Language Policy and the Typology of Scripts

David Bradley
La Trobe University

1. Typology of Orthographies

The main orthographic alternatives used in South, Southeast and East Asia can be classified into three: the character-syllabic Chinese and other related Sinospheric systems, the segmental Indospheric systems which write a syllable in a unitary way (not necessarily in sequence, and with an unmarked, inherent vowel), and alphabetic systems (mainly romanisations and Arabic systems) which write segments more or less in sequence. A contrary example is a script invented for Hmong described in Smalley et al. (1990) in which the vowel and tone of each syllable is written first, before the initial consonant. Among the non-Chinese Sinospheric systems, there is a general tendency to move to syllabic systems derived from a character system; there are also some independently invented syllabic systems, such as one for Lisu invented by Wang Renpo in the 1920s.

Some scholars have suggested that the Chinese system is appropriate for a language with very limited morphology and mainly single-syllable words, and this may be true to some extent. However, recent reconstructions of Proto-Sinitic suggest a certain amount of early affixation, later lost or replaced by additional tones or different segmental developments; is this monosyllabisation also to be attributed to the writing system? Conversely, with phonetic erosion compounding has become extremely frequent, especially in Mandarin where the erosion is most extreme.

What is not in doubt is that the less phonetic nature of the character systems means that they are inherently more stable. Sound change does not make the writing system archaic, and people whose spoken varieties become very different can continue to write with the same characters and feel a cultural unity. Thus the Han Chinese, despite millenia of divergence and lack of mutual intelligibility, still form one cultural and linguistic group; it is only in this century that the widespread use of Mandarin has also provided a spoken link among them. This cultural unity, largely derived from the writing system and the nature and status of the texts written with it, also accounts for the much higher level of linguistic entities referred to by the Chinese terms yuyan (usually translated as 'language') and fangyan (usually translated as 'dialect'). This tendency to link related groups together under one linguistic, cultural and political category has also been extended to the minorities of China.

The Chinese writing system was adopted by speakers of various genetically unrelated languages with polysyllabic stems and extensive affixation morphology, such as Japanese and Korean, as well as for typologically more similar but also unrelated Vietnamese. It was initially used to write Chinese, but when it came to be used to represent Japanese and Korean the typological differences eventually forced the development of syllable-based phonetic systems as well.

The Indospheric writing system spread mainly through the effects of Indian cultural influence, direct or mediated through another Indian-influenced society.
Most scholars believe that all such modern systems ultimately derive from Brahmi, which itself is usually ascribed to a Semitic (usually Aramaic) origin. In mainland Southeast Asia, the Mon script was spread first to the Burmans and subsequently to the Shans and others. Similarly, the source of the Thai script appears to be Khmer. This shows again that scripts spread by contact, and not according to genetic relationships among languages. In some cases, more recently created scripts for minorities have used a local majority Indospheric system: Devanagari, Bengali, Burmese, Thai and so on.

The Arabic system spread with the advent of Islam; while the romanised and related systems came with the European colonialists and were usually created by Christian missionaries. Over the last couple of hundred years, such scripts have been created for many minority groups who would traditionally have used one of the dominant majority languages if they needed to write.

In quite a few cases a new script was created for a language already written with another. Sometimes, as in the case of Vietnamese, Javanese and Malay, an earlier Sinospheric, Indospheric or Arabic script for a majority language was replaced by a romanisation. An earlier script may be replaced rapidly or slowly, and be preserved or completely forgotten; or may persist in restricted domains, as with Jawi (Arabic script for Malay) in Malaysia. Where more than one new script was created, the eventual choice was usually determined by nonlinguistic factors. Alternatively, more than one script may coexist among different religious denominations, in different political entities or within other subgroups or domains.

Choice of script is often based on historical, political, religious and other nonlinguistic factors. This choice has various major consequences: the cognitive consequences of using any writing system, and in particular a segmental, segmental-syllabic or character-syllabic system; the sociolinguistic consequences in terms of group identity; the historical linguistic consequences of greater stability and possible eventual development of monolingual diglossia; and so on. Scripts may evolve, be substantially reformed or be completely replaced; speakers may react to such evolution, reform or shift in various ways depending on nonlinguistic as well as linguistic factors. In particular, political change often leads to orthographic change. An established script, no matter how archaic or inadequate, may be accepted as part of a group’s identity; a new or reformed script which is phonologically much more adequate may be rejected. On the other hand, a script which is not well-established nor widely disseminated may more easily be replaced. Script selection and codification is usually the result of some kind of external influence, and usually has social consequences; script reform is more likely to be internally generated and related to political changes.

The level of linguistic unit for which a script is developed, and the taxonomy of linguistic relationship and intelligibility, also differs according to local situational factors. The level of classification of groups into linguistic entities may differ across borders. For example, in China there is a tendency to group distinct but related groups into larger nationalities, parallel to the situation of the Han Chinese; while in India the tendency is to local atomisation, parallel to the great diversity and localisation of castes. Often the same group may be classified as
distinct in one country, and have its own orthography and other trappings of a separate identity; but is included within a larger group in another country and thus lacks a separate political and orthographic identity there.

Apart from China, where romanisations became established as the official preference in the 1950s, governments prefer new scripts for minorities to use the same system as the national language, whether this is phonologically appropriate or not. This has the virtue that there can be some transfer of learning from one script to the other, and favours national unity. Christian missionaries, by contrast, have tended to prefer to use romanisations; these are easier for the missionaries, and easier to print; but do not assist in learning a national language not written in a romanisation. Minorities using romanised scripts may therefore feel more distinct, other things being equal.

In this paper, most of the examples will be drawn from Sino-Tibetan and especially Tibeto-Burman languages of East, Southeast and South Asia; of course parallel examples could also be chosen from other areas and other genetic groups.

2. Script Development Strategies

In most cases the early versions of a script contain various inadequacies; substantial revision may be desirable, but this may not take place if the script becomes well-established in an initial form. In many new segment-based Indospheric, romanised or Arabic scripts some existing symbols are used with different values, and new combinations of existing symbols and/or some new diacritic symbols are used. In a few cases more substantial changes are made; some examples are discussed in 2.1 below. Some examples of the more extensive creativity required to devise a character-based orthography are discussed in 2.2 below.

Speakers can be very attached to an existing script for their language, and may react very negatively to attempts to reform or replace it. The introduction of a new script for a nonliterary language may also encounter problems if the speech community is not motivated to develop literacy in their own language.

Most Tibeto-Burman and other languages of the Southeast and East Asian linguistic areas are tonal, while no segmental script provides a convenient means to represent tones; the strategies for doing so are discussed in section 3 below.

2.1 Modified Romanisations

In some cases, missionary scripts have used fairly radical adaptations of a romanisation. One such was first developed for Lisu in China in the early 1920s and is usually known as the 'Fraser' script after James Outram Fraser, the missionary who led its development. In China this script is known as 'old Lisu' (in contrast to 'new Lisu', the 1950s romanisation). It uses roman capital letters, upright and inverted, often in radically different values from the usual ones; for details see Bradley (1979) and Bradley & Kane (1981). One interesting feature is that the vowel /a/ is inherent in any consonant or consonant cluster, unless some other vowel follows it; and that the letter A indicates an initial glottal stop with the inherent vowel /a/. A similar orthography was also developed for Naxi, but is not
now used. The Fraser Lisu script continues to be used widely by Christian Lisu in China, Burma, Thailand and India, and has official status in the one Lisu Autonomous Prefecture in China, Nuijiang in northwestern Yunnan.

An even stranger-looking missionary attempt was first developed by Samuel Pollard and his Chinese associate Stephen Lee beginning in 1904 for Miao; it is usually known as the Pollard script. The development of this script and its extension to various varieties of Miao by Pollard and others is fully discussed in Enwall (1994). The basic principle is not unlike that of Korean Han'gul: a phonetic representation of the syllable composed of two parts, the initial consonant and the rhyme. In the Pollard script the initial consonant is large and central, and the rhyme is much smaller and peripheral; the position of the rhyme relative to the consonant indicates the tone of the syllable. Thus, the whole is designed to look like a character and represent an entire syllable. Related consonants (e.g. voiceless unaspirated, aspirated, prenasalised and voiced stops or affricates in the same place of articulation) are versions of the same basic consonant slightly modified in parallel fashion.\textsuperscript{1} Versions of the Pollard script were developed for five Tibeto-Burman languages of Yunnan, four of which are now included in the Yi nationality and one in the Hani nationality. What the missionaries called Nosu, Laka and Köpu are now known as Hei Yi, Gan Yi and Gepo in Chinese, and are all classified as subvarieties of Eastern Yi spoken in northeastern Yunnan. What the missionaries called Eastern Lisu is now known as Lipo, and included by the Chinese within the Central Yi variety spoken in north central Yunnan.\textsuperscript{2} Finally, Kadu of south central Yunnan (now known in the Chinese literature as Kaduo) is now classified as one of the subvarieties of Hani. In Yunnan, many Christian Miao and much smaller numbers of some of the groups in the Yi nationality still use the Pollard script, but in a modified form which writes the vowels (still in a smaller size) on the line after the consonant and then adds a further small letter to indicate the tone; in Guizhou some Christian Miao still use a form of the original Pollard script alongside 1950s romanisations.

2.2 From Character to Syllable

In Chinese, the character represents an entire syllable with its initial, rhyme and tone. Some characters are originally pictographic, but most characters consist of a ‘radical’ (semantic element) plus a ‘phonetic’; the traditional Chinese classification has four other categories of characters. Homophones are usually written differently. Though the number of ‘phonetics’ is much greater than the number of syllables, and the ‘phonetics’ are no longer phonetically very accurate (if they ever were!), this principle of combination contains the seeds of a shift to a syllabic system.

Such a shift has of course taken place in a partial way in Japanese, while also retaining over 2,000 kanji with various Chinese and Japanese readings. The

\textsuperscript{1} For example, the prenasalised consonants are all represented by adding a small prefixed C (which as a consonant represents /n/) before the main consonant; aspirated consonants are all represented by adding a small postscript ‘ after the main consonant.

\textsuperscript{2} Linguistically this is more closely related to Lisu, as the missionary name suggests.