ISSN 1799-2591
Theory and Practice in Language Studies, Vol. 15, No. 3, pp. 979-992, March 2025
DOI: https://doi.org/10.17507/tpls.1503.33

Determining the Exact Conditions That Affect Grapheme-Phoneme Correspondence Rules With Generative Phonology: A Case Study of Transliteration From the Balinese Script to Latin

I K. Paramarta*
Department of Balinese Language Education, Universitas Pendidikan Ganesha, Bali, Indonesia

I Wayan Rasna
Department of Indonesian Language Education, Universitas Pendidikan Ganesha, Bali, Indonesia

I B. Putrayasa
Department of Indonesian Language Education, Universitas Pendidikan Ganesha, Bali, Indonesia

I Nengah Suandi
Department of Indonesian Language Education, Universitas Pendidikan Ganesha, Bali, Indonesia

Kadek Wirahyuni
Department of Indonesian Language Education, Universitas Pendidikan Ganesha, Bali, Indonesia

***Abstract*—This study aims to investigate the precise conditions required to establish rules governing grapheme-phoneme correspondence (GPC) for six vowel graphemes in the Balinese script: <ᬳ> /a/, <ᬳᬶ> /i/, <ᬳᬸ> /u/, <ᬳᬾ> /e/, <ᬳᭀ> /o/, <ᬳᭂ> /ə/. The data were derived from 2,274 lemmas listed in the Balinese dictionary, which utilizes both Latin and Balinese scripts. The formulation of GPC rules is based on an analysis of underlying representations and morphemic structure conditions. Initial descriptions of the rules in ordinary language were subsequently converted into formal notations using the framework of four types of generative phonological rules. The GPC rules for the six vowel graphemes are derived from the independent underlying GPC <ᬳ> /ha/ and dependent markers for vowels: <◌ᬶ>/i/, <◌ᬸ>/u/, <◌ᬾ>/e/, <◌ᭀ>/o/, <◌ᭂ>/ə/. The analysis of morphemic structure conditions within canonical patterns reveals that syllable-initial positions serve as the environment that triggers changes in the underlying forms. Two key rules govern the GPC for these six vowel graphemes: (1) the deletion of voiceless glottal consonants and (2) sequential rules, which include the heightening of low vowels and the deletion of the voiceless glottal /h/. The exact conditions represent definitive statements that underpin the development of GPC alternation rules. These rules, which systematically map sequences of orthographic units onto their corresponding linguistic forms, provide a foundation for analyzing the transparency of a writing system.**

***Index Terms*—Balinese script, grapheme, phoneme, correspondence rule**

I. Introduction

The degree of orthography-phonology transparency is a critical factor in cross-writing system research, frequently associated with the systematic mapping of graphemes onto phonemes (Protopapas & Vlahou, 2009). This mapping system addresses both parallel and non-parallel correspondences between sequences of written units and linguistic units (Nag, 2014). Non-parallelism, observed as inconsistencies in grapheme-phoneme mapping, arises from the principle of interdependence between spoken and written language. Despite their relative (Pae, 2020; Pae & Wang, 2022) and autonomy, these systems exhibit consistent patterns of correspondence (Meletis & Dürscheid, 2022).

The inconsistency problem is one of the three primary challenges identified in the psycholinguistic grain size theory of reading acquisition, alongside the availability problem and the granularity problem. Ziegler and Goswami (2005), describe the inconsistency problem as a condition in which certain orthographic units have multiple pronunciations, and some phonological units have multiple spellings. These inconsistencies arise when a single orthographic unit

* Corresponding Author

© 2025 ACADEMY PUBLICATION

corresponds to multiple phonological units (Bosch & Daelemans, 1993) or when specific orthographic units lack phonological representation or are unpronounceable (Bassetti & Atkinson, 2015; Paramarta et al., 2024; Pike, 1956).

Inconsistencies resulting in ambiguous grapheme-phoneme mappings have significant implications across various domains related to the interaction between orthography and phonology. These include reading outcomes across writing systems (Burani et al., 2006; Jared, 2002; Moitra et al., 2024; Nag et al., 2010; Treiman et al., 1995), the accuracy of modern text-to-speech (TTS) systems (Gonzalvo & Podsiadło, 2014), the precision of akshara-to-sound conversions (Pandey, 2014), and the effectiveness of automated script converters for diverse writing systems (Rajan, 2018). Furthermore, such inconsistencies affect the acquisition and development of writing systems (Frost et al., 1987). The prevalence of orthography-phonology mismatches also represents a fundamental challenge in the study of specific writing systems (Sircar & Nag, 2014; Ziegler et al., 2010).

Ambiguous grapheme-phoneme mappings are also evident in the Balinese writing system, particularly concerning six orthographic units (graphemes) <ᬳ>, <ᬳᬶ>, <ᬳᬸ>, <ᬳᬾ>, <ᬳᭀ>, <ᬳᭂ>, which correspond to multiple phonemic units (Medera et al., 2003; Simpen AB, 1979; Tinggen, 1994). First, these graphemes represent the voiceless glottal consonant /h/ accompanied by the inherent vowel /a/, as seen in /ha/, or other vowels when combined with bound graphemes as vowel markers, resulting in forms such as /hi/, /hu/, /he/, /ho/, and /hə/. Second, under certain conditions, these graphemes may also map directly to full vowels /i/, /u/, /e/, /o/, and /ə/.

Inconsistencies in grapheme-phoneme mappings within the Balinese script writing system have adversely affected the accuracy of grapheme-phoneme conversion results, which are essential for automating the transliteration of Balinese script into Latin. The findings from conversion trials conducted with the ToLatin (Computer Science Undiksha, 2023) applications reveal that there are significant mapping errors concerning six orthographic units. These units should properly correspond to the phonemes represented as full vowels /a/, /i/, /u/, /e/, /o/, and /ə/. However, they are currently being incorrectly mapped as */ha/, */hi/, */hu/, */he/, */ho/, and */hə/. Figure 1 illustrates the errors in the transliteration results produced by the ToLatin applications for the words ᬳᬶᬗᬸᬄ */iŋuh/* 'confused, agitated', ᬩᬳᭀᬂ */baoŋ/* 'neck', and ᬘᬳᬸ */cau/* 'offering for rice plants'. The graphemes <ᬳᬶ>, <ᬳᭀ>, and <ᬳᬸ> in these three words should correspond to the vowel phonemes /i/, /o/, and /u/. Therefore, the correct transliterations into Latin should be *inguh*, *baong*, and *cau*, rather than **hinguh*, **bahong*, and **cahu*.

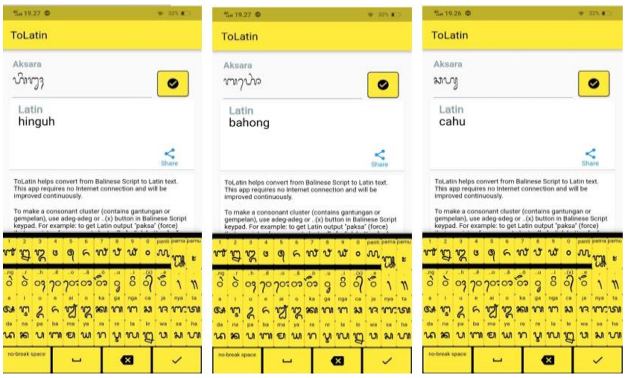

Figure 1. Illustrates the Errors in Grapheme-Phoneme Mapping Observed in the ToLatin Applications

This study investigates the specific conditions necessary to establish rules governing grapheme-phoneme alternations for six graphemes derived from a list of base lemmas in a Balinese-Latin script dictionary. Drawing on the framework of generative phonology, three key conditions were identified as the basis for determining grapheme-phoneme correspondence rules: (1) which segments undergo grapheme-phoneme changes, (2) the nature of the changes, and (3) the conditions under which these changes occur (Schane, 1973).

The grapheme-phoneme correspondence (GPC) framework offers a robust theoretical approach for representing orthographic-phonological knowledge. It provides a predictive and systematic method for identifying and defining consistent patterns in grapheme-phoneme correspondences (Coltheart et al., 1993, 2001).

II. Theoretical Framework

A. *Grapheme to Phoneme Correspondence (GPC) Rule*

Letter-to-sound rules, grapheme-to-phoneme correspondence (GPC) rules, or grapheme-to-phoneme conversion are essentially analogous to phonological rules in classical phonological studies, except that they specifically involve the conversion of grapheme sequences into phoneme sequences (Divay & Vitale, 1997). These rules are determined by "context-dependent" or "contex-sensitive" and "contex-free" (Burani et al., 2006).

© 2025 ACADEMY PUBLICATION

The transformation of spelling into a phonological form is referred to as the process of "recoding" (Neef, 2012). Neef introduces the term "recoding model of graphematic" as a theoretical reconstruction of the relationship between a writing system and the phonological system of a language. The input to this model is a graphematic representation, such as a sequence of written units within the graphematic word domain. Letters (*akshara*) are the fundamental abstract graphematic units, defined by correspondence rules. Each letter unit is governed by a specific type of rule. These correspondence rules can be highly complex, accommodating multiple possible options.

B. Generative Phonology

Generative phonology is a subfield of linguistic theory known as transformational generative grammar. According to this theory, the phonological component processes surface structures to yield a phonetic representation. In generative phonology, all linguistic forms are assumed to have an underlying form and a corresponding phonetic representation. The transformation of an underlying form into a derived form is governed by phonological rules, which can be seen as specifying the exact conditions under which a phonological process occurs (Schane, 1973).

The derivational approach, in relation to phonology and writing systems, has recently been considered outdated and has been largely supplanted by Optimality Theory (Prince & Smolensky, 1993). However, in practice, many analyses of orthography-to-phonology correspondence, particularly in computational studies of writing systems, continue to rely on the foundational principle of "finite-state power," using recoding rules within the derivational framework (Sproat, 2000).

Pastika (2005), based on the generative phonology theory of Chomsky and Halle (1968) and Schane (1973), identified 24 underlying phonemes in the Balinese language, consisting of 6 vowel phonemes and 18 consonant phonemes. He also identified 28 canonical morpheme patterns and 17 phonological rules. Pastika findings provide a basis for this study to: 1) identify the underlying phonemes of Balinese, 2) determine the exact conditions or specific contexts (positions and phonetic environments) that lead to changes in GPC, 3) develop and formalize GPC rules, and 4) establish guidelines for converting phonemic transcriptions into orthography.

C. Balinese Script Orthography

The Balinese script is divided into four based on its functions: 1) *Aksara Wreastra*, 2) *Aksara Swalalita*, 3) *Aksara Wijaksara*, and 4) *Aksara Modre* (Suasta, 2006). *Aksara Wreastra* also known as the ordinary script, is commonly used to write the modern Balinese language for everyday purposes. In contrast, the *Aksara Swalalita* serves not only for writing the modern Balinese language but also for Old Javanese and Sanskrit. Meanwhile, the *Aksara Wijaksara* and *Modre* are regarded as sacred and are typically used in the context of religious and mystical texts.

In this study, we focus on investigating the GPC rules in the *Aksara Wreastra*. The *Wreastra* script has 18 main orthographic units for consonants (*aksara wianjana*) and 6 orthographic units for vowels (*aksara suara*) (Medera et al., 2003; Tinggen, 1994). The *aksara wianjana* for consonants already carries the inherent vowel /a/, namely |ᬳ| /ha/, |ᬦ| /na/, |ᬘ| /ca/, |ᬭ| /ra/, |ᬓ| /ka/, |ᬤ| /da/, |ᬢ| /ta/, |ᬲ| /sa/, |ᬯ| /wa/, |ᬮ| /la/, |ᬫ| /ma/, |ᬕ| /ga/, |ᬩ| /ba/, |ᬗ| /ŋa/, |ᬧ| /pa/, |ᬚ| /ja/ |ᬬ|/ya/ |ᬜ| /ɲa/. The consonant letters also have variants in the form of conjunct-ligature (terms from Fedorova (2013)), collectively referred to as *pangangge aksara*, which include *pangangge ardasuara*, *pangangge tengenan*, *gantungan*, and *gempelan*. The *aksara suara* for vowels include: |ᬳ| /a/, |ᬳᬶ| /i/, |ᬳᬸ| /u/, |ᬳᬾ| /e/, |ᬳᭀ| /o/, and |ᬳᭂ| /ə/. The six orthographic units for the vowels were taken from one of the *aksara wianjana* |ᬳ| /ha/ (called *wisarga*) which is combined with diacritics for vowel markers or *pangangge suara* (|◌ᬶ|/i/, |◌ᬸ|/u/, |◌ᬾ|/e/, |◌ᭀ|/o/, |◌ᭂ|/ə/) as needed.

D. Grapheme

A grapheme is defined as the smallest unit that represents a form or a fundamental entity within a writing system. Meletis (2019) proposes three universal criteria for defining graphemes applicable across various writing systems: (1) lexical distinctiveness, (2) linguistic value, and (3) minimality. In the context of the Balinese script, the criterion of linguistic value pertains to the role of graphemes as visual symbols that connect the graphic units of writing with representations of language units, such as phonemes, sub-syllables, syllables, and morphemes (Astiti et al., 2023a, 2023b; Paramarta et al., 2024) or representations of abstractions of linguistic value, such as a null phoneme (Paramarta et al., 2024).

III. Research Method

This study employs a qualitative descriptive research design. The qualitative data includes six GPC for vowels in the Balinese script <ᬳ> /a/, <ᬳᬶ> /i/, <ᬳᬸ> /u/, <ᬳᬾ> /e/, <ᬳᭀ> /o/, and <ᬳᭂ> /ə/ that conducted from 2,274 lemmas listed in the Balinese dictionary uses Latin-Balinese script (Nala Antara et al., 2016). The sequence of grapheme units

© 2025 ACADEMY PUBLICATION

of the Balinese script in the compiled lemma list is then mapped to a sequence of phonemic units, converted from Latin orthographic transcription following the guidelines established by Pastika (2005).

Based on the mapping results of the GPC sequences from the complete lemma list, the underlying forms of the GPC were subsequently analyzed. The analysis refers to the identification of the underlying grapheme units within the Balinese script writing system, as established by Astiti (2023), and was conducted using the three criteria proposed by Meletis (2019). Meanwhile, the linguistic units represented by graphemes as phonemic segments refer to the underlying phonemic segments in the Balinese language as established by Pastika (2005). Each phonemic segment is assigned phonetic features using a binary system of plus (+) and minus (-) to indicate the presence or absence of specific phonetic parameters (Schane, 1973).

The specific conditions (phonetic environments and positions) that result in changes in GPC are analyzed through morpheme structure condition analysis. This analysis, conducted on the entire lemma list, aims to verify the realization of canonical patterns by adhering to the 28 canonical morpheme patterns of Balinese root morphemes as established by Pastika.

The exact conditions under which a GPC process occurs are first described in ordinary language. These descriptions are subsequently converted into formal notations to construct GPC rules, based on the notations defined within the four types of generative phonological rules: feature-changing rules, deletion and insertion rules, permutation and coalescence rules, and variable-based rules (Schane, 1973). The formal notation utilized includes the following: square brackets ([ ]) are used to denote a set of minimal distinctive features that characterize a segment or phonemic representation undergoing change; an arrow (→) indicates a transformation, specifying the segment that changes and the nature of the change; diagonal slashes (/) separate the environment from other rules; a hyphen (_) represents the specific part of the environment where the change occurs; **C** denotes the class of consonants, and **V** denotes vowels; the symbol (+) marks morpheme boundaries; the symbol (#) marks word boundaries; deletion is represented by the zero symbol (Ø); and brace notation ({ }) is used in rules to refer to alternative environments.

IV. Results and Discussions

The findings of this study address three main issues related to the determination of the exact conditions that affect the GPC rules for the six vowels in the Balinese script. First, an analysis of the underlying representation of the GPC. Second, an analysis of morpheme structure conditions in the realization of canonical patterns to identify the specific conditions (phonetic environments and positions) that lead to changes in the GPC. Third, the conversion of these exact conditions into formal notations for the construction of GPC rules. The complete results and discussion of the study are presented in the following sections.

A. The Underlying Form GPC

Based on the three main criteria for determining graphemes across writing systems by Meletis (2019), the underlying graphemes in the Balinese script writing system, particularly within the *wreastra* script type, are classified into two categories: free graphemes and bound graphemes. Free graphemes comprise 18 primary consonant graphemes, whereas bound graphemes encompass all diacritics used to mark bound vowels and conjunct-ligatures that signify bound consonants. These bound consonants function as paired variants of the primary consonant graphemes, demonstrating graphematic allography (Astiti, 2023; Meletis, 2020). Pastika (2005) identified underlying phonetic segments as representations of Balinese script graphemes in the realization of underlying phonemes in the Balinese language through an analysis of contrasts in identical environments.

(a). The Underlying Form GPC <ᬳ> /ha/

The grapheme-phoneme <ᬳ> /ha/ is established as a realization of the underlying GPC based on an analysis of grapheme criteria for lexical distinctiveness, demonstrated through minimal pair data that exhibit contrasts in identical environments, as illustrated in Data *(a). (1,2,3)*. The underlying GPC <ᬳ> /ha/ is only distributed in the final position.

Data (a).

1) <ᬲᬳ> /*saha*/ 'together' or 'with'

<ᬲᬧ> /*sapa*/ ‘greeting’

2) <ᬤᬳ> /*daha*/ ‘royal name’

<ᬤᬢ> /*data*/ ‘flat’

© 2025 ACADEMY PUBLICATION

3) <ᬫᬳ>-/*maha*/ ‘great’

<ᬫᬓ>-/*maka*/ ‘each’

The data above indicate that the grapheme <ᬳ> represents an underlying form categorized as a free grapheme. This underlying grapheme belongs to the primary consonant grapheme category, representing the linguistic unit of the voiceless glottal consonant /h/, which phonemically carries the distinctive features [-sonorant, +low]. It also inherently includes the vowel /a/ (without a specific marker), which is phonetically realized with the distinctive features [+back, +low].

(b). The Underlying Form GPC <◌ᬶ> /i/, <◌ᬸ> /u/, <◌ᬾ> /e/,<◌ᭀ> /o/, <◌ᭂ> /ə/

The grapheme-phonemes <◌ᬶ> /i/, <◌ᬸ> /u/, <◌ᬾ> /e/, <◌ᭀ> /o/, and <◌ᭂ> /ə/ are identified as realizations of the underlying GPC based on an analysis of grapheme criteria for lexical distinctiveness. This is demonstrated through minimal pair data that exhibit contrasts in identical environments. The underlying GPC is distributed in initial, medial, and final positions, except for <ᬳᭂ> /ə/, which does not appear in the final position of the morpheme (see Data (b), (c), (d), (e), and (f)).

Data *(b).*

1) <ᬮᬶᬫ>-/*lima*/ ‘five’

<ᬮᬫ>-/*lama*/ ‘failing a grade’

2) <ᬩᬗ᭄ᬓᬶᬢ᭄>-/*baŋkit*/ ‘delicious, interesting’

<ᬩᬗ᭄ᬓᭂᬢ᭄>-/*baŋkət*/ ‘ricefield’

3) <ᬩᬢᬶ>-/*bati*/ ‘profit’

<ᬩᬢᬸ>-/*batu*/ ‘stone’

Data (c).

1) <ᬲᬸᬓ>-/*suka*/ ‘like’

<ᬲᭀᬓ>-/*soka*/ ‘angsoka flower’

2) <ᬢᬓᬸᬮ᭄>-/*takul*/ ‘awkward’

<ᬢᬓᬶᬮ᭄>-/*takil*/ ‘wrap rice from areca leaves (*upih*)’

3) <ᬲᬭᬸ>-/*saru*/ ‘vague’

<ᬲᬭᬶ>-/*sari*/ ‘essence’

Data (d).

1) <ᬲᬾᬮ>-/*sela*/ ‘tubers’

<ᬲᬶᬮ>-/*sila*/ ‘base’

2) <ᬩᬭᬾᬲ᭄>-/*bares*/ ‘charitable’

<ᬩᬭᬶᬲ᭄>-/*baris*/ ‘*Baris* dance’

3) <ᬩᬮᬾ>-/*bale*/ ‘hall’

<ᬩᬮᬶ>-/bali/ ‘Balinese Island’

Data (e).

© 2025 ACADEMY PUBLICATION

1) <ᬩᭀᬤᬕ᭄>-/*bodag*/ ‘big basket’

<ᬩᬸᬤᬕ᭄>-/*budag*/ ‘strong enough to eat’

2) <ᬕᬾᬚᭀᬃ>-/*gejor*/ ‘earthquake’

<ᬕᬾᬚᬶᬃ>-/*gejir*/ ‘shake’

3) <ᬩᬦᭀ>-/*bano*/‘fish name for *Tylosurus crocodilus*’

<ᬩᬦᬶ>-/*bani*/ ‘brave’

Data (f).

1) <ᬲᭂᬮ>-/*səla*/ ‘leisure’

<ᬲᬶᬮ>-/*sila*/ ‘base’

2) <ᬜᭂᬄ>-/*ɲəh*/ ‘scared, amused’

<ᬜᬸᬄ>-/*ɲuh*/ ‘coconut’

3) <ᬳᬚᭂᬂ>-/*ajəŋ*/ ‘eat (high variety)’

<ᬳᬚᬸᬂ>-/*ajuŋ*/ ‘call for dad’

Data (b) indicate that the underlying GPC <◌ᬶ>-/i/ is bound to the primary consonant grapheme below it and is positioned above the primary consonant grapheme. Data (c) indicate that the underlying grapheme <◌ᬸ>-/u/ is bound to the primary consonant grapheme above it and is positioned below the primary consonant grapheme. Data (d) indicate that the underlying grapheme <◌ᬾ>-/e/ is bound to the primary consonant grapheme that follows it and is positioned to the left or before the primary consonant grapheme. Data (e) indicate that the underlying grapheme <◌ᭀ>-/o/ is bound to the primary consonant grapheme that both precedes and follows it and is positioned on both sides of the primary consonant grapheme. Data (f) indicate that the underlying grapheme <◌ᭂ>-/ə/ is bound to the primary consonant grapheme below it and is positioned above the primary consonant grapheme.

The data above indicate that the grapheme <◌ᬶ> corresponds to the vowel /i/ with the features [+sonorant, +high, -round], the grapheme <◌ᬸ> corresponds to the vowel /u/ [+syllabic, +high, +round], the grapheme <◌ᬾ> corresponds to the vowel /e/ [+syllabic, -high, -back], the grapheme <◌ᭀ> corresponds to the vowel /o/ [+syllabic, -high, +round], and the grapheme <◌ᭂ> corresponds to the vowel /ə/ [+syllabic, +back, -round, -low]. These five graphemes function as bound graphemes used to mark vowels in the Balinese script writing system.

Based on the criteria of lexical distinctiveness, linguistic value, and minimality, all underlying GPCs are demonstrated to distinguish lexical meaning through minimal pair data derived from a list of lemmas. Analysis using the criterion of linguistic value reveals that underlying graphemes can represent different linguistic units, such as combinations of consonant and vowel phonemes (/ha/) or individual vowel phonemes (/i/, /u/, /e/, /o/, and /ə/). Furthermore, from the perspective of linguistic levels, these underlying graphemes correspond to various levels of representation, including phonemes, syllables, and sub-syllabic units.

These findings reinforce the concept of linguistic compatibility proposed by Joyce and Meletis (2021), which suggests a typological classification of writing systems that better represents mixed linguistic units rather than relying solely on dominant linguistic representations. Additionally, testing through the criterion of minimality indicates that visually, all underlying graphemes are the smallest units within the Balinese script, classified under the *aksara wianjana* and *pangangge suara* categories (Medera et al., 2003), or categorized as "letters" by Lindqvist (2001) to distinguish them from punctuation, digits, and signs.

Rogers (1995) classifies graphemes based on their forms into two main groups: free graphemes and bound graphemes. The underlying GPC <ᬳ>-/ha/ belongs to the category of free graphemes, or in akshara-based systems, it is

© 2025 ACADEMY PUBLICATION

referred to as a main consonant grapheme (Fedorova, 2013; Gnanadesikan, 2017). In contrast, graphemes in the form of diacritics used as vowel markers fall under the category of bound graphemes, which Fedorova describes as "clothed".

The absence of an independent horizontal segmental space for the vowel graphemes <◌ᬶ>/i/, <◌ᬸ> /u/, <◌ᬾ> /e/, <◌ᭀ> /o/, and <◌ᭂ> /ə/ confirms their bound nature, meaning they cannot stand alone, particularly in post-consonantal positions. The vowel /a/ does not have a specific marker in the form of a bound grapheme. This aligns with the fundamental characteristics of the Balinese script writing system, which is categorized as an Abugida or Alphasyllabary type. In such systems, the main consonant grapheme (*akshara*) inherently includes a vowel, specifically the low vowel /a/ (Bright, 1996; Pandey, 2014; Vaid & Gupta, 2002).

The underlying GPC, when combined with other grapheme-phoneme elements to form Balinese words, undergoes various transformational processes, resulting in diverse realizations of derived GPCs. The environmental conditions surrounding the GPC are crucial factors influencing these transformations. This phenomenon aligns with the foundational principle of generative phonology, which asserts that changes from an underlying form to its derived forms are governed by precise conditions and can be systematically described through rules governing such transformations (Schane, 1973).

B. Morpheme Structure Condition

The Balinese morphemes base examined in this study comprise 2,274 listed lemmas. An analysis of the morpheme structure condition for the entire lemma list reveals that, out of the 28 canonical patterns of underlying Balinese morphemes reported by Pastika (2005), only eleven canonical patterns were identified to contain the GPC <ᬳ> /a/, <ᬳᬶ> /i/, <ᬳᬸ> /u/, <ᬳᬾ> /e/, <ᬳᭀ> /o/, and <ᬳᭂ>. The eleven canonical patterns and their frequencies of occurrence are as follows: (1) V.V, with 11 occurrences; (2) V.VC, with 104 occurrences; (3) V.CV, with 160 occurrences; (4) CV.V, with 40 occurrences; (5) V.CVC, with 1,045 occurrences; (6) CV.VC, with 243 occurrences; (7) VC.CVC, with 596 occurrences; (8) VC.CV, with 56 occurrences; (9) CV.V.CV, with 4 occurrences; (10) VC.CV.CVC, with 3 occurrences; and (11) CV.CV.V, with 12 occurrences.

The eleven canonical patterns identified in the lemma list were further analyzed based on the number of syllables and the position of the GPC. Table 1 presents the results of the analysis, showing the number of syllables and the positional distribution of the GPC within the eleven canonical patterns.

© 2025 ACADEMY PUBLICATION

TABLE 1
CANONICAL PATTERNS, NUMBER OF SYLLABLES, AND APPEARANCE POSITION

Canonical patterns	Number of syllables		Appearance position			Example (lemma)
	Bisyllabic	Trisyllabic	Initial	Middle	Final	
V.V	√		√		√	<ᬳᬳ>-/*a.a*/ ‘fig tree’ <ᬳᬳᬶ>-/*a.i*/ ‘sun’
V.VC	√		√	√		<ᬳᬳᬸᬤ᭄>-/*a.ud*/ ‘put it to release’ <ᬳᭀᬳᭀᬂ>-/*o.oŋ*/ ‘mushroom’
V.CV	√		√			<ᬳᬩᬸ>-/*a.bu*/ ‘ash’ <ᬳᬸᬤᬸ>-/*u.du*/ 'not selling well’
CV.V	√				√	<ᬘᬳᬶ>-/*ca.i*/ ‘you (male)’ <ᬘᬳᬸ>-/*ca.u*/ ‘offerings for rice ceremonies’
V.CVC	√		√			<ᬳᭀᬩᭀᬃ>-/*o.bor*/ ‘torch’ <ᬳᬸᬩᬤ᭄>-/*u.bad*/ ‘drug’
CV.VC	√			√		<ᬩᬳᭀᬲ᭄>-/*ba.os*/ ‘speech’ <ᬤᬳᬃ>-/*da.ar*/ ‘eat’
VC.CVC	√		√			<ᬳᬶᬫ᭄ᬧᬸᬲ᭄>-/*im.pus*/ ‘pick’ <ᬳᬸᬫ᭄ᬩᬄ>-/*um.bah*/ ‘wash’
VC.CV	√		√			<ᬳᬫ᭄ᬧᭀ>-/*am.po*/ ‘clay’ <ᬳᬸᬂᬲᬶ>-/*uŋ.si*/ ‘evacuate’
CV.V.CV		√		√		<ᬩᬳᬸᬮᬸ>-/*ba.u.lu*/ ‘padlock’ <ᬤᬳᬸᬲ>-/*da.u.sa*/ ‘names of plants for fences’
VC.CV.CVC		√	√			<ᬳᬸᬂᬲᬶᬮᬦ᭄>-/*uŋ.si.lan*/ ‘kidney’ <ᬳᭂᬫ᭄ᬧᬸᬕᬦ᭄>-/*əm.pu.gan*/ ‘full’
CV.CV.V		√			√	<ᬩᭂᬩᬳᬶ>-/*bə.ba.i*/ ‘evil spirits cause disease’ <ᬩᭂᬤᬳᬸ>-/*bə.da.u*/ ‘canoe’

Based on their syllable structure, eight canonical bisyllabic patterns (V.V, V.VC, V.CV, CV.V, V.CVC, CV.VC, VC.CVC, and VC.CV) and three canonical trisyllabic patterns (CV.V.CV, VC.CV.CVC, and CV.CV.V) were identified. Regarding the positional distribution of GPCs, five canonical patterns feature GPCs in the initial position (V.CV, V.CVC, VC.CVC, VC.CV, and VC.CV.CVC); one canonical pattern includes GPCs in both the initial and middle positions (V.VC); one pattern has GPCs in both the initial and final positions (V.V); two patterns exhibit GPCs in the middle position (CV.VC and CV.V.CV); and two patterns display GPCs in the final position (CV.V and CV.CV.V).

The morpheme structure condition refers to the constraints that define permissible sequences of phonemes within a languages canonical pattern (Schane, 1973). The permissible canonical patterns in a language play a significant role in shaping the orthography of its writing system by establishing systematic relationships between graphemic rules (graphotactics) and the principles governing GPC mapping (Coulmas, 1999; Lukatela et al., 1980; Pae & Wang, 2022; Pandey, 2014).

The eleven canonical patterns outlined above demonstrate that GPCs for the six vowels consistently appear in the syllable-initial position. As a result, the graphotactic rules of the Balinese script require all vowel phonemes in this position to be represented as independent graphemes. They cannot be rendered using bound graphemes as vowel markers attached to consonants, nor can they remain unmarked to indicate the inherent vowel /a/.

This feature is one of the fundamental characteristics of Abugida systems, classified as Alphasyllabaries, where vowels are represented by full-sized graphemes for independent positions—in particular, for syllable-initial positions (Fedorova, 2013; Pandey, 2014).

© 2025 ACADEMY PUBLICATION

In the orthography of the Balinese script, one approach to representing vowels in syllable-initial positions involves utilizing the akshara |ᬳ| /ha/, which is combined with vowel diacritics, known as *pangangge suara* (e.g., |◌ᬶ|/i/, |◌ᬸ|/u/, |◌ᬾ|/e/, |◌ᭀ|/o/, |◌ᭂ|/ə/) as needed (Medera et al., 2003; Simpen AB, 1979; Tinggen, 1994). The voiceless glottal phoneme /h/ in the primary consonant grapheme <ᬳ> /ha/ is silent or treated as "empty." Consequently, the GPC mapping for canonical patterns should be represented as <ᬳ> /a/, <ᬳᬶ> /i/, <ᬳᬸ> /u/, <ᬳᬾ> /e/, <ᬳᭀ> /o/, and <ᬳᭂ> /ə/. Mapping them as <ᬳ> */ha/, <ᬳᬶ> */hi/, <ᬳᬸ> */hu/, <ᬳᬾ> */he/, <ᬳᭀ> */ho/, and <ᬳᭂ> */hə/ would be inaccurate. This approach aligns with the "empty consonant device" strategy, commonly observed in Abugida-derived writing systems.

C. GPC Rules

Four types of generative phonological rules serve as the foundation for formulating GPC principles: (1) feature-changing rules, (2) deletion and insertion rules, (3) permutation and coalescence rules, and (4) variable-based rules (Schane, 1973). This study identifies two key GPC rules: deletion rules and sequentially ordered rules.

(a). Deletion Rules /h/ on Underlying GPC <ᬳ>-/ha/

The voiceless glottal consonants /h/ [-sil, +low] in the underlying GPC <ᬳ> /ha/ become zero <ᬳ> /Øa/ in the initial, medial, and final positions of the root morpheme, as demonstrated in the following lemma examples.

<**ᬳ**ᬩᬸ> /***a***.*bu*/ ‘ash’ (**V.**CV)

<**ᬳ**ᬫ᭄ᬧᭀ> /***a****m.po*/ ‘clay’ (**V**C.CV)

<ᬳᭀ**ᬳᬄ**> /*o.**a**h*/ ‘change (**V**.VC)

<ᬤ**ᬳᬃ**> /*da.**a**r*/ ‘eat’ (CV.**V**C)

<ᬳ**ᬳ**> /*a.**a***/ ‘fig tree’ (V.**V**)

<ᬲᬫᬸ**ᬳ**> /*sa.mu.**a***/ ‘forum’ (CV.CV.**V**)

The GPC rules for the deletion of the voiceless glottal consonant /h/ in the underlying GPC <ᬳ> /ha/ in formal notation are expressed as follows:

<ᬳ> /ha/ → <ᬳ> /Øa/ / { #__ ; #__# ; __# }

The rule above demonstrates that the voiceless glottal consonant /h/ [-sil, +low] in the underlying GPC <ᬳ> /ha/ undergoes elision in all positions (initial, medial, final) of the root morpheme. Upon closer inspection of the phonetic environment where this change occurs based on canonical patterns (syllable structure), it is evident that all instances appear in the syllable-initial position. Unfortunately, in generative phonological rules, there is no specific notation for marking syllable nodes. Considering that syllable nodes or boundaries significantly influence GPCs in an akshara-based writing system, Pandey (2014) uses the notation (σ) to mark syllable nodes or boundaries when formulating akshara-to-sound rules in Hindi. Pandey’s concept can be adapted in this study to simplify the GPC according to the basic principles of generative phonology, where abstraction and rules aim to streamline a system (Schane, 1973). The rule can be simplified by stating that the syllable-initial position becomes the environment where the underlying GPC <ᬳ> /ha/ changes to <ᬳ> /a/. The formal notation of the rule becomes:

<ᬳ> /ha/ → <ᬳ> /Øa/ / σ__

(b). Ordered Rule

The transformation of underlying GPCs into derived forms indicates the occurrence of various phonological processes. GPC rules play a crucial role in elucidating and explaining the mechanisms by which a writing system functions. Data analysis reveals that several root morphemes, represented as lemmas, undergo multiple rule applications

© 2025 ACADEMY PUBLICATION

due to grapheme addition or interaction. These processes occur sequentially. Table 2 illustrates that the underlying GPC is positioned in the upper-left corner, while the processes on the left demonstrate the application of GPC rules that transform the underlying form into derived forms. The resulting derived forms are positioned in the lower-left corner.

TABLE 2
ORDERED RULES

Left		Right	
1. Ordered rule for <ꦲꦶ>/i/			
Underlying	#<ꦫ+ꦲ+◌ꦶ+ꦤ> /ra+ha+i+na/#	Underlying	# <ꦫ+ꦲ+◌ꦶ+ꦤ> /ra+ha+i+na/#
Heightening low vowel	#<ꦫ+ꦲꦶ+ꦤ> /ra+hi+na/#	Deletion voiceless glottal	#<ꦫ+ꦲꦶ+ꦤ> /ra+ai+na/#
Deletion voiceless glottal	#<ꦫ+ꦲꦶ+ꦤ> /ra+i+na/#	Heightening low vowel	#<ꦫ+ꦲꦶ+ꦤ> /ra+ii+na/#
Derived	<ꦫꦲꦶꦤ> /*raina*/ 'day, afternoon'	Derived	<ꦫꦲꦶꦤ> */*raiina*/
2. Ordered rule for <ꦲꦸ>/u/			
Underlying	#<ꦢ+ꦲ+◌ꦸ+ꦱ> /da+ha+u+sa	Underlying	#<ꦢ+ꦲ+◌ꦸ+ꦱ> /da+ha+u+sa/#
Heightening low vowel	#<ꦢ+ꦲꦸ+ꦱ> /da+hu+sa/#	Deletion voiceless glottal	#<ꦢ+ꦲꦸ+ꦱ> /da+au+sa/#
Deletion voiceless glottal	#<ꦢ+ꦲꦸ+ꦱ> /da+u+sa/#	Heightening low vowel	#<ꦢ+ꦲꦸ+ꦱ> /da+uu+sa/#
Derived	<ꦢꦲꦸꦱ> /*dausa*/ 'plants for fences'	Derived	<ꦢꦲꦸꦱ> */*dauusa*/
3. Ordered rule for <ꦲꦺ>/e/			
Underlying	#<ꦕ+ꦺ◌+ꦲ+ꦺ◌+◌ꦁ> /ca+e+ha+e+ŋ/#	Underlying	#<ꦕ+ꦺ◌+ꦲ+ꦺ◌+◌ꦁ> /ca+e+ha+e+ŋ/#
Heightening low vowel	#<ꦕꦺ+ꦲꦁ> /ce+heŋ/#	Deletion voiceless glottal	#<ꦕꦺ+ꦲꦁ> /ce+aeŋ/#
Deletion voiceless glottal	#<ꦕꦺ+ꦲꦁ> /ce+eŋ/#	Heightening low vowel	#<ꦕꦺ+ꦲꦁ> /ce+eeŋ/#
Derived	<ꦕꦺꦲꦁ> /*ceeŋ*/ 'measuring rice'	Derived	<ꦕꦺꦲꦁ> */*ceeeŋ*/
4. Ordered rule for <ꦲꦺꦴ>/o/			
Underlying	#<ꦧ+ꦲ+ꦺ◌ꦴ+◌ꦁ> /ba+ha+o+ŋ/	Underlying	#<ꦧ+ꦲ+ꦺ◌ꦴ+◌ꦁ> / ba+ha+o+ŋ/#
Heightening low vowel	#<ꦧ+ꦲꦺꦴ+◌ꦁ> /ba+ho+ŋ/#	Deletion voiceless glottal	#<ꦧ+ꦲꦺꦴ+◌ꦁ> /ba+ao+ŋ/#
Deletion voiceless glottal	#<ꦧ+ꦲꦺꦴꦁ> /ba+oŋ/#	Heightening low vowel	#<ꦧ+ꦲꦺꦴ+◌ꦁ> /ba+oo+ŋ/#
Derived	<ꦧꦲꦺꦴꦁ> /*baoŋ*/ 'neck'	Derived	<ꦧꦲꦺꦴꦁ> */*baooŋ*/
5. Ordered rule for <ꦲꦼ>/ə/			
Underlying	#<ꦏ+◌ꦼ+ꦲ+◌ꦼ+ꦠ+◌꧀> /ka+ə+ha+ə+ta+Ø/#	Underlying	#<ꦏ+◌ꦼ+ꦲ+◌ꦼ+ꦠ+◌꧀> /ka+ə+ha+ə+ta+Ø/#
Heightening low vowel	#<ꦏ+ꦲꦼ+ꦠ꧀> /kə+hə+t/#	Deletion voiceless glottal	#<ꦏ+ꦲꦼ+ꦠ꧀> /ka+aə+t/#
Deletion voiceless glottal	#<ꦏ+ꦲꦼ+ꦠ꧀> /ka+ə+t/#	Heightening low vowel	#<ꦏ+ꦲꦼ+ꦠ꧀> /kə+əə+t/#
Derived	<ꦏꦲꦼꦠ꧀> /*kəət*/ 'slice'	Derived	<ꦏꦲꦼꦠ꧀> */*kəəət*/

The ordered rule described above demonstrates that the heightening of the low vowel occurs when the underlying GPC <ꦲ> /ha/ is combined with bound graphemes for vowel marking, specifically <◌ꦶ>/i/, <ꦲꦸ>/u/, <ꦲꦺ>/e/, <ꦲꦺꦴ>/o/, and <ꦲꦼ>/ə/. This combination results in the transformation of the vowel /a/ with the features [+low] into /i/, /u/, /e/, /o/, and /ə/ with the features [-low], followed by the deletion of the voiceless glottal consonant /h/ to zero /Ø/. This rule posits that the derived form on the left results in the correct derived form because the GPC rules are applied sequentially. Specifically, the heightening of the low vowel rule must be applied before the voiceless glottal /h/ deletion rule. Reversing the order of these rules, as seen in the derived form on the right, produces an unacceptable or unattested form (*). These rules are applicable across all positions: initial, medial, and final morpheme positions. However, for the

© 2025 ACADEMY PUBLICATION

GPC <ᬳᭂ>-/ə/, the rule is not found to apply in the final position of the morpheme, as the realization of the vowel /a/ in the final morpheme undergoes weakening and is consistently pronounced as [ə] in Balinese (Pastika, 2005).

V. Conclusion

The exact conditions represent definitive statements that underpin the development of GPC alternation rules. These rules, which systematically map sequences of orthographic units onto their corresponding linguistic forms, provide a foundation for analyzing the transparency of a writing system. The first exact condition stipulates that the GPCs <ᬳ> /a/, <ᬳᬶ> /i/, <ᬳᬸ> /u/, <ᬳᬾ> /e/, <ᬳᭀ> /o/, and <ᬳᭂ> are derived forms of the underlying GPC <ᬳ> /ha/, which is classified as a free grapheme, and <◌ᬶ>/i/, <◌ᬸ>/u/,<◌ᬾ>/e/,<◌ᭀ>/o/, <◌ᭂ>/ə/ are bound graphemes used as vowel markers. The second exact condition addresses the morpheme structure, which explains the surrounding GPC environment that triggers the transformation of the underlying GPC into the derived GPC, such as canonical patterns, positions within the morpheme (initial, medial, final), and syllable positions (initial, medial, final). The third exact condition concerns the rule for the transformation of the underlying GPC <ᬳ> /ha/ into <ᬳ> /a/, which is expressed as a voiceless glottal consonant /h/ deletion occurring at syllable-initial positions. The transformation rule for the derived GPCs <ᬳᬶ> /i/, <ᬳᬸ> /u/, <ᬳᬾ> /e/, <ᬳᭀ> /o/, <ᬳᭂ> /ə/, formed by the combination of the free underlying GPC <ᬳ> /ha/ and the bound vowel-marking GPCs <◌ᬶ>/i/, <◌ᬸ>/u/,<◌ᬾ>/e/,<◌ᭀ>/o/, <◌ᭂ>/ə/, is expressed through the sequential application of heightening low vowel rules and voiceless glottal /h/ deletion rules. These sequential rules apply across all positions (initial, medial, and final) within the morpheme, with exception of the GPC <ᬳᭂ> /ə/, which does not occur in the final position of the morpheme.

A comprehensive analysis of the cognitive linguistic mechanisms that explain the systematic mapping of orthography to phonology, across various sublexical levels of representation, is crucial for advancing universal theories of reading and writing across diverse writing systems. This is especially pertinent to minority writing systems, such as the Balinese script, which have not been thoroughly explored. Fundamental concepts in phonology—such as distinctive features, phonological processes, phonological rules, and language unit boundaries—including sublexical representations such as syllable structure, syllable weight, and mora—are firmly rooted in linguistic theory and can be leveraged to identify regularities in orthography-phonology mapping. This framework holds the potential to address ambiguities in orthography-phonology correspondences. However, these concepts have yet to be widely applied to explain the complexities of writing systems, thereby offering significant opportunities for future research.

Acknowledgements

The authors would like to thank Ministry of Education, Culture, Research, and Technology, Universitas Pendidikan Ganesha, Bali, Indonesia (Grant No. 903/UN48.16/LT/2024).

References

[1] Astiti, S. L. W. C. (2023). *Analisis Representasi Unit Linguistik Grafem Aksara Bali Dalam Kamus Bali – Indonesia Beraksara Latin Dan Bali Tahun 2016* [Analysis of the Linguistic Unit Representation of Balinese Graphemes in the 2016]. Balinese-Indonesian Dictionary Written in Latin and Universitas Pendidikan Ganesha.

[2] Astiti, S. L. W. C., Paramarta, I. K., & Martha, I. N. (2023a). Representasi Fonem dalam Grafem Aksara Bali pada Kamus Bali-Indonesia Beraksara Latin dan Bali Tahun 2016 [Phoneme Representation in Balinese Graphemes in the 2016 Balinese-Indonesian Dictionary Written in Latin and Balinese Script]. *Jurnalistrendi: Jurnal Linguistik, Sastra, Dan Pendidikan*, *8*(1), 76–89. https://doi.org/10.51673/jurnalistrendi.v8i1.1509

[3] Astiti, S. L. W. C., Paramarta, I. K., & Martha, I. N. (2023b). Representasi Morfem dalam Grafem Aksara Bali pada Kamus Bali-Indonesia Beraksara Latin dan Bali Tahun 2016 [Morpheme Representation in Balinese Graphemes in the 2016 Balinese-Indonesian Dictionary Written in Latin and Balinese Script]. *Jurnal Pendidikan Dan Pembelajaran Bahasa Indonesia*, *12*(1), 91–102.

[4] Bassetti, B., & Atkinson, N. (2015). Effects of orthographic forms on pronunciation in experienced instructed second language learners. *Applied Psycholinguistics*, *36*(1), 67–91.

[5] Bosch, A. van den, & Daelemans, W. (1993). Data-oriented methods for grapheme-to-phoneme conversion. *Proceedings of the Sixth Conference on European Chapter of the Association for Computational Linguistics*, 45–53.

[6] Bright, W. (1996). The Devanagari Script. In P. T. Daniels & W. Bright (Eds.), *The World's Writing System* (pp. 384–391). Oxford University Press.

[7] Burani, C., Barca, L., & Ellis, A. W. (2006). Orthographic complexity and word naming in Italian: Some words are more transparent than others. *Psychonomic Bulletin and Review*, *13*(2), 346–352. https://doi.org/10.3758/BF03193855

© 2025 ACADEMY PUBLICATION

[8] Chomsky, N., & Halle, M. (1968). *Sound Pattern of English*. Harper and Row.
[9] Coltheart, M., Curtis, B., Atkins, P., & Haller, M. (1993). Models of reading aloud: Dual-route and parallel-distributed-processing approaches. *Psychological Review*, *100*(4), 589–608. https://doi.org/10.1037/0033-295x.100.4.589
[10] Coltheart, M., Rastle, K., Perry, C., Langdon, R., & Ziegler, J. (2001). DRC: A dual route cascaded model of visual word recognition and reading aloud. *Psychological Review*, *108*(1), 204–256. https://doi.org/10.1037/0033-295X.108.1.204
[11] Computer Science Undiksha. (2023). *ToLatin* (2.0.2). Google Playstore. https://play.google.com/store/apps/details?id=id.ac.undiksha.tolatin
[12] Coulmas, F. (1999). *The Blackwell Encyclopedia of Writing System*. Blackwell Publishers.
[13] Divay, M., & Vitale, A. J. (1997). Algorithms for Grapheme-Phoneme Translation for English and French: Applications for Database Searches and Speech Synthesis. *Computational Linguistics*, *23*(4), 495–523.
[14] Fedorova, L. L. (2013). The development of graphic representation in Abugida writing: The Akshara's grammar. *Lingua Posnaniensis*, *55*(2), 49–66. https://doi.org/10.2478/linpo-2013-0013
[15] Frost, R., Katz, L., & Bentin, S. (1987). Strategies for visual word recognition and ortographic dept: A Multilingual comparison. *Journal of Experimental Psychology: Human Perception and Performance*, *13*, 10–115.
[16] Gnanadesikan, A. E. (2017). Towards a typology of phonemic scripts. *Writing Systems Research*, *9*(1), 14–35. https://doi.org/10. 1080/17586801.2017.1308239
[17] Gonzalvo, X., & Podsiadło, M. (2014). Text-To-Speech with cross-lingual Neural Network-based grapheme-to-phoneme models. *Interspeech*, 3–7.
[18] Jared, D. (2002). Spelling-sound consistency and regularity effects in word naming. *Journal of Memory and Language*, *46*(4), 723–750. https://doi.org/10.1006/jmla.2001.2827
[19] Joyce, T., & Meletis, D. (2021). Alternative criteria for writing system typology. *Zeitschrift Für Sprachwissenschaft*, *40*(3), 257–277. https://doi.org/10.1515/zfs-2021-2030
[20] Lindqvist, C. (2001). *Scandinavian writing systems in comparison*. De Gruyter.
[21] Lukatela, G., Popadic, D., & Ognjenovic, P. (1980). Lexical decision in a phonologically. *Memory and Cognition*, *8*(2), 124–132.
[22] Medera, I. N., Suasta, I. B. M., & Japa, I. W. (2003). *Pedoman Pasang Aksara Bali* [Guidelines for Writing Balinese Script]. Dinas Kebudayaan Provinsi Bali.
[23] Meletis, D. (2019). The grapheme as a universal basic unit of writing. *Writing Systems Research*, *0*(0), 257–277. https://doi.org/10.1080/17586801.2019.1697412
[24] Meletis, D. (2020). Types of allography. *Open Linguistics*, *6*(1), 249–266. https://doi.org/10.1515/opli-2020-0006
[25] Meletis, D., & Dürscheid, C. (2022). Writing Systems and Their Use An Overview of Grapholinguistic. In *De Gruyter Mouton*. https://doi.org/10.1515/9783110757835
[26] Moitra, S., Chacón, D. A., & Stockall, L. (2024). How long is long? Word length effects in reading correspond to minimal graphemic units: An MEG study in Bangla. *PLoS ONE*, *19*(4 April), 1–13. https://doi.org/10.1371/journal.pone.0292979
[27] Nag, S. (2014). Akshara-phonology mappings: The common yet uncommon case of the consonant cluster Akshara-phonology mappings: The common yet uncommon. *Writing Systems Research*, *6:1*(August 2014), 105–119. https://doi.org/10.1080/17586801.2013.855621
[28] Nag, S., Treiman, R., & Snowling, M. J. (2010). Learning to spell in an alphasyllabary: The case of Kannada. *Writing Systems Research*, *2*(1), 41–52. https://doi.org/10.1093/wsr/wsq001
[29] Nala Antara, I. G., Suardiana, I. W., & Wisnu, I. W. G. (2016). *Kamus Bali-Indonesia beraksara Latin dan Bali Edisi II [Balinese dictionary - Indonesian with Latin and Balinese script]*. Badan Pembina Bahasa Aksara, dan Sastra Bali.
[30] Neef, M. (2012). Graphematics as part of a modular theory of phonographic writing systems. *Writing Systems Research*, *4*(2), 214–228. https://doi.org/10.1080/17586801.2012.706658
[31] Pae, H. K. (2020). Script effects as the hidden drive of the human mind, cognition, and culture. In *Sociolinguistic Studies* (Vol. 17, Issues 1–3). Springer. https://doi.org/10.1558/sols.23457
[32] Pae, H. K., & Wang, M. (2022). The effects of writing systems and scripts on cognition and beyond: An introduction. *Reading and Writing*, *35*(6), 1315–1321. https://doi.org/10.1007/s11145-022-10289-z
[33] Pandey, P. (2014). Akshara-to-sound rules for Hindi. *Writing Systems Research*, *6*(1), 54–72. https://doi.org/10.1080/17586801.2013.855622
[34] Paramarta, I. K., Indrawan, G., & Rai, I. B. (2024). Silent Letters in the Balinese Script |… ɿ| Adeg Adeg: A Graphetic and Graphematic Feature Analysis. *Theory and Practice in Language Studies*, *14*(4), 1206–1218. https://doi.org/10.17507/tpls.1404.28
[35] Paramarta, I. K., Martha, I. N., & Astiti, S. L. W. C. (2024). Pemetaan Aksara-Silabel Berbasis Grafem Aksara Bali dalam Kamus Bali – Indonesia Beraksara Latin dan Bali Tahun 2016 [Grapheme-Based Syllable-Script Mapping in the 2016 Balinese-Indonesian Dictionary Written in Latin and Balinese Script]. *GHANCARAN: Jurnal Pendidikan Bahasa Dan Sastra Indonesia*, *5*(2), 338–351. https://doi.org/10.19105/ghancaran.v5i2.9899
[36] Pastika, I. W. (2005). *Fonologi Bahasa Bali Sebuah Pendekatan Generatif Transformasi* [Phonology of the Balinese Language: A Transformational Generative Approach]. Pustaka Larasan.
[37] Pike, K. L. (1956). *Phonemic: A Technique for Reducing Language to Writing*. University of Michigan Press.
[38] Prince, A., & Smolensky, P. (1993). *Optimality Theory: Constraint Interaction in Generative Grammar. RuCCS Technical Report 2*.
[39] Protopapas, A., & Vlahou, E. L. (2009). A comparative quantitative analysis of Greek orthographic transparency. *Behavior Research Methods*, *41*(4), 991–1008. https://doi.org/10.3758/BRM.41.4.991
[40] Rajan, V. (2018). *Aksharamukha: Script Converter*. https://www.aksharamukha.com/about
[41] Rogers, H. (1995). Optimal orthographies. In I. Taylor & D. R. Olson (Eds.), *Scripts and literacy: Reading and learning to read alphabets, syllabaries and characters (Neuropsychology and Cognition 7* (pp. 31–43). Springer.
[42] Schane, S. A. (1973). *Generative Phonology Foundations of modern linguistics Prentice-Hall foundations of modern*

© 2025 ACADEMY PUBLICATION

linguistics series. Prentice-Hall.
[43] Simpen AB, I. W. (1979). *Pasang Aksara Bali* [Rules of Writing in Balinese Script]. Dinas Pengajaran Provinsi Daerah Tingkat I Bali.
[44] Sircar, S., & Nag, S. (2014). Akshara-syllable mappings in Bengali: A language-specific skill for reading. In H. Winskel & P. Padakannaya (Eds.), *South and Southeast Asian psycholinguistics* (pp. 202–211). Cambridge University Press.
[45] Sproat, R. (2000). *A computational theory of writing systems*. Cambridge University Press.
[46] Suasta, I. B. M. (2006). Ejaan Aksara Bali Dalam Pelestarian Budaya Bali [Orthography of Balinese Script in the Preservation of Balinese Culture]. In *Kumpulan Makalah Kongres Bahasa Bali VI*.
[47] Tinggen, I. N. (1994). *Celah-Celah Kunci Pasang Aksara Bali* [Key Gaps in the Rules of Writing Balinese Script]. RHIKA DEWATA.
[48] Treiman, R., Mullennix, J., Bijeljac-Babic, R., & Richmond-Welty, E. D. (1995). The Special Role of Rimes in the Description, Use, and Acquisition of English Orthography. *Journal of Experimental Psychology: General*, *124*(2), 107–136. https://doi.org/10.1037/0096-3445.124.2.107
[49] Vaid, J., & Gupta, A. (2002). Exploring word recognition in a semi-alphabetic script: The case of Devanagari. *Brain and Language*, *81*(1–3), 679–690. https://doi.org/10.1006/brln.2001.2556
[50] Ziegler, J. C., Bertrand, D., Tóth, D., Csépe, V., Reis, A., Faísca, L., Saine, N., Lyytinen, H., Vaessen, A., & Blomert, L. (2010). Orthographic depth and its impact on universal predictors of reading: A cross-language investigation. *Psychological Science*, *21*(4), 551–559. https://doi.org/10.1177/0956797610363406
[51] Ziegler, J. C., & Goswami, U. (2005). Reading acquisition, developmental dyslexia, and skilled reading across languages: A psycholinguistic grain size theory. *Psychological Bulletin*, *131*(1), 3–29. https://doi.org/10.1037/0033-2909.131.1.3

I K. Paramarta was born in Sulahan Village, Bangli Regency, Bali Province, Indonesia. He completed his Doctor of Linguistics at Udayana University (2015). His research interests are the study of the Balinese language from a comparative historical linguistic and sociolinguistic perspective, and the study of the Balinese script system from a grapholinguistic perspective. He is a lecturer in the Balinese Language Education Study Program, Faculty of Languages and Arts at Universitas Pendidikan Ganesha, Bali Indonesia.

I Wayan Rasna is a linguistics professor at the Ganesha University of Education. He has 35 years of experience conducting research and community service focusing on linguistics, especially those related to educational linguistics, anthropological linguistics, and education in cultural contexts.

I B. Putrayasa is a professor at the Ganesha University of Education in Singaraja, Bali. He completed his doctoral degree at the Indonesian University of Education, Bandung (1998-2001), and earned a professorship on October 1, 2004. Currently he is active in various academic activities, such as teaching in the Undergraduate, Masters and Doctoral Programs at Undiksha Singaraja.

I Nengan Suandi was born in Klungkung on December 31, 1956. He completed his doctoral at Udayana University in 2007. He was appointed professor in 2010 in the field of linguistics. Before becoming a lecturer at the Ganesha University of Education.

© 2025 ACADEMY PUBLICATION

Kadek Wirahyuni earned her Doctoral in Language Education from Universitas Pendidikan Ganesha in 2022. Her research interests focus on the study of public speaking. She is currently a lecturer in the Indonesian Language Education Study Program at Universitas Pendidikan Ganesha, Bali, Indonesia.

© 2025 ACADEMY PUBLICATION