

The Linguistic Architecture of Old Khmer: A Technical Analysis of Philip N. Jenner's Lexicographical Corpus

Published: April 13, 2025 | Index ID: #185

The study of **Old Khmer** represents one of the most significant undertakings in Southeast Asian historical linguistics and epigraphy. Central to this field is the monumental work of **Philip N. Jenner**, specifically his dual-volume contribution: *A Dictionary of Pre-Angkorian Khmer* and *A Dictionary of Angkorian Khmer*. Published under the **Pacific Linguistics** series by the Australian National University (ANU), these works provide a rigorous structural and etymological framework for understanding the evolution of the Khmer language from the 6th century through the decline of the Angkorian Empire. This article provides an in-depth technical analysis of Jenner's methodology, the linguistic shifts identified between Pre-Angkorian and Angkorian periods, and the broader implications for Mon-Khmer studies.

1. Historical and Epigraphical Context

Old Khmer is categorized primarily through its epigraphical record—inscriptions carved onto stone stelae, lintels, and doorframes of temples. These records are divided into two distinct chronological phases:

- **Pre-Angkorian Khmer (6th to 8th Century)**: Corresponding to the periods of Funan and Chenla. These inscriptions are characterized by a higher degree of linguistic conservatism and a specific set of orthographic conventions.
- **Angkorian Khmer (9th to 14th Century)**: Corresponding to the rise of the Khmer Empire under Jayavarman II and his successors. This period shows significant evolution in syntax, the absorption of extensive Sanskrit vocabulary, and shifts in phonological structures.

Jenner's research involves the dissection of over **1,100 inscriptions**. This corpus is not merely a list of words but a mapping of the socio-political and religious transformations of the Khmer people. The dictionaries serve as a bridge between the archaic forms of the language and **Middle Khmer**, and eventually, **Modern Khmer**.

2. Technical Methodology: Lexicographical Reconstruction

Philip N. Jenner's approach to lexicography in the context of Old Khmer is built upon **comparative philology** and **structural linguistics**. The methodology can be broken down into four primary phases:

Phase I: Epigraphical Collation

The primary data source consists of *estampages* (ink rubbings) of stone inscriptions. Jenner utilized the established corpus compiled by previous scholars like George Coedès but applied a more rigorous linguistic lens. Each entry in the dictionary is cross-referenced with its specific inscription number (e.g., K. numbers), ensuring that the context of every word is preserved.

Phase II: Morphological Decomposition

Khmer is a non-tonal, **Austroasiatic language** known for its complex system of **affixation**. Jenner's technical breakthrough was the systematic identification of prefixes and infixes that were productive in Old Khmer but have since become frozen or lost in Modern Khmer. He utilized a formulaic approach to identify the **root (R)** and the **derivative (D)**.

Phase III: Etymological Tracing

A significant portion of the Old Khmer lexicon consists of loanwords from **Sanskrit** and **Pali**. Jenner's dictionaries distinguish between indigenous Mon-Khmer roots and borrowed Indic terms. This involves tracing the phonetic adaptations of Sanskrit sounds (such as the reduction of clusters) into the Khmer phonological system.

Phase IV: Semantic Mapping

Because the inscriptions are often legalistic (land grants, lists of temple offerings, slave registers), Jenner had to reconstruct meanings based on repetitive usage patterns. This requires a deep understanding of the **Angkorian administrative hierarchy** and religious rituals.

3. Core Linguistic Mechanics of Old Khmer

To understand Jenner's work, one must grasp the mechanical differences between the two stages of the language. Below is a technical comparison of the phonological and morphological traits identified in his research.

Phonological Shifts

Pre-Angkorian Khmer possessed a vowel system that was likely simpler than the modern registers but complex in its vowel length distinctions. Jenner's analysis suggests that the **Great Vowel Shift** in Khmer, which led to the modern distinction between O-series and A-series vowels, had its roots in the late Angkorian period.

Table 1: Comparative Features of Pre-Angkorian and Angkorian Khmer

Feature	Pre-Angkorian (6th-8th C)	Angkorian (9th-14th C)	Linguistic Impact
Orthography	High Sanskrit influence; archaic scripts.	Standardization of the Khmer script.	Stabilization of literary forms.
Affixation	Highly productive infixes (-n-, -m-).	Increased use of compound words.	Shift from synthetic to analytic structure.
Honorifics	Developing social hierarchy terms.	Extremely complex royal/divine lexicon.	Reflects the Devaraja (God-King) cult.
Loanwords	Primarily religious/technical Sanskrit.	Widespread integration of Indic terms.	Lexical enrichment and stratification.

4. Morphological Analysis: The Infix System

One of the most technically demanding aspects of Jenner's work is the explanation of **Khmer morphology**. Unlike English, which relies heavily on suffixes, Old Khmer utilized **prefixes** and **infixes**. Jenner categorizes these into functional classes:

The Causative Prefix (p- / b-)

This prefix was used to transform an intransitive verb into a transitive or causative verb. For example, a root meaning "to be clean" would be prefixed to mean "to cleanse."

The Nominalizing Infix (-n-)

Jenner identifies the **-n- infix** as a primary tool for creating nouns from verbs. This is a technical hallmark of the Mon-Khmer family. In his dictionary, Jenner meticulously traces how a verb like **cam* (to wait/guard) becomes *cnam* (a guard or a period of time/year).

The Instrumental Infix (-m-)

This infix often denotes the instrument or the person performing an action. Jenner's dictionary entries for words containing the -m- infix provide a window into the daily technology and social roles of the Angkorian era.

5. Lexical Categories and Statistical Distribution

Jenner's corpus allows for a statistical analysis of the lexicon. Based on the **Pacific Linguistics Series 597 and 598**, the distribution of word types reveals the priorities of the Khmer civilization.

Lexical Category	Percentage (Approx.)	Primary Source Context			
Administrative/ Legal	35%	Land boundaries, taxes, legal disputes.	Ritual/Religious	30%	Offerings, deities, temple ceremonies.
Onomastics (Names)	20%	Lists of slaves (gnam), donors, and officials.			
Toponyms (Places)	10%	Village names, canals, and field locations.			
Abstract Concepts	5%	Merit (punya), time, and existence.			

6. Technical Case Study: Tracing the Word 'Vrah'

A quintessential example of Jenner’s lexicographical depth is his treatment of the word **Vrah** (modern *Preah*). In Jenner’s *Dictionary of Pre-Angkorian Khmer*, he tracks this term through various contexts:

- 1. As an adjective: Denoting sacredness or excellence.
- 2. As a noun: Referring to a deity or a king.
- 3. As a prefix: Creating sacred compounds (e.g., Vrah Pada - the sacred feet).

Jenner notes that the term originates from an indigenous root rather than Sanskrit, despite its central role in Hindu-Buddhist contexts. By analyzing 1,100+ inscriptions, he demonstrates how *Vra* evolved from a simple marker of respect to a foundational element of the state’s ideological apparatus.

7. Computational and Digital Humanities Applications

In recent years, the work of Philip N. Jenner has been transitioned into digital formats, notably through the **SEALANG Projects** and **Doug Cooper’s** editorial contributions. This digital transformation allows for computational linguistic analysis that was impossible when the volumes were first compiled.

Searchable Metadata and Regular Expressions

Modern researchers use Jenner’s data to perform **regex (regular expression) searches** across the corpus. This allows for the identification of phonetic patterns and the testing of hypotheses regarding **vowel mergers**. For example, a researcher can query all instances where a specific Sanskrit *sibilant* was replaced by a Khmer *aspirate*, providing empirical data for phonological evolution models.

Formula for Phonetic Alignment

In comparative studies, researchers apply an alignment formula to Jenner’s data to measure the distance between Old Khmer and its modern descendants:

$$D(L1, L2) = \sum (c_i * w_i) / N$$

Where:**D** is the linguistic distance.**c_i** is the cost of phonemic change (mutation, deletion, or insertion).**w_i** is the

frequency weight of the lexeme in Jenner's dictionary. N is the total number of compared tokens.

8. Challenges and Limitations in Old Khmer Lexicography

While Jenner's dictionaries are authoritative, they are not without technical challenges. Scholars must navigate several operational hurdles when using these texts:

- **Orthographic Inconsistency:** Scribes in different provinces often used varying spellings for the same word. Jenner provides "variant" entries, but the researcher must exercise judgment in reconciling these.
- **Lacunae (Gaps):** Stone erosion often leaves partial words. Jenner uses brackets [...] to denote reconstructed text, which requires a deep knowledge of formulaic inscription styles to interpret correctly.
- **Semantic Shift:** The danger of anachronism—applying modern meanings to ancient words—is a constant risk. Jenner mitigates this by strictly adhering to the contemporary Sanskrit definitions of borrowed terms.

9. Practical Implementation Guide for Researchers

For students and epigraphers using *A Dictionary of Pre-Angkorian Khmer*, the following workflow is recommended for accurate translation:

Step-by-Step Translation Protocol

1. **Identify the Inscription Number (K. No):** Locate the source of the text within the EFEO (École française d'Extrême-Orient) catalogue.
2. **Transliterate the Script:** Convert the Old Khmer script into a Romanized system (e.g., Cœdès or Jenner system).
3. **Isolate the Root:** Strip away prefixes and infixes to find the base lexeme using Jenner's morphological guides.
4. **Cross-Reference Series 597 and 598:** Check if the word exists in both the Pre-Angkorian and Angkorian dictionaries to see if the meaning or form shifted over time.
5. **Analyze the Syntactic Context:** Determine if the word is acting as a Vv (Verb), Nn (Noun), or Aj (Adjective) based on its position relative to particles like ai (at/in) or gi (is/are).

10. Broader Implications for Southeast Asian Linguistics

The significance of Jenner's work extends far beyond the borders of Cambodia. By documenting the earliest recorded stages of a major Mon-Khmer language, he provides the comparative data necessary to reconstruct **Proto-Mon-Khmer**. His dictionaries allow linguists to track the "Sanskritization" of Southeast Asia, a process that redefined the intellectual and cultural landscape of the region for a millennium.

Furthermore, the data compiled by Jenner serves as a critical resource for **historical Thai linguistics**. Since Early Thai (Sukhothai period) borrowed heavily from Angkorian Khmer, Jenner's work is essential for Thai etymologists attempting to untangle the complex layers of their own national language.

The preservation of this knowledge, originally disseminated through the **ANU Open Research** repository and **Pacific Linguistics**, ensures that the structural nuances of Old Khmer remain accessible to a global audience. As we move further into the digital age, Philip N. Jenner's meticulous lexicographical foundations remain the indispensable bedrock for all future inquiries into the ancient voices of the Khmer people. His work does more than define words; it reconstructs a lost world, one inscription at a time.

Tags: [#Old Khmer](#) [#Philip N. Jenner](#) [#Linguistics](#) [#Angkorian History](#) [#Lexicography](#)

