined its user to inhabit. It is then no surprise that *Beilu yiyu* was rich in equestrian and tactical terms [5], or that *zazi* volumes lavished attention on measures and bargaining [9]. These choices were not accidental; they encoded a sociology of tasks.

The politics of labels has to be read with care. Terms like “Huihui” (回回), “Gaochang” (高昌), “Baiyi” (百譯), or “Xifan” (西番) changed meanings across time and space. A polyglot dictionary published in the fifteenth century and another compiled in the eighteenth might assign the same exonym, “Xifan”, for instance, to different constellations of speakers [30,33]. To avoid reifying such fluid terms, one must align their usage with date and locale and recognize them as documentary pointers rather than stable ethnic essences. Doing so clarifies how the state tracked flows and contacts; it does not license us to harden names into people.

The materiality of the books also answered to bureaucratic expectations. Columns segregated languages and scripts; rules of alignment guaranteed that translations met at the eye; marginalia and interlinear glosses accelerated reading; classification hierarchies signaled to the user how to navigate; and the physical size of a volume and the quality of its paper and binding told us whether it was meant for the desk or for the saddlebag [21]. Every such feature mediated between page and practice. In this sense, the dictionaries belonged to the same family of artifacts as forms, seals, gazetteers, and maps. They were bureaucratic media that allowed the late imperial state to imagine itself as a multilingual whole.

10. Conclusions

The passage from the 'Phags-pa experiments in the Yuan to the domain-based bilingualism of the Ming and the Manchu-centered polyglossia of the Qing maps a transformation in how the late imperial state handled difference. What began as a regime for pronouncing foreign sounds became a curriculum for pairing meanings and drafting documents, and then, at its most elaborate, became a machine for integrating scripts and standardizing names [1]. Throughout, the books were not passive shelves of data. They shaped their users by habituating them to look up the world through particular grids and to slot words into routine moves. They thus converted plurality into procedure.

This is why it is misleading to segregate dictionaries from the practical arts of governing. The Ming *Huayi yiyu* did not merely record what Mongolian or Tibetan words corresponded to Chinese lexemes; it also instructed the clerk where to place those words in letters that enacted relations among polities [3]. The Qing *Pentaglot* (*Wuti Qing wen jian*) did not merely allow users to gloss a term across languages; it also trained a Manchu-literate officer to pronounce Tibetan and Turkic words according to a shared code, and it did so at a time when the empire sought to fold Inner Asian spaces into a centralized apparatus [7]. In both examples, lexicography disclosed itself as a political technology.

The vitality of that technology is measurable beyond the court. *Zazi* and *Lao qida* handbooks attended to the pragmatics of exchange and travel, extending standardization into the world of inns, shops, and caravans [9,10]. The Joseon *Nogeoldae* and *Bak Tongsa* taught Beijing speech to foreign students, transforming the Ming and Qing conventions of dialogic interaction into a portable curriculum that could be transplanted into other languages while retaining its narrative skeleton [11,12]. These civilian and pedagogical streams did not duplicate the courtly corpus; they augmented it. Together, all three produced a multilingual infrastructure with cross-cutting lines of circulation: documentary and conversational, symbolic and technical, as well as metropolitan and frontier.

A brief coda points to late-Qing continuities. Institutions such as the Tongwen Guan (同文館, established 1862) and the Translation House of the Kiangnan Arsenal (江南製造局翻譯館, established 1868) extended the lexicographic logic of governance into new epistemic terrains by standardizing Chinese terminology for mathematics, engineering, international law, and the natural sciences [34]. In another respect, multilingual administrative practices persisted in frontier regions such as the Jirim League in the eastern part of Inner Mongolia, where late-Qing attempts to coordinate Manchu, Mongolian, and Chinese literacies made language policy itself a site of political contestation [35]. Their bilingual and technical glossaries did not replace the earlier polyglot infrastructure; they retooled it for Sino-Western knowledge transfer, making the management of equivalence a pillar of modernization.

This perspective resonates with recent scholarship that foregrounds translation as a constitutive practice of early modern and modern China [36,37]. By centering the page-level technologies of equivalence—classification, templates, and transliteration—this essay complements that literature with a material history of the dictionaries that made multilingual governance legible, speakable, and administratively actionable.

If a final moral may be drawn, it is that the history of polyglot lexicons and encyclopedic word-books in late imperial China cannot be reduced to a parade of titles or to a nostalgic celebration of learned culture. It is a history of moving parts in working systems. The Yuan supplied phonetic scaffolding and, in the Sino-Mongol glossary, an early semanticization [1]; the Ming set a curriculum that bound words to procedures [3]; the Qing built a Manchu pivot that multiplied languages while preserving a shared discipline of form [7]. Through these books, a multilingual empire learned to pronounce itself, to define its terms, and to inscribe its spaces.

This historical trajectory from phonetic key to administrative database offers a powerful lens through which to view contemporary statecraft. The challenges of governing multilingual populations, standardizing data across disparate systems, and deploying language as a tool of political integration have not vanished; they have merely migrated to new technological platforms. The story of these polyglot lexicons is therefore not just a history of the book, but a foundational chapter in the longer story of how states seek to manage diversity through the codification of knowledge—a process as relevant in the age of algorithms as it was in the age of empires.

Funding: This research was funded by Academia Sinica (AS-NTU-114-11) and the National Science and Technology Council of Taiwan (111-2410-H-001-005-MY3).

Institutional Review Board Statement: Not applicable.

Informed Consent Statement: Not applicable.

Data Availability Statement: No new data were created or analyzed in this study. Data sharing is not applicable to this article.

Conflicts of Interest: The author declares no conflict of interest. The funders had no role in the design of the study; in the collection, analyses, or interpretation of data; in the writing of the manuscript; or in the decision to publish the results.

References

1. Jia, J.Y.; Zhu, F. (Eds.) *Menggu Yiyu Nüzhen Yiyu Huibian*, 1st ed.; Tianjin Guji Chubanshe: Tianjin, China, 1990.
2. Tailian Guofeng Chubanshe (Ed.) *Menggu Yiyu Liu Zhong*; Tailian Guofeng Chubanshe: Taipei, Taiwan, 1974.
3. Huo, Y.J. *Huayi Yiyu*; Shumu Wenxian Chubanshe: Beijing, China, 1988.
4. Apatóczky, A.B. (Ed.) *Yiyu: An Indexed Critical Edition of a Sixteenth-Century Sino-Mongolian Glossary*; Global Oriental: Folkestone, UK, 2009.
5. Mao, Y.Y. *Dengtan Bijiu, Yiyu*; Tailian Guofeng Chubanshe: Taipei, Taiwan, 1974.
6. Qing Emperor Gaozong. *Yuzhi Zengding Qing Wen Jian*; Shijie Shuju: Taipei, Taiwan, 1986.
7. Qing Emperor Gaozong. *Yuzhi Wuti Qing Wen Jian*; Minzu Chubanshe: Beijing, China, 1957.
8. Fu, H. *Qinding Xiyu Tong Wen Zhi*; Toyo Bunko: Tokyo, Japan, 1961–1963.
9. Anonymous. *Man-Han Tongwen Xinchu Duixiang Menggu Zazi*; Da Mo Chang Wen Cheng Tang: Beijing, China.
10. Zhuang, J.F., Ed. and Translator; *Qingyu Lao Qida*, rev. ed.; Wen Shi Zhe Chubanshe: Taipei, Taiwan, 1984.
11. Kyujanggak (Ed.) *Nogeoldae Eonhae*; Kyujanggak: Seoul, Republic of Korea, 2003.
12. Kyujanggak (Ed.) *Bak Tongsa Eonhae*; Kyujanggak: Seoul, Republic of Korea, 2004.
13. Nakano, M. *A Phonological Study in the Phags-Pa Script and the Meng-Ku Tzu-Yün*; Australian National University Press: Canberra, Australia, 1971.
14. Coblin, W.S. *A Handbook of 'Phags-Pa Chinese*; University of Hawai'i Press: Honolulu, HI, USA, 2007.
15. Rossabi, M. Communications and “Communicators” in the Yuan and Early Ming. *J. Chin. Hist.* **2025**, *9*, 271–285. [[CrossRef](#)]
16. Chen, Y. Study on ‘Huitong guan’ as the Ming Central Hospitality Facility [Mindai chūō shukuhaku shisetsu “Kaitōkan” shōkō]. *Nishōgakusha Daigaku Jinbun Ronsō* **2021**, *107*, 73–93.
17. Chu, T.C. The Siyi Guan and Its Students in Ming Dynasty [Ming Dai De Si Yi Guan Yu Yi Zi Sheng]. Ph.D. Dissertation, Department of History, National Chengchi University, Taipei, Taiwan, 2020.
18. Kuribayashi, H. *A Study of the Huayi Yiyu [Kai Yakugo (Kōshu Hon) no Kenkyū]*; Shokado: Kyoto, Japan, 2019.
19. Rykin, P. The Sino-Mongolian Glossary Dada yu/Beilu yiyu and the Problem of its Dating. In *Central Asian Sources and Central Asian Research*; Reckel, J., Ed.; Universitätsverlag Göttingen: Göttingen, Germany, 2016.
20. Jiang, Q. *A Study of the Qing Wen Jian of Kangxi [Kangxi Yuzhi Qing Wen Jian Yanjiu]*, 1st ed.; Yanshan Chubanshe: Beijing, China, 2001.
21. Jiang, Q. *A Study of the Pentaglot Dictionary [Yuzhi wuti Qing Wen Jian Yanjiu]*, 1st ed.; Yanshan Chubanshe: Beijing, China, 2017.
22. Millward, J.A.; Dunnell, R.W.; Elliott, M.C.; Forêt, P. (Eds.) *New Qing Imperial History: The Making of Inner Asian Empire at Qing Chengde*; Routledge: New York, NY, USA, 2004.
23. Schorkowitz, D.; Chia, N. (Eds.) *Managing Frontiers in Qing China: The Lifanyuan and Libu Revisited*; Brill: Leiden, The Netherlands, 2017.

24. Deng, K.; Xiu, L. Interpreting for Governance: The Manchu Conceptualization of “tongshi” in Early Qing China. *Acta Orient. Acad. Sci. Hung* **2025**, *78*, 151–174. [[CrossRef](#)]
25. Dai, Y.Z. *A Study on Zazi in Huizhou during the Ming and Qing Dynasties* [Ming Qing Huizhou Zazi Yanjiu]; Shanghai Jiaoyu Chubanshe: Shanghai, China, 2017.
26. Kim, T.J. *History of Literary Chinese Literature in Korea* [Chaoxian Hanwen Xue Shi]; Zhang, L.G., Translator; Shehuikexue Wenxian Chubanshe: Beijing, China, 1996.
27. Kyujanggak (Ed.) *Chephae Mongeo/Mongeo Ryuhae*; Kyujanggak: Seoul, Republic of Korea, 2006.
28. Li, T.Z. *A Linguistic Study of Four Versions of Laoqida* [Lao Qida si Zhong Banben Yuyan Yanjiu]; Yuyan Chubanshe: Beijing, China, 2003.
29. Jin, J.K. The Messengers of Knowledge/Information in Eighteenth-and Nineteenth-Century East Asia: Translating Officials [18 19 segi dong-Asia wa jisig jeongbo ui mesinjeo, Yeoggwan]. *Hangukhanmunhag-Yeongu* **2011**, *47*, 105–138.
30. Nishida, T. *A Study of the Tibetan-Chinese Vocabulary Hsi-fan-kuan I-yu, an Introduction to Tibetan Linguistics* [Seibankan Yakugo no Kenkyū: Chibetto Gengo Gaku Josetsu]; Shokado: Kyoto, Japan, 1970.
31. Vedal, N. Literati of the Garrisons: The Civil Service Translation Examination and Manchu Literary-Intellectual Culture in the Eighteenth and Nineteenth Centuries. *J. Asian Stud.* **2023**, *82*, 339–361. [[CrossRef](#)]
32. Bian, H.; Saarela, M.S. *The Manchu Mirrors and the Knowledge of Plants and Animals in High Qing China*; Harvard University Press: Cambridge, MA, USA, 2025.
33. Miyake, S.; Matsukawa, T. (Eds.) *Facsimile and Study of Huayi Yiyu (Four Versions of Xifan Yiyu and One Version of Guoluo Yiyu)* [Kai Yakugo (Seiban Yakugo Yonshu Karo Yakugo Isshu) Ein to Kenkyu]; Shokado: Kyoto, Japan, 2015.
34. Elman, B.A. *On Their Own Terms: Science in China, 1550–1900*; Harvard University Press: Cambridge, MA, USA, 2005.
35. He, J. *Ruling the Mongols of Manchuria: Language, Literacy, and Power in Late Qing Borderlands*; Amsterdam University Press: Amsterdam, The Netherlands, 2025.
36. Nappi, C. *Translating Early Modern China: Illegible Cities*; Oxford University Press: New York, NY, USA, 2021.
37. Wright, D. *Translating Science: The Transmission of Western Chemistry into Late Imperial China, 1840–1900*; Brill: Leiden, The Netherlands, 2000.

Disclaimer/Publisher’s Note: The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.