

The Akshara as a Graphematic Unit

Arvind Iyengar

Abstract. The written unit called an akshara is emblematic of several South and South-East Asian writing systems. In its most basic form, an akshara comprises a combination of free and bound written elements, and typically denotes a phonological vowel or a consonant-vowel sequence. Yet, there remain several open questions on the definitional limits of an akshara, leaving the concept somewhat fuzzy. For instance, what phonological values could an akshara potentially have? Conversely, is the phonological value of a written entity sufficient grounds to determine its status as an akshara? Further, if a particular written sequence is pronounced differently in different languages written in the same script, does this impact on the aksharic status of the written entities? Finally, to what extent is the akshara an inherent characteristic of certain writing systems? Is its existence in any way determined by externally imposed orthographic norms and practices? This paper addresses these questions and more in attempting to constrain and define the akshara.

1. Introduction

1.1. Overview

Since the 1990s, the popularization of the terms *abugida* and *alphasyllabary* (Bright, 1999; Daniels and Bright, 1996) has coincided with an increase in scholarly interest on the unit of writing known as the *akshara*.¹ Originating as a phonological concept (Rimzhim, Katz, and Fowler, 2014), the akshara is now commonly understood as a written unit, often characterized as a graphic syllable (Salomon, 2007) or orthographic syllable (Sproat, 2000). The prototype of the akshara is a combination of ‘free’ and ‘bound’ written elements that corresponds to a phonological [V] or [CV] sequence (Gnanadesikan, 2017; Nag and Perfetti, 2014). The

Arvind Iyengar  0000-0002-7303-1524

Discipline of Linguistics, School of Humanities, Arts and Social Sciences, University of New England, Armidale NSW 2351, Australia

E-mail: arvind.iyengar@une.edu.au

1. Sanskrit /əkṣərə/; English usu. /'ʌkʃəɹə/.

Y. Haralambous (Ed.), *Grapholinguistics in the 21st Century 2022. Proceedings*
Grapholinguistics and Its Applications (ISSN: 2681-8566, e-ISSN: 2534-5192), Vol. 10.
Fluxus Editions, Brest, 2024, pp. 419–436. <https://doi.org/10.36824/2022-graf-iyen>
ISBN: 978-2-487055-06-3, e-ISBN: 978-2-487055-07-0

akshara is considered iconic of the Brahmic writing systems prevalent across South and South East Asia (Gnanadesikan, 2021), due to which they are sometimes termed *aksharic* systems.²

At the same time, much of the research thus far on aksharic writing has approached the subject matter from a cognitive, psycholinguistic and educational lens (Joshi and McBride, 2019; Winskel and Padakanaya, 2014). Critiques explicitly addressing the graphematic aspects of the akshara are rare. In the words of Joyce and Meletis (2021), existing research on aksharic writing has focused primarily on its processing fits, and less on its linguistic fits. As a result, the akshara is yet to be rigorously analyzed using contemporary grapholinguistic approaches. Moreover, and akin to the term *grapheme* (Daniels, 2018), authors over the years have differed in their interpretation of the akshara, in the process adversely impacting the term's semantic connotations and theoretical consistency.

In light of the above, there is a need to constrain, refine and define the conceptual scope of an akshara, in order to ensure terminological precision and epistemological robustness. With this aim in mind, I propose in this paper certain criteria for labelling a unit of writing an akshara, while also highlighting aspects requiring further scrutiny.

1.2. Terminology

Although the study of writing systems is currently a “hot” discipline (Sproat, 2018, p. 269), it is also a fairly recent development. As a result, it invariably grapples with issues of vacillating and often competing terminology for one and the same concept. Conversely, a given term may be used in a polysemous manner by different authors, in the process leading to inadvertent ambiguity. Against this background, I provide below a summary of the terms used in this paper and their definitional scope.

In the literature, the basic unit of a writing system has been variously designated as *sign*, *symbol* or *letter*. However, as Meletis (2020, p. 78) observes, these appellations “lead double lives as lay terms and quasi-technical terms”. Equally ambiguous is the term *grapheme*, due to which it has been notably rejected by Daniels (2018). Consequently, I label the basic unit of writing a *graph*, and consider it to be the written counterpart to a phonological *segment*. Individual or sequences of

2. Sanskrit /əkṣərikə/; English usu. /'ɑ:kʃəɹɪk/. The term serves as the adjectival counterpart of the noun *akshara*. The en-Latn spelling |aksharic| is modelled on |Sanskritic|, |Vedic|, |yogic|, |karmic| and the like, and is becoming increasingly prevalent in recent anglophone scholarship (Gnanadesikan and Judson, 2021; Vaid, 2022).

graphs are enclosed in pipes (| |) where needed. I define *script* as a superset or macro-inventory of graphs used to graphize or create a written form for a spoken language. When coupled with a particular language, the script-language pair constitutes a *writing system* (henceforth ‘WS’). The conceptualization of script and language as two distinct but essential components of a WS follows the lead of Meletis (2020), Neef (2015), and Weingarten (2013). Accordingly, Latin (or Roman), Arabic, Devanagari and Canadian Syllabics are scripts, in that they represent inventories of graphs used to graphize one or more spoken languages. When paired with a spoken language, they form WSs such as German–Latin, Kashmiri–Arabic, Nepali–Devanagari or Inuktitut–Syllabics. It follows that the same script may be used by multiple languages, as in German–Latin and Swahili–Latin, while a particular language may be written in more than one script, as in Kashmiri–Arabic and Kashmiri–Devanagari. Each unique combination of script and language results in a distinct WS. To ensure compactness, I adopt the convention followed by the website *ScriptSource* and specify a WS using the ISO codes for its constituent language and script. Thus, German–Latin, Kashmiri–Arabic, Nepali–Devanagari or Inuktitut–Syllabics may be designated de-Latn, ks-Arab, ne-Deva and iu-Cans, respectively. Assigning WSs binomial monikers of this kind ensures componential transparency while also highlighting the subtle but crucial distinction between a script and WS on the one hand, and between a script and ISO tag for greater specificity (Raymond, 2020). For instance, ne-Deva-IN may be used to denote Nepali written in Devanagari in India, and de-Latn-1996 to refer to German written according to the 1996 spelling reforms (Johnson, 2005).

From the above, it is evident that the *graphetic* properties of a WS, entailing matters of graph size, shape, position and other typological considerations (Altmann and Fengxiang, 2008; Meletis, 2020, p. 393), only concern the script component. This makes a graphetic analysis of a script or WS analogous to the phonetic analysis of a spoken language. In contrast, the *graphematics* or *graphematic system* of a WS refers to the underlying correspondences between the graphs of its script component on the one hand, and the phonological segments, morphemes or other units of its comparable to the phonological dimension of a spoken language. For convenience, I denote ‘graphematic’ and ‘phonological’ by $|\gamma|$ and $|\phi|$, respectively (Haralambous, 2019). *Orthography* denotes the explicit rules or implicit conventions that constrain graphematic possibilities and, in some cases, override them (Honda, 2021; Meletis, 2020; Neef, 2015). This makes orthography an optional component of a WS which, when present, may have graphematic as well as sociolinguistic ramifications (Bunčić, Lippert, and Rabus, 2016; Joyce and Meletis, 2021). Table 1 summarizes the scope of the terms described.

Other key terms appearing in this paper concern the various types and subtypes of WSs in question. All WSs discussed herein are *segmen-*

TABLE 1. Dimensions of writing systems

<i>Term</i>	<i>WS component(s) involved</i>	<i>Objects of analysis</i>
<i>graphetic</i>	script	typographic & calligraphic characteristics, design & material properties...
<i>graphematic</i>	script + language	γ - ϕ correspondences, allograph distribution...
<i>orthographic</i>	script + language	rules or conventions governing (allo)graphs, spelling, punctuation...
<i>sociolinguistic</i>	script + language	WS used by who, when, how, for what purpose...

taries, in that their minimum grain size is the ϕ -segment (Gnanadesikan, 2017). More precisely, all WSs mentioned in this paper are vowelised segmentaries, in that they generally denote individual consonant and vowel ϕ -segments (henceforth [C] and [V], respectively) using distinct graphs or graphetic elements. That said, these WSs may also comprise graphs corresponding to sequences or clusters of ϕ -segments, such as [CV] or [CCV]. Where required to describe graphs based on their phonological values, I refer to them as [C]-grams, [V]-grams, [CV]-grams and so on. Regardless of its phonological value, a graph or written element may be graphetically central, as in Latin-script $|e|$, or graphetically peripheral, as in the so-called acute accent $|\acute{e}|$. In other words, written elements may be graphosegmental or graphosubsegmental (Meletis, 2020, 97ff; Osterkamp and Schreiber, 2021, 172 footn. 2).

In the grapholinguistic literature, reference is most often made to three subtypes of vowelised segmentary: *alphabet*, *alphasyllabary* and *abugida*. These terms are defined by Daniels (1996) thus:

ALPHABET

a type of writing system that denotes consonants and vowels

ALPHASYLLABARY

a writing system in which vowels are denoted by subsidiary symbols not all of which occur in a linear order (with relation to the consonant symbols) that is congruent with their temporal order in speech

ABUGIDA

a type of writing system whose basic characters denote consonants followed by a particular vowel [...]

(*ibid.*, p. xxxix)

Of these, the terms *alphasyllabary* and *abugida* have often been used as near-synonyms by scholars and laypersons alike. However, both Daniels (*ibid.*, 4, footnote) and Bright (1999) have clarified the conceptual dis-

creteness of the two terms. According to Bright (ibid., p. 45), an alphasyllabary denotes [V]s with distinct allographs in complementary distribution. Following a [C]-gram, an alphasyllabary indicates a [V] using a specific set of allographs, known variously as secondary, dependent or bound graphs. In most other positions, a [V] is indicated with another set of allographs, termed primary, independent or free graphs. It should be noted that the ‘secondary’, ‘dependent’ or ‘bound’ nature of a graph does not require it to be graphetically diminutive or subsegmental. With reference to Table 1, the dependent or bound nature of a graph is a graphematic-orthographic restriction on its occurrence, while its γ-(sub)segmentality is a graphetic aspect. It is entirely possible for a graph in a WS to satisfy one of these conditions but not the other. In the context of an alphasyllabary, bound [V]-grams may well be graphetically central.

In contrast to an alphasyllabary, Bright (ibid., p. 45) lays down the defining characteristic of an abugida as the presence of an identifiable subset of graphs with phonological value [CV₀], where [V₀] is a so-called inherent, default or unwritten vowel φ-segment. The label *inherent* stems from the fact that the graphs in question have no overt graphetic element corresponding specifically to [V₀]. The inherent vowel in a [CV₀]-gram may be overridden or suppressed by various graphematic and/or orthographic means, depending on the WS in question.

2. The Prototypical Akshara

The distinction between an alphasyllabary and an abugida has been further distilled by Gnanadesikan (2017) as follows:

The definition of *alphasyllabary* focuses on the arrangement of the signs into syllable-like structures (the *akṣara*), while the definition of *abugida* includes the use of an ‘inherent’ or ‘default’ vowel. Thus, the Daniels and the Bright typologies disagree as to whether to give priority to *what* [phonological] segments are being represented or to how the representation of those segments is arranged.

(ibid., 9–10, emphasis in original)

Gnanadesikan’s observation underscores Bright’s (1999) clarification on the labels alphasyllabary and abugida being indicative of distinct graphematic behaviors. Also present in her observation is a reference to the written unit known as the *akṣara*.³ The prototypical akshara has a

3. The Latin-script spelling |akṣara| used by Gnanadesikan (2017) is based on conventions of the International Alphabet for Sanskrit Transliteration (IAST). Per the terminology and conventions outlined in Section 1.2, the IAST may be considered an instantiation of the Sanskrit–Latin writing system (sa-Latn-iastr).

phonological value of [V] or [CV]. Among the latter subtype, [C] is typically indicated by an independent or free written element and [V] by an element that is graphematically dependent on, or bound to, the free [C]-gram as its base. Thus, the prototypical akshara satisfies the conditions for alphasyllabic writing. Included in the prototype are graphs with a phonological value of [CV₀], where [V₀] is the inherent or default vowel ϕ -segment that has no distinct graphetic manifestation. Consequently, the prototypical akshara also qualifies as abugidic in nature. Table 2 presents a selection of aksharas with phonological values of [V] and [CV] from the Hindi–Devanagari, Tibetan–Tibetan and Sinhala–Sinhala WSs.

TABLE 2. Aksharas denoting [V] and [CV]

hi-Deva			bo-Tibt			si-Sinh		
अ	इ	उ	ཨ	ཨི	ཨུ	අ	ඉ	උ
na	ni	nu	na	ni	nu	na	ni	nu
न	नि	नु	ན	ནི	ནུ	න	නි	නු

In sum, the akshara is typical of WSs that are alphasyllabaries as well as abugidas, due to which such WSs—including the ones in Table 2—are known as *aksharic* systems (Gnanadesikan, 2021, p. 304; Iyengar, 2023).

3. Beyond the Prototype

Whereas the prototypical or *minimal* akshara is relatively straightforward to identify and narrow down on, there is as yet no scholarly consensus on what the definitional limits of an akshara should be. That is, the boundaries of the *maximal* akshara remain fuzzy. This prompts the question of how much a written unit can depart from the aksharic prototype before it is considered to no longer exhibit the essential properties of an akshara. To address this question, we first need to scrutinize the characteristics used to identify the minimum or lower limits of an akshara, and verify whether they prove adequate for identifying its maximum or upper limits. If not, how can we best plug any conceptual gaps that exist, in order to devise an epistemologically and theoretically robust definition of the akshara?

3.1. Written Elements and Their Phonological Values

The prototypical akshara is defined as a written unit that corresponds either to a vowel ϕ -segment [V] or a consonant-vowel sequence [CV]. However, several WSs commonly portrayed as aksharic also feature written units corresponding to a series or cluster of consonants followed by a vowel, such as [CCV], [CCCV] and so on. The question that emerges here is: how many [C]s can precede a [V] within the context of a single akshara? Put differently, is there an upper limit on the number of pre-vocalic [C]s that a single akshara can denote? Furthermore, are these upper limits WS-specific, or can they be applied to multiple WSs? In brief, can an akshara be plausibly and reliably defined as written unit denoting no more than a certain number of [C]s in sequence, followed by a single [V]? Or are there instances of putative aksharas terminating in a [C], or denoting only a [C]? If yes, is the prototype of the akshara—a written unit corresponding to [V] or [C*V]—theoretically tenable? Table 3 addresses some of these questions by providing examples of written sequences with phonological values of [CCCV] and [C] from various aksharic WSs.

TABLE 3. Written sequences from aksharic WSs denoting [CCCV] and [C]

Row 1	hi-Deva [tʃjə] त्स्य	kn-Knda [tʃjɐ] ತ್ಸ್ಯ	si-Sinh [tʃja] ත්ස්‍යත්සිය
Row 2	ml-Mlym [n] ന	bn-Beng [t] ত	mni-Mtei [k] ꯏ

From an emic perspective, each of the hi-Deva and Kannada–Kannada (kn-Knda) examples in Row 1 of Table 3 would usually—although not always—be identified as an individual akshara, in the process legitimizing the existence of aksharas with value [CCCV] in these WSs. Such aksharas are also attested in homoscriptal WSs, such as Marathi–Devanagari (mr-Deva) or Tulu–Kannada (tcy-Knda), unless explicitly proscribed by WS-specific orthographic rules or conventions. Unlike hi-Deva and kn-Knda, however, si-Sinh does not provide for [tʃja] to be written in a manner generally identifiable as a single akshara. In fact, si-Sinh allows for [tʃja] to be transcribed in two ways. Of these, [ත්ස්‍ය] may be emically interpreted as comprising two aksharas, while [ත්සිය] might well be described as having three aksharas. The issue of homophonous heterography will be examined further in Section 3.2. For now, the examples from Row 1 of Table 3 suggest that evidence is sparse for a univer-

sal upper limit on the number of pre-vocalic [C]s that a single akshara can denote. To the extent such limits can be identified, they are invariably WS-specific. Consequently, the phonological criterion of number of pre-vocalic [C]s appears unreliable for purposes of defining an akshara.

Row 2 of Table 3 provides examples of graphs denoting only [C] from Malayalam–Malayalam (ml-Mlym), Bengali–Bengali (bn-Beng) and Manipuri–Meetei Mayek (mni-Mtei). Evidence from the grapholinguistic and computational literature points to these graphs being considered independent units in their respective WSs (Constable, 2004; Everson, 2007; Mohanan, 1996; 2007). Based on this evidence, there appears to be a strong case for the graphs in question to be deemed aksharas in their own right, in the process establishing the existence of aksharas that are simple [C]-grams. Such a conclusion sits uneasily with the prototype of an akshara described in Section 2 as a written unit whose phonological value is [V]-final.

Among the properties of the prototypical akshara outlined in Section 2, the [CV] subtype was characterized by an independent or free written element corresponding to [C], optionally appended with a dependent or bound element denoting [V]. This begs the question: are there instances of prospective aksharas where the *bound* element denotes [C]? If yes, can the bound element in an akshara also have other phonological values, such as ϕ -suprasegmentals? Table 4 provides examples to this end.

TABLE 4. Bound elements in aksharic WSs denoting [C] and ϕ -suprasegmentals

Row 1a	kn-Knda 𑌕 𑌕		te-Telu 𑌕 𑌕		sa-Gran 𑌕 𑌕	
Row 1b	rja: 𑌕𑌕		(ʷ)o:m 𑌕𑌕		ʈsjeiᶱ(i) 𑌕𑌕𑌕𑌕	
Row 2a	or-Orya 𑌕 <i>vowel nasalisation</i> 𑌕		my-Mymr 𑌕 <i>creaky tone</i> 𑌕		pa-Guru 𑌕 <i>consonant gemination</i> 𑌕	
Row 2b	mū 𑌕		kē 𑌕		uəɖ:i 𑌕𑌕	
Row 3a	ne-Deva 𑌕	ta-Taml 𑌕	si-Sinh 𑌕	kn-Knda 𑌕	te-Telu 𑌕	my-Mymr 𑌕
Row 3b	kAk 𑌕𑌕	kek 𑌕𑌕	kak 𑌕𑌕	kek 𑌕𑌕	kek 𑌕𑌕	kɛʔ 𑌕𑌕

Row 1a in Table 4 shows that kn-Knda, Telugu–Telugu (te-Telu) and Sanskrit–Grantha (sa-Gran) comprise certain bound graphs that effectively act as [C]-grams in themselves. This phenomenon is further detailed in Row 1b, which shows potential aksharas in these WSs with phonological values [CCV], [VC], and [CCCVC], wherein one of the [C]s is denoted by the bound [C]-gram in Row 1a. Rows 2a and 2b illustrate how bound aksharic elements in Odia–Odia (or-Orya), Burmese–Burmese (my-Mymr) and Punjabi–Gurmukhi (pa-Guru) may stand for ϕ -suprasegmentals such as nasalization, tone or even consonant gemination.⁴ Of these, pa-Guru | $\dot{\text{c}}$ | is particularly intriguing as it represents gemination of the phonological [C] represented by the following [C]-gram. With reference to the pa-Guru example in Row 2b, if | $\dot{\text{c}}$ | [vəḍ:i] is considered to comprise two aksharas, it would imply that | $\dot{\text{c}}$ | is graphematically situated within the first akshara but manifests phonologically under the second. Rows 3a and 3b depict a set of graphematically analogous bound graphs from various aksharic WSs, collectively known by the generic name *virama*.⁵ It is the viramas that pose perhaps the greatest challenge to a phonology-based definition of the akshara. As exemplified in Row 3b, the bound viramas are affixed to a free [CV₀]-gram to overtly indicate that [V₀] should not be pronounced.⁶ Due to this function, viramas are also known as vowel killers (Gnanadesikan, 2021) or zero vowel markers (Bright, 1996, p. 387). Since the viramas effectively denote [∅], it is unclear what category they fall under in terms of phonological value.

The examples in Table 4 add to the body of evidence for aksharas that deviate from the [V]-final prototype described in Section 2. Rows 1a and 1b tell us that several WSs feature aksharic candidates that are phonologically [C]-final, denote a phonological [C] by a graphematically bound element, or both. Rows 2a and 2b implicitly reveal that the sound value of the prototypical akshara is typically conceived of in terms of [C]s and [V]s—namely, in terms of ϕ -segments. This is despite the vast majority of aksharic WSs having provisions for denoting ϕ -suprasegmentals such as nasalization and tone. The example of pa-Guru | $\dot{\text{c}}$ | also drives home the point made by some authors (Gnanadesikan, 2017) that the graphematic boundaries of an akshara may not align with the boundaries of a phonological syllable. Finally, the existence of the bound graph known

4. The adjective [oɾia] is increasingly being rendered in en-Latn as [Odia] in place of the previously common [Oriya]. The adjective ‘Burmese’ remains common in anglophone linguistic circles, although the en-Latn form [Myanmar] is also encountered.

5. From Sanskrit /vireḥma/ ‘stop, pause’; English usu. /vɪˈɪɑ:mə/.

6. In certain WSs, the virama may have additional graphematic functions, such as forming part of complex bound graphs—as in si-Sinh |කෙ| [ke:]—or indicating specific phonological values depending on the graphematic environment—as in my-Mymr |ကေ| [keʔ].

as virama potentially results in aksharas that are [∅]-final. It remains to be seen how this aspect can be suitably captured in a phonology-based definition of the akshara. In all, the cases described and analyzed in this subsection strongly prompt us to revisit the prototype of an akshara described in Section 2, and reflect on whether it is at all possible to robustly and rigorously define an akshara in phonological terms alone.

3.2. Heterophonous Homography and Homophonous Heterography

In order to arrive at a holistic view of the situation, the WS-specific evidence provided in Section 3.1 needs to be complemented by a cross-WS or macro perspective. In this regard, Table 5 demonstrates how the same written sequence may be varying pronounced in different languages—or, more precisely, in different WSs.

TABLE 5. Heterophonous homography in WSs based on the Devanagari and Tibetan scripts

sa-Deva	ॐ८६		bo-Tibt	ke:ṽ	
ne-Deva	ॠ३ॠ	अंश	dz-Tibt	ke:ṽ	ཀེ
hi-Deva	ॐ८		lbj-Tibt	skət	ཀེཏ
mr-Deva	ॐ८		bft-Tibt	skət	

As seen in Table 5, the graph sequence |अंश| is a valid lexical item in Devanagari-script Sanskrit, Nepali, Hindi and Marathi. Similarly, the Tibetan-script |ཀེ| is a well-formed string in the written forms of the Tibetan, Dzongkha, Ladakhi and Balti languages. This reveals that a homographic sequence of prospective aksharas may be heterophonously interpreted depending on the language—and, by extension, the WS—in question. As evident, the heterophonous interpretations implicate not just the target ϕ -segments but also the ϕ -syllables indicated. For instance, Devanagari-script |अंश| is interpreted as disyllabic [ॐ८६] and [ॠ३ॠ] in the context of Sanskrit and Nepali, respectively, but as monosyllabic [ॐ८] and [ॐ८] in Hindi and Marathi, respectively. This phenomenon ties in with the observation in Section 3.1 that the graphematic boundaries of a written akshara may not align with the phonological boundaries of a spoken syllable.

A cross-WS analysis reveals that the converse—namely, homophonous heterography—is also attested. Table 6 provides examples of such occurrences.

In Table 6, each heterographic pair has the same phonological value, despite the members of each pair generally being viewed as differing in akshara count. Together with the examples from Table 5, it emerges

TABLE 6. Homophonous heterography in aksharic WSs

hi-Deva	əɖɖa	अड्डा अड्डा
ta-Taml	ləkɻmi	லக்ஷ்மி லக்ஷ்மி
ml-Mlym	ṇaṇma	നന്നമ നന്ന

that there is a large degree of variation across aksharic WSs in the correspondence between their written and spoken forms. In such a scenario, it proves problematic to define one in terms of the other. This further strengthens the case against using a phonological yardstick to define a written unit such as the akshara.

3.3. Innate Property or Orthographic Convention?

Of the examples in Table 6, it emerges that the homographic sequences in Hindi–Devanagari and Tamil–Tamil are in free variation. In most contexts, users may opt for either form based on personal preference. However, selecting between the two Malayalam–Malayalam forms has orthographic implications. The ml-Mlym sequence |നന്ന| is characteristic of the pre-1970s “traditional” orthography, whereas |നന്നമ| follows the simplified or “modern” orthography (Mohanani, 1996, p. 424).

The Lao-Lao (lao-Lao) presents perhaps the strongest argument against the use of a phonological metric to define an akshara. Traditionally, lao-Lao was an alphasyllabary-cum-abugida, much like the neighboring WS of Thai–Thai (th-Thai) (Diller, 1996). However, following the 1975 revolution in Laos, a new official orthography for lao-Lao was promulgated, which required all phonological [V]s to be overtly represented in writing (Gnanadesikan, 2021). The new rules effectively resulted in the elimination of [CV₀]-grams from the lao-Lao inventory, since the former [V₀] would now be written just like every other [V]—with a bound [V]-gram in postconsonantal position. The discarding of the inherent [V₀] feature meant that lao-Lao in the reformed orthography no longer qualified as an abugida. Gnanadesikan (ibid., p. 314) observes that lao-Lao’s loss of the inherent [V₀] was in no way the result of natural evolution, but was entirely attributable to government-imposed orthographic reform.

The examples of ml-Mlym and lao-Lao testify to orthographic conventions having the ability to alter or impact the phonological value of a written sequence and, consequently, its akshara count. This finding refutes any assumption that the akshara is an innate property of a WS.

Rather, it shows that aksharic WSs are impacted by top-down orthographic conventions just as WSs of other typological categories are.

Therefore, if an akshara's boundaries may vary independently of its phonological values, and its very existence impacted by external orthographic norms, there seems to be scant justification for using a phonological point of departure for defining this unit of writing. Rather, a robust, generalizable and expandable definition of the akshara appears more likely to emerge from a holistic approach, centered primarily on graphematic criteria.

4. Constraining the Akshara

Some of the conclusions in Section 3 have already been alluded to in the scholarly literature, albeit in disparate contexts. For instance, Nag (2017) notes that certain aksharic WSs may indicate specific consonant ϕ -segments by bound graphs, which is corroborated by the examples in Table 4. Conversely, Padakannaya, Pandey, Saligram, and Ranga Rao (2016) state that an akshara may have a variety of phonological values, be it a ϕ -segment, a ϕ -syllable or part of a syllable. This observation aligns with the examples of Table 5, which highlight the inherently variable overlap of an akshara with a ϕ -syllable. Tying together the above claims, Gnanadesikan (2021, p. 327) writes that “[t]he syllable-akshara mapping breaks down with [graphetically] more complex aksharas”. Gnanadesikan's observation that an akshara's phonological value becomes harder to predict with increasing graphetic complexity drives home the core argument in this paper—namely, that phonological value proves increasingly less reliable as a defining criterion for the akshara as we move away from the simple prototype.

My argument, therefore, is essentially to pivot from a phonology-dependent or dependentialist definition of the akshara towards one that is more phonology-independent or autonomistic (Haralambous and Dürst, 2019; Meletis, 2020). To wit, the akshara will always be *associated* with phonology, but should not be *decided* by it. This distinction is a subtle but crucial one. Accordingly, I propose that, across writing systems, an akshara is best conceived of in graphe(ma)tic terms, comprising one mandatory free graph and zero or more optional bound graphs. Table 7 details this conception using the examples from Table 6. Individual aksharas are shown separated by a γ -segmental space, while bound graphs are denoted in red.

In the proposed conception of an akshara, the notion of a free graph includes so-called conjuncts, ligatures and stacked-graphs compressed into one γ -segmental space, exemplified in Table 7 by $|\text{ॠॡ}|$, $|\text{ॢ}|$ and $|\text{ॣ}|$, respectively. Such a stance is based on the graphematic properties of the complex graphs in question being akin to the properties of simple

TABLE 7. Homophonous heterography with varying aksharic counts

hi-Deva	əɖɖa	अ ड ड	3 aksharas
		अ ड्	2 aksharas
ta-Taml	ləkfmi	ல க் ஷ் மி	4 aksharas
		ல க்ஷ் மி	3 aksharas
ml-Mlym	ṇṇnmē	ന ന് മ	3 aksharas
		ന ന്മ	2 aksharas

free graphs. For instance, |க்|, |ന്| and |ड्| each occupies a distinct γ -segmental space, possesses an identifiable phonological value, and is able to take on bound graphs, among other things. Hence, it makes sense to consider such complex graphs graphematically equivalent to simple free graphs in the aksharic context.

In theory, the number of optional bound graphs appendable to a free graph within the bounds of a single akshara could be infinite. In practice, however, the number of bound graphs one can affix to or juxtapose with the free aksharic nucleus would depend on WS-specific factors. These include the number of bound graphs available in that WS's graph inventory, any WS-specific graphematic-orthographic restrictions on the co-occurrence of certain bound graphs, and so on.

As evident, the above conceptualization of an akshara is intertwined with, but not reliant on, the phonological values of its constituent elements. Since an akshara is now being defined on graphematic terms, it could have any sound value. Thus, the free and bound elements of an akshara may denote a ϕ -segment, a ϕ -suprasegmental or a null ϕ -segment. Aksharic elements of the last type include viramas (Table 4, Rows 3a and 3b) and the shared graphetic bases used to construct independent [V]-grams in certain aksharic WSs, such as bo-Tibt |ཨ| and my-Mymr |အ|. ⁷ Besides, if an akshara is conceived of as a written unit centered around one free graph, it follows that one can add or remove as many bound

7. Such phonologically empty elements used as graphetic bases to construct independent [V]-grams are known by a variety of names: *vowel carriers*, *vowel bearers*, *vowel holders* or *vowel support letters*, among several others (Gill, 1996; Gnanadesikan, 2009; 2017; Iyengar, 2018; Salomon, 2007). In the context of Korean–Hangeul (kr-Hang), the terms *zero consonant* (King, 1996) and *dummy consonant* (Pae, 2011) have also been used. However, the last two terms have connotations of the graphetic elements in question being associated with a phonological [C], when, in reality, they simply occur in a paradigmatic relationship with graphematic [C]-grams or [CV₀]-grams. Iyengar (2023) labels the phonologically empty elements in question *kenograms*, from Greek |κενός| ‘empty’.

elements as graphematically permissible without impacting the akshara count. However, if one adds or removes a free graph, the akshara count would increase or decrease accordingly. This idea is reflected in the examples in Table 7.

5. Open Questions and Further Refinement

The robustness of the definition of an akshara proposed in this paper will only be vindicated by testing it on as many purportedly aksharic WSs as possible—namely, on WSs that are alphasyllabaries as well as abugidas. Doing so will bring to light WS-specific phenomena and how well the proposed definition of an akshara accounts for them. In fact, there are already several questions at this stage worth exploring further.

At the granular level, if an akshara should not be defined in terms of its phonological value, how far can a putative akshara depart from being fundamentally phonographic in nature? Put differently, can a logogram—comprising graphetic elements that cannot be compartmentalized and individually associated with specific phonological values—qualify as an akshara? For instance, ta-Taml |ஸ்ரீ| [sri: ~ ɕri:] is commonly described as a ‘ligature’, but is effectively logographic in nature (Amalia Gnanadesikan, personal communication, 10 June 2022). Its sound value is invariant, and it cannot take on additional bound graphs. Consequently, one may argue that ta-Taml |ஸ்ரீ| does not entirely satisfy the aksharic criteria proposed in this paper. Along similar lines, there exist homophonous-heterographic pairs such as sa-Deva |ओम्| and |ॐ|, both pronounced [o:m], and ta-Taml |ஓம்| and |ஐ|, both realized as [(^w)o:m]. Whereas the graphematic sequences |ओम्| and |ஓம்| are clearly phonographic and, therefore, aksharic in nature, |ॐ| and |ஐ| are evidently logograms whose aksharic status is contestable. Hence, more research is required into the question of an akshara’s compositional transparency (Meletis, 2020) and on setting out criteria thereunder for a graph to qualify as aksharic in nature.

At a macro or WS level, there remains room for the phonological associations of an akshara to be scrutinized further. In Section 2, the prototypical akshara was portrayed as a written unit displaying alphasyllabic as well as abugidic properties. Should this be a hard criterion for a graph—and, by extension, its source WS—to qualify as aksharic? Can a WS that is either an alphasyllabary or an abugida, but not both, be legitimately considered aksharic in nature? For instance, Divehi–Thaana (dv-Thaa) and Sindhi–Arabic (sd-Arab) both write postconsonantal [V]s using bound allographs, in the process qualifying as alphasyllabaries (Bright, 1999; Gnanadesikan, 2017; Iyengar, 2023). However, dv-Thaa and sd-Arab do not feature [CV₀]-grams and are, therefore, not abugidas. Conversely,

various South Asian languages when written in Braille—known collectively as Bharati Braille—do not feature bound [V]-grams, but do contain [CV₀]-grams (IIT Madras, 2020; Sproat, 2010). Hence, the Bharati Braille systems are not alphasyllabaries, but they are abugidas.⁸ Against this background, to what extent can dv-Thaa, sd-Arab and Bharati Braille be considered aksharic in nature, and their written units subject to the definition of akshara proposed in this paper? The question also applies to Korean–Hangeul (kr-Hang), which appears to satisfy the conditions for an alphasyllabary but not for an abugida (Iyengar, 2023; King, 1996). WSs based on Canadian Aboriginal Syllabics, such as those for Cree (cr-Cans), Inuktitut (iu-Cans) and Carrier (crx-Cans), present a particularly intriguing test case for the definition of akshara. These WSs denote a [C] by a specific graph shape and the subsequent [V] by the orientation of the graph, as in cr-Cans and iu-Cans |C| [ta] and |∩| [ti] (Harvey, 2003; Nichols, 1996). A sole [V] is indicated by a phonologically empty graph or kenogram (see footnote 7) appropriately oriented, as in cr-Cans and iu-Cans |<| [a] and |Δ| [i]. If one likens graph orientation to a bound [V]-gram in postconsonantal position, there are grounds to argue that cr-Cans, iu-Cans and crx-Cans are alphasyllabic and, hence, typologically similar to dv-Thaa and sd-Arab. Consequently, these WSs need to be considered collectively when making decisions on their aksharic status.

Aside from its graphematic implications, the question of whether non-abugidic or non-alphasyllabic WSs can be considered aksharic is also relevant from a sociolinguistic perspective. In several South Asian ‘letter’ is etymologically derived from the Sanskrit term /əkṣərɐ/ (Amalia Gnanadesikan, personal communication, 21 October 2021). Consequently, categorizing the South Asian WSs of dv-Thaa, sd-Arab and Bharati Braille as non-aksharic may prove dissonant with popular emic views on the basic graphematic unit of these WSs. At the same time, classifying sd-Arab as aksharic might sit uneasily with the pervasive portrayal in the literature of Arabic-script-based WSs as abjads, which persists despite scholarly evidence to the contrary (Bright, 1999; Gnanadesikan, 2017; Iyengar, 2023). Moreover, since the label *aksharic* remains semiotically associated with Brahmic WSs, it remains to be seen whether the Ethiopic WSs of Ge‘ez, Amharic and Tigrinya in the Ge‘ez script (gez-Ethi, am-Ethi and ti-Ethi, respectively) will be readily characterized as aksharic despite evidently satisfying the criteria proposed in this paper. For this reason, a definition of the akshara as a graphematic unit should, wherever possible, take into account and align with sociolinguistic perceptions of this unit of writing. To this end, I have attempted in this paper to address the graphematic aspect. A complementary sociolinguistic treatment of the subject matter is not only highly desirable but

8. For additional examples of WSs that are either alphasyllabaries or abugidas, but not both, see Gnanadesikan and Judson (2021) and Iyengar (2023).

imperative if we are to arrive at a comprehensive and epistemologically robust definition of the elusive unit of writing that is the akshara.

References

- Altmann, Gabriel and Fan Fengxiang, eds. (2008). *Analyses of script: Properties of characters and writing systems*. Berlin/New York: Mouton de Gruyter.
- Bright, William (1996). "The Devanagari script." In: *The world's writing systems*. Ed. by Peter T. Daniels and William Bright. New York: Oxford University Press, pp. 384–390.
- (1999). "A matter of typology: Alphasyllabaries and abugidas." In: *Written Language & Literacy* 2.1, pp. 45–55.
- Bunčić, Daniel, Sandra L. Lippert, and Achim Rabus, eds. (2016). *Biscriptality: A sociolinguistic typology*. Heidelberg: Universitätsverlag Winter.
- Constable, Peter (2004). *Review of Bengali Khanda Ta and PRI-30 feedback*. Unicode.
- Daniels, Peter T. (1996). "The study of writing systems." In: *The world's writing systems*. Ed. by Peter T. Daniels and William Bright. New York: Oxford University Press, pp. 3–17.
- (2018). *An exploration of writing*. Sheffield: Equinox.
- Daniels, Peter T. and William Bright (1996). "Abbreviations, conventions, and definitions." In: *The world's writing systems*. Ed. by Peter T. Daniels and William Bright. New York: Oxford University Press, pp. xxxix–xlv.
- Diller, Anthony (1996). "Thai and Lao writing." In: *The world's writing systems*. Ed. by Peter T. Daniels and William Bright. New York: Oxford University Press, pp. 457–466.
- Everson, Michael (2007). *Proposal for encoding the Meitei Mayek script in the BMP of the UCS*. Unicode. <https://unicode.org/wg2/docs/n3206.pdf>.
- Gill, Harjeet Singh (1996). "The Gurmukhi script." In: *The world's writing systems*. Ed. by Peter T. Daniels and William Bright. New York: Oxford University Press, pp. 395–398.
- Gnanadesikan, Amalia E. (2009). *The writing revolution: Cuneiform to the internet*. Malden, MA: Wiley-Blackwell.
- (2017). "Towards a typology of phonemic scripts." In: *Writing Systems Research* 9.1, pp. 14–35.
- (2021). "Brahmi's children: Variation and stability in a script family." In: *Written Language & Literacy* 24.2, pp. 303–335.
- Gnanadesikan, Amalia E. and Anna P. Judson (2021). "Diversity in writing systems: Embracing multiple perspectives." In: *Written Language & Literacy* 24.2, pp. 167–170.

- Haralambous, Yannis (2019). "Phonocentrism in Greece: Side effects of two centuries of diglossia." In: 12th international workshop on writing systems and literacy (March 25–28, 2019). University of Cambridge: Association for Written Language and Literacy.
- Haralambous, Yannis and Martin Dürst (2019). "Unicode from a linguistic point of view." In: *Graphemics in the 21st Century. Brest, June 13–15, 2018. Proceedings*. Grapholinguistics and Its Applications, Vol. 1. Brest: Fluxus Editions, pp. 127–166.
- Harvey, Christopher (2003). *Some general aspects of the Syllabics orthography - The nature of syllabics*. Languagegeek. <https://www.languagegeek.com/typography/syllabics/syl2.html>.
- Honda, Keisuke (2021). "A modular theoretic approach to the Japanese writing system: Possibilities and challenges." In: *Grapholinguistics in the 21st Century 2020. Proceedings*. Ed. by Yannis Haralambous. Grapholinguistics and Its Applications, Vol. 5. Brest: Fluxus Editions, pp. 621–643.
- IIT Madras (2020). *Introduction to Bharati Braille. Acharya: Multilingual computing for literacy and education*.
- Iyengar, Arvind V. (2018). "Variation in Perso-Arabic and Devanāgarī Sindhī orthographies: An overview." In: *Written Language & Literacy* 21.2, pp. 169–197.
- (2023). "More matters of typology: Alphasyllabaries, abugidas and related vowelised segmentaries." In: *Written Language & Literacy* 26.1, pp. 30–56.
- Johnson, Sally (2005). *Spelling trouble? Language, ideology and the reform of German orthography*. Clevedon: Multilingual Matters.
- Joshi, R. Malatesha and Catherine McBride, eds. (2019). *Handbook of literacy in akshara orthography*. Cham, Switzerland: Springer.
- Joyce, Terry and Dimitrios Meletis (2021). "Alternative criteria for writing system typology." In: *Zeitschrift für Sprachwissenschaft* 40.3, pp. 257–277.
- King, Ross (1996). "Korean writing." In: *The world's writing systems*. Ed. by Peter T. Daniels and William Bright. New York: Oxford University Press, pp. 218–227.
- Meletis, Dimitrios (2020). *The nature of writing: A theory of grapholinguistics*. Grapholinguistics and Its Applications, Vol. 3. Brest: Fluxus Editions.
- Mohanan, K P (1996). "Malayalam writing." In: *The world's writing systems*. Ed. by Peter T. Daniels and William Bright. New York: Oxford University Press, pp. 420–426.
- (2007). *Theoretical and practical aspects of Cbillus in Malayalam.unicode*.
- Nag, Sonali (2017). "Learning to read languages of South Asia." In: *Learning to read across languages and writing systems*. Ed. by Ludo Verhoeven and Charles Perfetti. Cambridge: Cambridge University Press, pp. 104–126.

- Nag, Sonali and Charles Perfetti (2014). "Reading and writing: Insights from the alphasyllabaries of South and Southeast Asia (special issue)." In: *Writing Systems Research* 6.1, pp. 1–9.
- Neef, Martin (2015). "Writing systems as modular objects: proposals for theory design in grapholinguistics." In: *Open Linguistics* 1.1, pp. 708–721.
- Nichols, John D. (1996). "The Cree syllabary." In: *The world's writing systems*. Ed. by Peter T. Daniels and William Bright. New York: Oxford University Press, pp. 599–611.
- Osterkamp, Sven and Gordian Schreiber (2021). "<Th>e ubi<qu>ity of polygra<ph>y and its significan<ce> for <th>e typology of <wr>iti<ng> systems." In: *Written Language & Literacy* 24.2, pp. 171–197.
- Padakannaya, Prakash et al. (2016). "Visual-orthographic complexity of Akshara and eye movements in reading: A study in Kannada alphasyllabary." In: *Writing Systems Research* 8.1, pp. 32–43. ISSN: 1758-6801.
- Pae, Hye K. (2011). "Is Korean a syllabic alphabet or an alphabetic syllabary?" In: *Writing Systems Research* 3, pp. 103–115.
- Raymond, Martin (2020). *Presenting writing system information*. ScriptSource. https://scriptsource.org/cms/scripts/page.php?item_id=entry_detail&uid=r7jbnbjkfw.
- Rimzhim, Anurag, Leonard Katz, and Carol A Fowler (2014). "Brāhmī-derived orthographies are typologically Āksharik but functionally predominantly alphabetic." In: *Writing Systems Research* 6.1, pp. 41–53.
- Salomon, Richard G. (2007). "Writing systems of the Indo-Aryan languages." In: *The Indo-Aryan languages*. Ed. by George L Cardona and Dhanesh Jain. Oxon: Routledge, pp. 75–114.
- Sproat, Richard W. (2000). *A computational theory of writing systems*. Cambridge: Cambridge University Press.
- (2010). *Language, technology, and society*. New York: Oxford University Press.
- (2018). "Peter T. Daniels, 2018, 'An Exploration of Writing'." In: *Written Language & Literacy* 21.2, pp. 269–278.
- Vaid, Jyostna (2022). "Biscriptality: a neglected construct in the study of bilingualism." In: *Journal of Cultural Cognitive Science* 6, pp. 135–149.
- Weingarten, Rüdiger (2013). "Comparative graphematics." In: *Typology of writing systems*. Ed. by Susanne R. Borgwaldt and Terry Joyce. Amsterdam: John Benjamins, pp. 13–39.
- Winkel, Heather and Prakash Padakannaya, eds. (2014). *South and South-east Asian psycholinguistics*. New York: Cambridge University Press.