



An Optimality-Theoretic Analysis of Vietnamese Loanwords in the Bahnar Language

Tam T. T. Le^a , Quy T. Nguyen^b , Long S. T. Nguyen^c , Tho T. Quan^{d*} 

^a *University of Social Sciences and Humanities (USSH), Vietnam National University, Hanoi, Vietnam. Email: tamltt@vnu.edu.vn*

^b *University of Social Sciences and Humanities (USSH), Vietnam National University, Ho Chi Minh City, Vietnam. Email: nguyentranquy@hcmussh.edu.vn*

^c *Ho Chi Minh City University of Technology (HCMUT), Vietnam National University, Ho Chi Minh City, Vietnam. Email: long.nguyencse2023@hcmut.edu.vn*

^d *Ho Chi Minh City University of Technology (HCMUT), Vietnam National University, Ho Chi Minh City, Vietnam. Email: qttho@hcmut.edu.vn*

Received: 30 August 2025 | Received: in revised form 20 November 2025 | Accepted 20 December 2025

APA Citation:

Le, T. T. T., Nguyen, Q. T., Nguyen, L. S. T., Quan, T. T. (2025). An Optimality-Theoretic Analysis of Vietnamese Loanwords in the Bahnar Language *Eurasian Journal of Applied Linguistics*, 11(3), 296-307.
Doi: <http://dx.doi.org/10.32601/ejal.11326>

Abstract

this study aims to develop a phonological adaptation model of Vietnamese loanwords in the Binh Định dialect of Bahnar, an Austroasiatic language of Vietnam's Central Highlands. It also aims to construct a constraint hierarchy reflecting the phonological patterns of a minority language, and document and analyze lexical change in real time under sociolinguistic influences to identify common semantic domains in the loanword inventory. A corpus of 5,769 loanwords was compiled from dictionaries, field notes, and elicited data, and each item was coded for phonological features. Quantitative counts of adaptation strategies were combined with Optimality Theory (OT) tableaux to model constraint rankings. Results show that coda deletion is the dominant strategy (≈60%), followed by consonant substitution (25%) and vowel epenthesis (15%). Vietnamese tones are neutralized in borrowing, aligning with Bahnar's register-based prosody. Tableau analyses indicate that a ranking where *CODA and *COMPLEX outrank IDENT-IO and MAX-IO predicts these adaptations. Additionally, reduplication occurs in 9.5% of loanwords, underscoring its role in morphophonological integration. Interdialectal comparison reveals both convergence across Bahnar varieties and localized innovations in Binh Định. The findings support Haugen's model of borrowing and demonstrate that Bahnar speakers actively reshape foreign items to fit native phonotactics and sociolinguistic norms. This study provides the first systematic OT account of Bahnar loanword phonology, and contributes to comparative Austroasiatic studies, and offers a replicable framework for analyzing contact-induced change in minority languages.

© 2025 EJAL & the Authors. Published by Eurasian Journal of Applied Linguistics (EJAL). This is an open-access article distributed under the terms and conditions of the Creative Commons Attribution license (CC BY-NC-ND) (<http://creativecommons.org/licenses/by-nc-nd/4.0/>).

Keywords: Loanword Adaptation, Optimality Theory, Bahnar Language, Lexical Borrowing Typology, Austroasiatic Phonology, Tone Simplification

* Corresponding Author.

Email: qttho@hcmut.edu.vn

DOI: <http://dx.doi.org/10.32601/ejal.11326>

Introduction

The Bahnar language belongs to the Bahnar group within the Bahnar branch of the Mon-Khmer languages, which are part of the Austroasiatic language family (Thomas, 1966). In addition to Bahnar, this group includes Rengao, Sedang, Halang, Jeh, Monom, Kêyong, Hre, and Cua. The Bahnar community numbers approximately 287,000 people (Nam & Duc, 2021), with most people living in Kon Tum, Gia Lai, and Binh Dinh provinces. Gia Lai has the largest Bahnar population, estimated at about 190,000. This region is a linguistic contact zone among Bahnaric, Viet-Muong, and Cham languages, shaped by long-standing intercommunity interactions. In the past, such contact primarily occurred through trade and marriage; today it is increasingly facilitated by public education, mass media, and state administration.

Loanwords have long been viewed as indicators of language contact and drivers of language change (Haugen, 1950; Thomason & Kaufman, 2023; Winford, 2003). In the context of globalization, lexical borrowing is increasingly shaped by asymmetrical social and political relations. Borrowed words often reflect the influence of dominant languages on minority language ecologies. They not only expand the lexicon of the recipient language but also serve as conduits for cultural and administrative concepts. Vietnamese and Bahnar differ substantially in typology. Bahnar is sesquisyllabic, non-tonal, and morphologically complex, with numerous prefixes and infixes. Vietnamese is isolating, tonal (six tones), rich in segmental contrasts, and has minimal affixation. These differences create significant challenges for lexical borrowing, especially in phonological and morphological adaptation.

Despite extensive Vietnamese-Bahnar contact, there is limited empirical documentation of how loanwords are phonologically adapted in Bahnar, particularly using formal frameworks like Optimality Theory (OT). This gap hinders understanding of lexical change in minority Austroasiatic languages facing assimilation pressures. Studying the status of loanwords in Bahnar, and distinguishing borrowed from inherited vocabulary, is a complex challenge due to the long history and diversity of contact situations in an area with hundreds of languages. While the genetic affiliation of Bahnar was once debated, current evidence confirms that it retains a strong core of basic vocabulary inherited from Mon-Khmer.

This topic was selected due to the urgent need to document endangered minority languages like Bahnar amid rapid Vietnamese dominance, providing insights into ongoing lexical evolution and supporting implications for preservation in Southeast Asia's multilingual contexts. This study models the adaptation of Vietnamese loanwords in Bahnar from the perspective of OT, clarifying how Bahnar speakers reconcile the faithfulness of Vietnamese input forms with the markedness constraints of their native phonology. This study applies the OT framework of (Prince & Smolensky, 2004) to model the adaptation of Vietnamese loanwords in Bahnar. OT posits that surface forms are shaped by the interaction between universal and language-specific constraints. In loanword phonology, speakers negotiate between faithfulness constraints (e.g., IDENT-IO, MAX-IO) and markedness constraints (e.g., CODA, COMPLEX, *x*, *j*).

Phenomena such as coda deletion, epenthesis, substitution, and tone simplification were examined as outcomes of constraint interaction. This links the historical and typological background to the research objectives by emphasizing how structural mismatches drive innovative adaptations, contributing novel empirical evidence from a large corpus. Drawing on a systematically annotated corpus of 5,769 loanwords, this study aims to (i) develop a phonological adaptation model for underrepresented Austroasiatic languages; (ii) construct a constraint hierarchy reflecting the phonological patterns of a minority language; (iii) document and analyze lexical change in real time under sociolinguistic influences; (iv) identify common semantic domains in the loanword inventory. Overall, this study proposes an analytical framework for investigating borrowing in contexts where minority languages face structural and cultural pressures from dominant languages such as Vietnamese. This investigation is innovative in its scale and application of OT to Bahnar dialects, offering a replicable model for similar contact scenarios.

Literature Review

Loanwords

Loanwords are lexical items adopted from a source language into a recipient language, used in both speech and writing without necessarily translating their meaning (Haspelmath & Tadmor, 2009). (Haugen, 1950) defines loanwords as units that preserve both the sound form and meaning of the original. (Haspelmath & Tadmor, 2009) distinguish two major types: (i) Direct borrowing (adoption): retaining the original sound form and spelling (e.g., *cà phê* (*café*), *nhà trẻ* (*kindergarten*)); (ii) Adaptation: modifying phonemes, morphology, or syntax to fit the recipient language (e.g., *virus* → *fāyruṣ* in Arabic). In language-contact contexts, direct borrowing is often associated with bilingualism and carries concepts and ideas from the source language (Yule, 2006). Loanword recognition can be based on phonological or morphological features, or both (Ilmina, 2016).

Vietnamese itself has borrowed extensively from Chinese, French, English, Russian, and other sources

(Alves, 2006), and hence often Vietnamese is described as a donors' language. Alves (2006) estimates that Chinese-origin words account for 90% of Vietnamese loanwords, followed by French (4%), Thai (3%), and smaller proportions from English, Cham, and other languages. Borrowing in Vietnamese reflects historical layers: ancient Chinese, colonial French, and modern English. Thompson (1963) proposes three main motivations for using loanwords: clarifying meaning, shortening descriptive expressions into single-word units, and naming concepts without existing equivalents.

Nguyễn (2003) identifies the Mon-Khmer as the oldest loanword layer in Vietnamese, followed by Thai-Tay, Chinese, and Indo-European layers. Vietnamese lexicographical works (Nguyễn, 2003) show a dense distribution of loanwords across specialized fields, providing a basis for hypothesizing how Bahnar may have received Vietnamese vocabulary in semantic domains such as administration, technology, and education.

Bahnar As a Recipient Language

Bahnar belongs to the Northern Bahnar branch of the Austroasiatic language family, which also includes Vietnamese (Alves, 2006). Some basic vocabulary items in Bahnar are cognate with Vietnamese, particularly words for natural phenomena, numerals, and body parts (Ferlus, 1999). The Bahnar phonological system, however, differs significantly: it is non-tonal, has a sesquisyllabic structure, and employs prefixes and infixes, while Vietnamese is tonal, isolating, and has minimal affixation. These differences necessitate systematic phonological adaptation when Vietnamese words are borrowed. Similar adaptation processes are observed in Tay-Nung (Mai & Le Phan, 2025), for example: *ket* (Tay-Nung) from Vietnamese *hát*; *van* (TN) from Vietnamese *văn*; and the *sắc* (high) tone mapping to the *nặng* tone in Tay-Nung.

Optimality Theory and Adaptation Strategies

Trang (2021) describes "repair strategies" that may occur during perception, articulation, or both as adaptation measures. Scholvin (2022) emphasizes the importance of identifying which language's perceptual filter-source or target-dominates in adaptation. In Bahnar, markedness constraints such as *CODA* and */*often outrank faithfulness, leading to systematic changes such as coda deletion, epenthesis, substitution, and tone simplification. Studies like (LaCharité & Paradis, 2005; Trang 2021) show that such adaptations are not random but reflect the phonological constraints typical of Austroasiatic languages.

Social factors also play a crucial role in borrowing. Language policy, education, and the prestige of the source language strongly influence the degree of loanword adoption (Scholvin, 2022). In Bahnar-speaking regions, Vietnamese dominates in education, administration, and mass media facilitating vocabulary expansion through borrowing. Studies on Bahnar and other minority languages (Bùi et al., 2006; Phong, 2005; Van Van, 2020) indicate that lexical change is not purely linguistic but also socio-cultural. Borrowing can serve communicative needs, stylistic innovation, identity expression (Eckert, 2000), and integration into national norms in a multi-ethnic society. By integrating historical, typological, and sociolinguistic perspectives, this paper establishes a solid theoretical foundation for analyzing how Bahnar speakers receive and adapt Vietnamese vocabulary at multiple structural levels.

Recent empirical research on loanword adaptation in Vietnamese and Austroasiatic contexts deserve special mention. Scholvin and Meinschaefer (2018) empirically analyzed the phonological integration of French loanwords into Vietnamese, focusing on tone and syllable structure adjustments using acoustic data from native speakers, revealing systematic constraint-based repairs similar to those in Bahnar. Alves (2024) conducted a corpus-based study of Austroasiatic components in Vietnamese, quantifying loanword proportions across semantic domains and highlighting bidirectional borrowing in contact zones, which parallels the Vietnamese-to-Bahnar flow observed here. Additionally, Brunelle (2009) provided acoustic evidence of tone coarticulation in Northern Vietnamese, offering a baseline for comparing tone simplification in Bahnar adaptations. These studies create impact by demonstrating data-driven patterns of adaptation, motivating further exploration of underrepresented languages like Bahnar and addressing arguments on phonological dominance in contact (e.g., markedness over faithfulness).

Verdonschot et al. (2022) experimentally investigated phonological encoding in Vietnamese, using picture-word interference tasks to show that the initial phonological unit in Vietnamese word production is the phoneme, with implications for tone adaptation in loanwords, as contour tones are simplified under native constraints. Coto-Solano and Trần (2021) explored priming effects of tones in visual processing of Vietnamese, revealing how tonal cues influence lexical access, relevant to Bahnar's tone restructuring. Ha (2022) examined code-mixing and loanwords in Vietnamese vocabulary, identifying patterns of integration from Chinese, French, and English, which inform the sociolinguistic motivations for Vietnamese loanwords in Bahnar. Cao (2022) analyzed lexical-phonological patterns of Tai loans in Vietnamese, highlighting segmental adaptations and diagnostics for contact timing, paralleling Austroasiatic influences. Alves (2024) provided an updated overview of Austroasiatic components in Vietnamese, identifying nearly 200 etyma and noting bidirectional borrowing, while considering first-millennium CE regional loanwords related to material culture, offering chronological context for Vietic-Austroasiatic interactions. Person (2022) described the phonology of Li Xei, an understudied Bahnaric language, documenting syllable structures and segmental inventories similar to

Bahnar, aiding comparative analyses. Sidwell (2024) reconstructed 500 Proto-Austroasiatic etyma, providing a lexical database for tracing inherited vs. borrowed elements in Bahnar. Mahdi (2024) reviewed Austroasiatic loanwords in Austronesian languages, highlighting areal diffusion patterns applicable to Mainland Southeast Asia contacts.

Comparative Optimality Theory (OT) studies from other languages motivate the framework's application here. Allen (2020) synthesized research on English loanwords in Japanese, emphasizing cognate effects and perceptual adaptations. Abbasi and Khanam (2025) applied OT to English loanwords in Urdu, modeling variable adaptations in constraint rankings. Alhoody (2019) offered an OT account of English loanwords in Qassimi Arabic, focusing on segmental repairs like substitution and epenthesis. Abdulrazzaq and Al-Ubaidy (2023) contrasted OT with perceptual and phonological models in loanword adaptation, advocating for integrated approaches. Shafi (2017) explored variable adaptation of English loanwords in Mirpur Pahari using OT, addressing suprasegmental variation. Naika (2021) analyzed morphophonological adaptations of Oluwanga loanwords via Optimality Theory (OT), highlighting constraint interactions in African contexts. These empirical insights underscore gaps in OT applications to Austroasiatic minority languages like Bahnar, motivating this study's corpus-based analysis.

Methodology

Research Design

This study adopted a mixed-method research design, combining phonetic modeling within the OT framework, acoustic analysis, and cross-dialect comparison. This design enables the identification of both systematic patterns in loanword adaptation and regional variations within the Bahnar corpus.

Sampling and Research Procedure

The primary dataset was drawn from a 320,000-word Bahnar corpus compiled from materials in Binh Dinh, Gia Lai, and Kon Tum provinces. Sampling targeted diverse sources for representativeness: radio broadcasts (n=150 hours), folk tales (n=50 texts), educational materials (n=100 recordings), and field data (n=50 speakers). A total of 5,769 tokens were annotated for origin, strategy, and domain. From these, a subset of 210 was selected for tone analysis based on frequency and dialect balance.

Data Collection

Data included radio broadcasts, folk tales, and recordings from educational materials. Additional comparative sources include: Ba Na Dialect Handbook (Pham, 2019). Vietnamese-Ba Na, Ba Na-Vietnamese Comparative Lexicon (Nguyễn, 2006) Bahnar-Vietnamese Dictionary (Phu et al., 2018). Field recordings collected in Kon Tum and Gia Lai (2021-2023). Recordings were obtained via audio sessions with native speakers (5 Bahnar, 5 Vietnamese males, aged 40s) in 2024, using standardized prompts for elicitation.

Data Analysis

Each lexical entry is annotated for three variables: lexical origin, borrowing strategy, semantic domain. Phonetic transcription used IPA-aligned Bahnar orthography. Acoustic analysis in Praat extracted F0 values (10 points per syllable, z-normalized). OT tableaux modeled patterns (e.g., constraint rankings). Statistical summaries used tables for distributions; interdialectal comparisons via lexical matching.

Results and Analysis

Distribution and Classification of Loanwords

From a 320,000-word Bahnar corpus, 3,663 items were identified as borrowings from Vietnamese, illustrating the extensive integration of external lexicon across a wide range of domains. Following Haugen (1950) three-component model, these borrowings are classified into three main types: Direct loanwords; Loan translations (calques); and Semantic extensions

These Vietnamese borrowings in Bahnar occur in different structural forms, as presented in Table 1.

Table 1: Types and distribution of Vietnamese borrowings in Bahnar

Type of borrowing	Quantity	Percentage (%)
Direct loanwords	2,701	73.7
Loan translations (calques)	532	14.5
Semantic extensions	430	11.7

A majority of Vietnamese borrowings in Bahnar are direct loanwords (2,701 items, 73.7%). These items transfer both the phonological shape and meaning from Vietnamese, often undergoing phonetic modifications

to fit Bahnar phonotactics. Examples include *phủ bộ* → *uiban* and *xoa Sành* → *sapong*. A smaller portion consists of loan translations (calques) (532 items, 14.5%), in which Vietnamese concepts are expressed using native Bahnar elements in descriptive form. For instance, *máy bay* ‘airplane’ is rendered as *may biet bay*. Finally, semantic extensions (430 items, 11.7%) involve native words whose meanings expand under Vietnamese influence. An example is *tơ* (originally ‘only’), which has also come to mean ‘paper’ under the influence of Vietnamese *giấy tơ*.

An examination of the source languages shows that Vietnamese borrowings in Bahnar derive from different strata of Vietnamese vocabulary, including Sino-Vietnamese, native Vietnamese, and French/English mediated through Vietnamese. Table 2 outlines the source language categories of Vietnamese borrowings in Bahnar.

Table 2: Source Language Categories of Vietnamese Borrowings in Bahnar

Source category	Quantity	Percentage (%)
Sino-Vietnamese	2,147	58.6
Native Vietnamese	1,209	33.0
French/English via Vietnamese	307	8.4

A majority of Bahnar borrowings are of Sino-Vietnamese origin (2,147 items, 58.6%), concentrated in formal registers such as politics and administration. Examples include terms for *chính sách* (policy) and *quốc hội* (congress). Native Vietnamese borrowings account for 1,209 items (33.0%) and are more frequent in everyday activities and productive domains, such as *váy* ‘skirt’ and *gấp* ‘fold’. Finally, French and English items mediated through Vietnamese represent 307 borrowings (8.4%), typically referring to modern concepts, cultural imports, or technological innovations, including *In tơ nét* (internet), *bơ* ‘butter’, and *buýt* ‘bus’.

When classified by semantic domain, Vietnamese borrowings in Bahnar also reveal distinct patterns of distribution, with certain fields showing much higher concentrations than others. Table 3 illustrates distribution of Vietnamese borrowings in Bahnar in general semantic domain.

Table 3: Distribution of Vietnamese Borrowings in Bahnar by General Semantic Domain

Domain	Quantity	Percentage (%)
Administration and politics	1,270	34.7
Education and science	982	26.8
Agriculture and livelihood	587	16.0
Health and hygiene	344	9.4
Technology and mass media	257	7.0
Other	223	6.1

Most Vietnamese borrowings in the Bahnar corpus occur in administration and politics (1,270 items, 34.7%), followed by education and science (982 items, 26.8%). Agriculture and livelihood (587 items, 16.0%) and health and hygiene (344 items, 9.4%) also constitute substantial portions. Borrowings in technology, mass media, and other fields together account for less than 15% of the total.

Using the classification of the World Loanword Database (Haspelmath & Tadmor, 2009), loanwords are further distributed by detailed semantic field in the Bahnar corpus as shown in Table 4.

Table 4: Distribution by Detailed Semantic Field in the Bahnar Corpus

Domain	Count
Agriculture and vegetation	27
Animals	14
Body and health	22
Clothing and grooming	19
House and construction	13
Kinship and person	19
Law and politics	10
Motion and location	7
Perception and cognition	7
Social and behavior	12
Tools and technology	15

The largest category relates to agriculture and vegetation (27 items), reflecting the importance of traditional livelihoods. Health-related vocabulary (22 items) and kinship/person terms (19 items) also appear prominently, indicating borrowing in areas of daily life and personal relationships. More abstract categories such as motion, perception, and cognition contain fewer items (7 each), likely because these concepts are already well expressed in native Bahnar or are less salient in daily discourse. The law-politics domain, while institutionally significant, shows a smaller number of borrowings (10 items), possibly due to limited

representation in the corpus.

This suggests borrowings are concentrated in domains tied to institutional authority, education, and mass communication. Moreover, the high proportion of Sino-Vietnamese forms reflects their historical prestige and entrenched role in formal contexts.

Segment Transformation Models

Analysis of Vietnamese-derived borrowings reveals that adaptation strategies are applied systematically, reflecting the dominance of phonotactic constraints in Bahnar. The most recurrent transformations include: (i) Fricative substitution: retroflex /ʂ/ is replaced by alveolar /s/ (e.g., [ʂo] ‘numbers’ → [so] ‘numbers’); (ii) Affricate simplification: retroflex /tʃ/ is realized as alveopalatal /c/ (e.g., [tʃiŋ] ‘school’ → [ciŋ]); (iii) Liquid simplification: palatal nasal /ɲ/ is replaced by lateral /l/ or deleted entirely (e.g., [zaɲ] ‘name’ → [zal] ‘name’), indicating a shift from palatal to simpler coronal articulations. (iv) Coda deletion: final consonants are regularly omitted (e.g., [kuok] ‘nation’ → [kuo]); and (v) Vowel epenthesis: vowels are inserted to break consonant clusters and match Bahnar syllable patterns (e.g., *committee* → *uiban*, *law* → *pápaluat*).

Restructuring Tones in Loanwords

Fundamental frequency (F0) analysis of 210 high-frequency loanwords reveals systematic simplification of Vietnamese tone contours in Bahnar. In general, Vietnamese contour tones are restructured into level or near-level pitch shapes, consistent with the absence of contrastive tones in Bahnar. Recordings from five Bahnar speakers and five Vietnamese speakers were analyzed using Praat, with F0 values extracted at ten equidistant points per syllable and normalized per speaker. The results helped identify five types of tones namely, Hôi (hôi) tone, Nặng (nặng) tone, Sắc (sắc) tone, Huyền (huyền) tone, and Ngang (ngang) tone.

Hôi (hôi) tone

In Bahnar forms, the rising-falling contour of the Vietnamese *hôi* tone is flattened into a low-mid trajectory, starting at approximately 134.8 Hz and ending at 128.6 Hz. In Vietnamese, the tone starts slightly lower (132.6 Hz) but drops more sharply to 117.2 Hz. Bahnar realizations therefore maintain a higher pitch throughout the syllable, with reduced pitch excursion. Figure 1 graphically presents the comparison of *hôi* tone between Bahnar and Vietnamese.

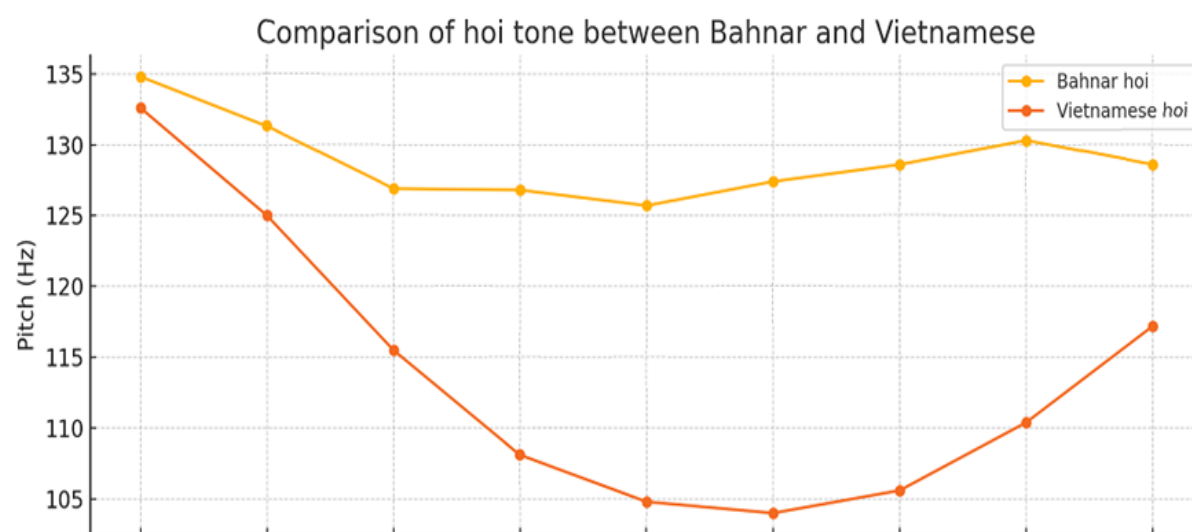


Figure 1: Pitch Contour of "hôi" tone in Vietnamese and Bahnar Loanwords

Nặng (nặng) Tone

Bahnar productions of *nặng* tone begin at about 132.3 Hz and remain relatively stable, ending at 133.0 Hz. In contrast, Vietnamese realizations start much lower (108.2 Hz) and rise slightly to 109.6 Hz. This reanalysis transforms the low register into a level tone, eliminating the marked low-falling contour. Figure 2 graphically presents the comparison of *nặng* tone between Bahnar and Vietnamese.

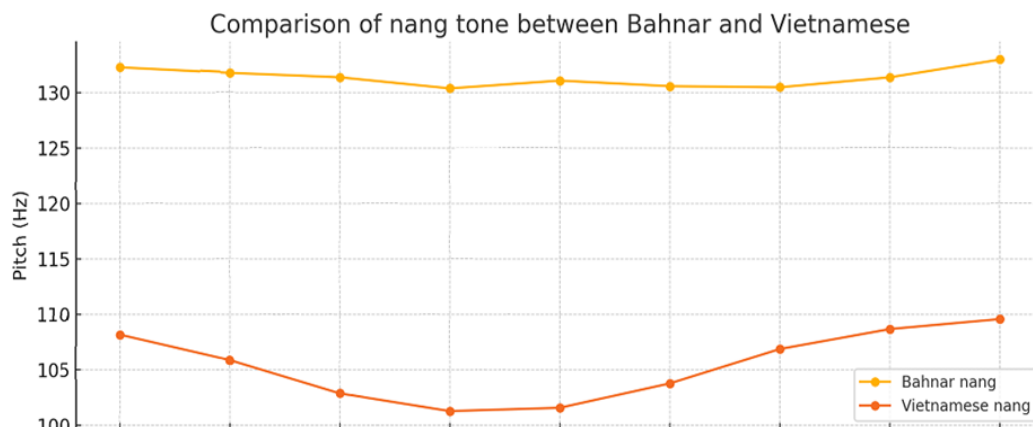


Figure 2: Pitch Contour of "nặng" tone in Vietnamese and Bahnar Loanwords

Sắc (sac) tone

In Bahnar forms, pitch in sac tone starts high (171.5 Hz) and falls slightly to 165.3 Hz, whereas in Vietnamese it starts lower (152.8 Hz) and rises to 168.0 Hz. The original high-rising contour is therefore recast as a high-level tone. Figure 3 graphically presents the comparison of sac tone between Bahnar and Vietnamese.

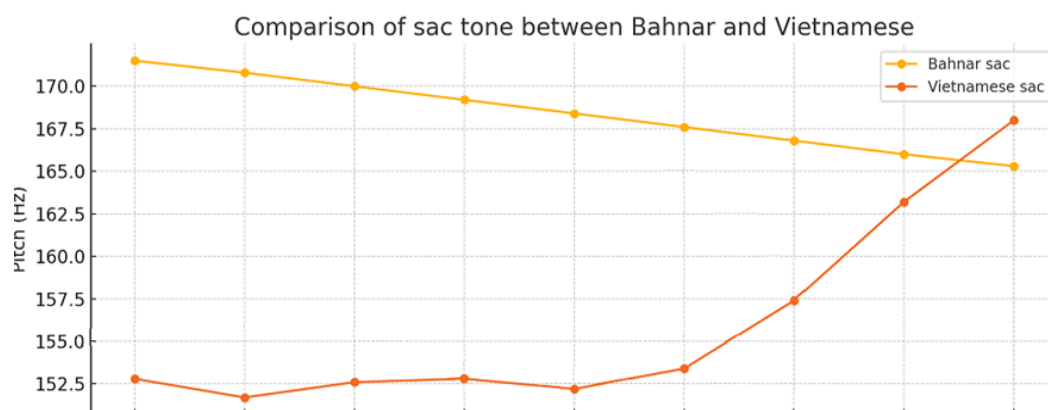


Figure 3: Pitch Contour of "sắc" Tone in Vietnamese and Bahnar Loanwords

Huyền (huyen) tone

Both languages exhibit a falling pitch in terms of huyen tone, but Bahnar realizations fall more gradually, from 146.2 Hz to 131.1 Hz, ending slightly higher than Vietnamese (147.5 Hz to 127.6 Hz). Figure 4 graphically presents the comparison of huyen tone between Bahnar and Vietnamese.

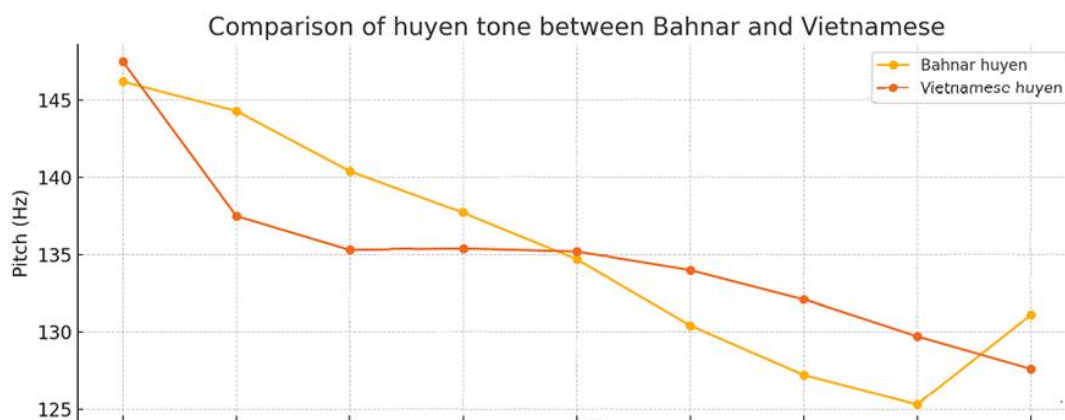


Figure 4: Pitch Contour of "huyền" Tone in Vietnamese and Bahnar Loanwords

Ngang (ngang) Tone

The ngang tone in Bahnar starts at 165.2 Hz (Fig. 5) and falls to 143.5 Hz, while Vietnamese begins at 146.9 Hz and shows a slight rise to 139.0 Hz by the end. The Bahnar version thus has a higher onset and a

gentle falling pattern. Figure 5 graphically presents the comparison of ngang tone between Bahnar and Vietnamese

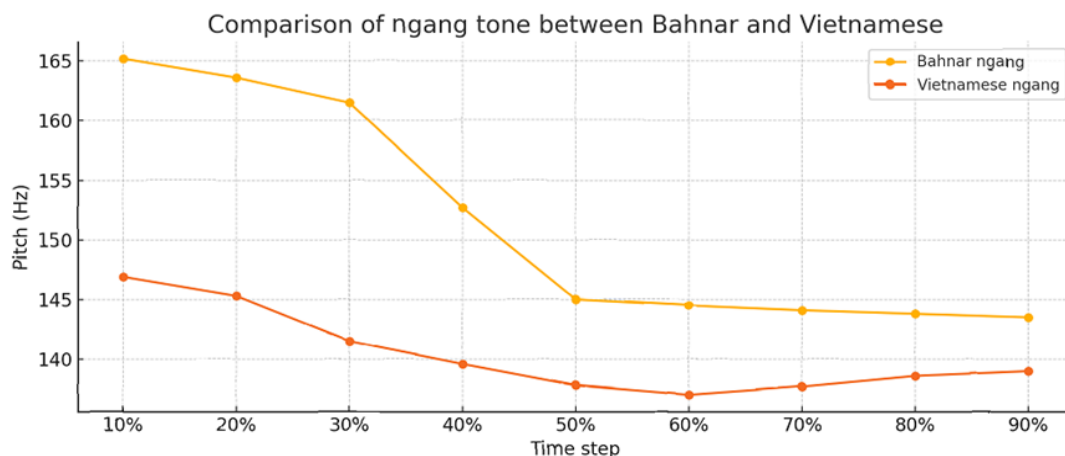


Figure 5: Pitch Contour of "ngang" tone in Vietnamese and Bahnar Loanwords

Among the 285 loanwords analyzed, 27 items (9.5%) exhibited reduplication. Two main types were attested. Total reduplication involved repeating the entire form without alteration, as in *bô bô* (*talkative*) and *xanh xanh* (*light blue*). Partial reduplication copied only a portion of the word, usually a consonant or a rhyme, as in *chập chồn* (*flicker*), *lập cập* (*trembling*). These reduplicated borrowings are most frequently found in affective speech, sound symbolism, and adverbial or action-descriptive contexts, showing that reduplication is a recurrent strategy in the Bahnar adaptation of Vietnamese items.

Interdialectal Comparison

To evaluate the degree of homogeneity and variation in loanword adaptation among Bahnar dialects, a cross-dialectal comparison was also conducted between (i) a large corpus collected in Binh Dinh province, and (ii) the lexical database of the Bahnar Kon Tum dialect. This comparison enabled the identification of borrowing patterns that are stable across the region, as well as the detection of local linguistic innovations. Table 5 exhibits lexical form comparison across three Bahnar dialects viz., Binh Dinh, Gia Lai, and Kon Tum.

Table 5: Lexical form Comparison Across Three Bahnar Dialects (Binh Dinh, Gia Lai, Kon Tum)

Criterion	Number	Percentage (%)
Total number of compared lexical items	285	100
Identical (or nearly identical) in all dialects	172	60.4
Found only in Binh Dinh	51	17.9
Found only in Gia Lai	33	11.6
Found only in Kon Tum	29	10.2

The cross-dialectal comparison of 285 loanwords across Bahnar varieties showed that 172 items (60.4%) were identical or nearly identical in all three dialects (Binh Dinh, Gia Lai, Kon Tum). The remaining items exhibited regional distribution: 51 words (17.9%) were found only in Binh Dinh, 33 words (11.6%) only in Gia Lai, and 29 words (10.2%) only in Kon Tum.

Examples of region-specific borrowings in Binh Dinh include *abẽn* 'skirt' ← Vietnamese *váy*, *adon* 'stack' ← *xếp*, and *bươ bươ* 'side by side, simultaneously', a reduplicated form not attested elsewhere. These cases illustrate that while the majority of Vietnamese borrowings are shared across Bahnar dialects, certain items remain localized, reflecting uneven diffusion of lexical innovations. The high frequency of loanwords exclusively in the Binh Dinh dataset occurred potentially due to scenarios like greater institutional exposure (e.g., school-based Vietnamese, regional radio broadcasts); higher levels of bilingualism and Vietnamese fluency; and integration of recent or colloquial borrowings. Examples include: *abẽn* 'skirt' ← Vietnamese *váy*; *adon* 'stack' ← Vietnamese *xếp*; and *bươ bươ* 'side by side, simultaneously' ← expressive reduplication unique to Binh Dinh. Overall, most loanwords were shared across dialects, though each region also retained a subset of unique borrowings.

Discussion

The findings of this study provide robust empirical support for the hypothesis that Vietnamese loanwords in Bahnar undergo a systematic process of adaptation shaped by native phonotactic constraints and sociolinguistic factors. This section interprets the main results in light of Optimality Theory (OT), language contact typology, and dialectal phonological variation, and explores their implications for minority language vitality, identity negotiation, and structural convergence.

Constraint Re-ranking in the OT Grammar of Bahnar

The phonemic substitutions documented in Bahnar loanwords such as /ʁ/ → /s/, /ɰ/ → /c/, and regular deletion of coda consonants can be accounted for within the framework of Optimality Theory. These patterns reflect a constraint ranking in which markedness constraints (*CODA, *COMPLEX, *RETROFLEX) consistently outrank faithfulness constraints (e.g., IDENT-IO[continuant], IDENT-IO[place], MAX-IO). This ranking aligns with patterns widely attested in Mon-Khmer languages, where the canonical CV syllable is preferred, complex consonant clusters are avoided, and articulatorily demanding retroflex segments are eliminated.

The dominance of markedness constraints indicates that Bahnar phonology prioritizes structural simplicity and perceptual clarity. Crucially, these strategies recur across hundreds of loanword tokens, showing that they are not isolated or sporadic but rather manifestations of a stable phonological system shared among speakers. In reshaping non-native inputs, Bahnar speakers ensure that borrowed words conform to the permissible syllable structures of their language, while still retaining enough phonetic substance to remain recognizable. These findings underscore the explanatory power of OT in modeling loanword adaptation, consistent with observations in other contact settings (LaCharité & Paradis, 2005; Trang, 2021).

Tone Restructuring and Functional Reduction

The neutralization of Vietnamese contour tones into level or near-level tones in Bahnar loanwords reflects a process of systematic phonologization rather than imperfect imitation. Acoustic results showed that rising–falling shapes such as *hỏi* and *ngã* were realized as mid-level trajectories, while *sắc* was reinterpreted as a high-level pitch and *nặng* as a mid-level tone with minimal excursion. This simplification can be modeled within OT as *CONTOUR* » *IDENT-TONE*, whereby complex tonal contours are disfavored and simplified categories are preferred (Yip, 2002).

Typologically, Bahnar exhibits a low functional load for tone, relying instead on register distinctions and segmental contrasts. As a result, tonal contrasts that are not essential for lexical or grammatical differentiation are reduced. This process aligns with the principle of cognitive economy: lexical recognition is preserved while eliminating unnecessary acoustic complexity. Similar tendencies have been reported in other Austroasiatic languages, such as Kammu Mahdi (2024), suggesting that contour simplification is a broader areal and genetic feature.

Morpheme–Phoneme Integration Via Reduplication

Reduplication emerges as a productive mechanism for integrating loanwords into Bahnar's morphophonological system. In the present dataset, 9.5% of Vietnamese borrowings exhibit total or partial reduplication, demonstrating that speakers actively restructure foreign lexical items to align with native prosodic and semantic patterns. Total reduplication (e.g., *ùn ùn* (one after another), *đăm đăm* (staring)) conveys emphasis or continuity, while partial reduplication (e.g. *gay gắt* (harsh), *êm đềm* (calm), *ê chề* (miserable)) frequently occurs in affective or ideophonic contexts.

This finding is consistent with classic descriptions of Bahnar reduplication. Banker (1964) documented multiple types, including formations that express contrary-to-expectation meanings, emotive nuances such as anger or displeasure, sequences of consecutive actions, and intensification. These patterns reveal that reduplication is deeply embedded in Bahnar grammar, not a marginal or ad hoc process. Austroasiatic languages more broadly show similar tendencies toward reduplicative morphology (Diffloth, 1976), suggesting that Bahnar's adaptation of Vietnamese loans reflects long-standing areal and genealogical features.

Functionally, reduplication in loanwords serves several roles including (i) Phonological assimilation by adjusting syllable structure and rhythm to Bahnar's prosodic templates; (ii) Morphological marking of plurality, aspect, or intensity; and (iii) Sociolinguistic marking by embedding borrowed items into stylistic registers familiar to Bahnar speakers. Taken together, these processes show that reduplication is not simply a stylistic ornament but a structural strategy of domestication, ensuring that borrowed forms become fully integrated within Bahnar's morphophonological and sociocultural system.

Social Motivation and Lexical Stratification

Loanword distribution across semantic domains reflects patterns of social prestige and institutional influence. Vietnamese-origin terms dominate in administration, education, and healthcare, which are domains where Vietnamese holds official status. In contrast, core vocabulary in domains such as kinship and ritual exhibits high resistance to borrowing, illustrating a diglossic distribution in which Bahnar fulfills everyday intimate communicative functions, while Vietnamese occupies formal, technical, and institutional spheres.

Bahnar speakers are not passive recipients. They selectively reshape borrowed forms, actively preserving linguistic identity. This aligns with theories of language ideology and identity performance (Eckert, 2000;

Scholvin, 2022), where speakers regulate assimilation as an expression of group belonging. As noted by Trang (2021) and Nguyen (2010), language contact in Vietnam is closely linked to educational expansion and globalization. Loanword use may therefore serve not only as a sign of external influence but also as a means of enriching the lexicon for stylistic, indexical, and register-marking purposes.

Dialectal Variation and Lexical Diffusion

The comparison between the Bahnar dialects of Binh Định and Kon Tum shows both convergence and divergence. Shared loanwords, which account for 60.4% of the corpus, reflect a common phonological constraint architecture, while dialect-specific forms in Binh Định (e.g., *abên* 'skirt', *adon* 'stack') point to localized innovations and uneven contact exposure. This pattern supports a gradient model of lexical diffusion (Labov, 1994; William, 1969), in which high-frequency and socially salient loanwords diffuse widely, whereas low-frequency or context-specific items remain localized. From a theoretical perspective, the findings suggest that OT constraint rankings in contact settings are not entirely fixed but may vary across dialects under the influence of social and geographic factors (LaCharité & Paradis, 2005; Trang, 2021). Comparative studies of other Bahnaric varieties (Person, 2022; Sidwell, 2024) would further clarify the dynamics of convergence and divergence.

Conclusion

This study provides a comprehensive analysis of Vietnamese-origin loanwords in the Bahnar dialect of Binh Dinh, drawing on a dataset of 5,769 entries. By combining semantic field classification following the WOLD framework with source language identification, the research offers new insights into the nature of lexical borrowing in a Southeast Asian minority language under long-term contact with Vietnamese. First, the semantic distribution of loanwords reveals a marked concentration in domains such as public administration, technology, education, and sanitation, reflecting areas of sustained socio-cultural and institutional interaction. The consistency of these patterns across the two surveyed areas, Binh Dinh and Kon Tum, indicates a systematic borrowing process shaped by national policies, media exposure, and formal education.

Second, analysis of source languages shows that the majority of borrowings are of non-standardized or uncertain origin, most likely from local varieties of Vietnamese, alongside a substantial proportion of Sino-Vietnamese forms and a smaller number from French and English. This distribution mirrors the historical stratification of lexical influence in Vietnamese described by Alves (2006), and the intermediary role of Vietnamese in transmitting foreign-origin vocabulary into Bahnar.

Finally, applying Haugen's three-component model in combination with the sociolinguistic frameworks of Thomason and Kaufman (2023) and Winford (2003), the study finds that borrowing in Bahnar is predominantly lexical and phonological, with minimal evidence of morphosyntactic transfer. The adaptation of loanwords to Bahnar phonotactic constraints points to a controlled and selective integration process, consistent with patterns observed in other languages of Mainland Southeast Asia. Beyond its contribution to the description of Bahnar, this research proposes an analytical framework that can be replicated to examine contact-induced phonological adaptation and lexical diffusion in other minority language contexts. These findings would also advance OT applications in contact linguistics, informing language policy for minority preservation (e.g., bilingual curricula) and typology studies in Southeast Asia.

This study faced several limitations as well. First, it draws data primarily on elicited and broadcast corpora without incorporating naturalistic conversational data, which might reveal emergent trends. Second, the analysis focuses on phonological and lexical adaptation, leaving pragmatic and syntactic integration for future investigation. Third, the dataset is skewed toward adult speakers in formal contexts, while younger speakers and informal media such as social networks may display distinct borrowing patterns. Last, but not the least, the dialectal comparison is limited to Binh Dinh and Kon Tum, excluding other Bahnar-speaking regions. Addressing these gaps will provide a more comprehensive picture of loanword dynamics in Bahnar. The future research can incorporate naturalistic data from younger speakers and informal media; expand to pragmatic/syntactic integration and additional dialects.

Acknowledgement.

This research is funded by Vietnam Ministry of Science and Technology under the Program "Supporting Research, Development, and Technology Application of Industry 4.0" (KC-4.0/19-25), via the project "Development of a Vietnamese-Bahnaric Machine Translation and Bahnaric Text-to-Speech System (All Dialects)" (KC-4.0-29/19-25).

Authors' contribution.

Each of the first three authors (Tam T. T. Le, Quy T. Nguyen, Long S. T. Nguyen) contributed equally

and are co-first authors of this study.

References

- Abbasi, K., & Khanam, T. (2025). Optimality Theory and Phonology of English Loanwords in Urdu. *Journal of Arts and Linguistics Studies*, 3(3), 5071-5084. <https://doi.org/10.71281/jals.v3i3.447>
- Abdulrazzaq, A. H., & Al-Ubaidy, S. M. A. (2023). Models of phonological loanword adaptation: The optimality model as opposed to the perceptual and phonological models. *Al-Adab Journal*(144), 17-24. <https://doi.org/10.31973/aj.v3i144.3881>
- Alhoody, M. M. A. (2019). *Phonological adaptation of English loanwords into Qassimi Arabic: an optimality-theoretic account* [Newcastle University]. <http://theses.ncl.ac.uk/jspui/handle/10443/4662>
- Allen, D. (2020). An overview and synthesis of research on English loanwords in Japanese. *Vocabulary Learning and Instruction*, 9(1), 33-50. <https://doi.org/10.7820/vli.v09.1.Allen.a>
- Alves, M. (2006). Linguistic research on the origins of the Vietnamese language: An overview. *Journal of Vietnamese Studies*, 1(1-2), 104-130. <https://doi.org/10.1525/vs.2006.1.1-2.104>
- Alves, M. (2024). An Updated Overview of the Austroasiatic Components of Vietnamese. *Languages*, 9(12), 377. <https://doi.org/10.3390/languages9120377>
- Banker, E. M. (1964). Bahnar reduplication. *Mon-Khmer Studies*, 1, 119-134. <http://www.icaal.org/archives/mks/pdf/1:119-134.pdf>
- Brunelle, M. (2009). Tonal Coarticulation in Northern Vietnamese. https://www.internationalphoneticassociation.org/icphs-proceedings/ICPhS2003/papers/p15_2673.pdf
- Bùi, M. D., Trần, H. T., Bùi, B. L., & học, V. d. t. (2006). *Dân tộc Ba Na ở Việt Nam*. Nhà xuất bản Khoa học xã hội. <https://books.google.com.pk/books?id=ml1QAQAAMAAJ>
- Cao, V. (2022). Lexical-phonological patterns of Tai loans in Vietnamese: diagnoses and implications. <https://www.academia.edu/103298641>
- Coto-Solano, R., & Trần, D. (2021). Priming Effects of Tones in Visual Processing of Vietnamese. Proc. TAI 2021, https://www.isca-archive.org/tai_2021/cotosolano21_tai.pdf
- Diffloth, G. (1976). Expressives in Semai. *Oceanic linguistics special publications*(13), 249-264. <https://www.jstor.org/stable/20019159>
- Eckert, P. (2000). Linguistic Variation as Social Practice: The Linguistic Construction of Identity in Belten High. <https://www.jstor.org/stable/3086949>
- Ferlus, M. (1999). Les dishannonies tonales en viet-muong et leurs implications historiques. *Cahiers de linguistique Asie orientale*, 28(1), 83-100. <https://doi.org/10.1163/19606028-90000066>
- Ha, P. (2022). Code mixing and loan words in the Vietnamese vocabulary. *Eurasian Journal of Applied Linguistics*, 8(1), 143-148. https://eric.ed.gov/?q=vietnamese&ff1=dtvSince_2018&id=EJ1474680
- Haspelmath, M., & Tadmor, U. (2009). *Loanwords in the World's Languages: A Comparative Handbook*. De Gruyter. <https://books.google.com.pk/books?id=HnKeVbwTwYc>
- Haugen, E. (1950). The analysis of linguistic borrowing. *Language*, 26(2), 210-231. <https://doi.org/10.2307/410058>
- Ilmina, H. (2016). *Borrowed words in "World Student" magazine* Universitas Negeri Islam Maulana Malik Ibrahim]. <http://etheses.uin-malang.ac.id/id/eprint/5212>
- Labov, W. (1994). Principles of linguistic change: Internal factors. In: Blackwell. <https://www.wiley.com/en-us/Principles+of+Linguistic+Change%2C+Volume+1%3A+Internal+Factors-p-9780631179146>
- LaCharité, D., & Paradis, C. (2005). Category preservation and proximity versus phonetic approximation in loanword adaptation. *Linguistic inquiry*, 36(2), 223-258. <https://doi.org/10.1162/0024389053710666>
- Mahdi, W. (2024). Austroasiatic loanwords in Austronesian languages. *Journal of the Humanities of Indonesia*, 25(3). <https://doi.org/10.17510/wacana.v25i3.1780>
- Mai, N. C., & Le Phan, T. N. (2025). Language Endangerment And Endangered Languages In Vietnam. *Tạp Chí Khoa Học Ngôn Ngữ Và Văn Hóa*, 9(1), 63-79. <https://doi.org/10.63506/jilc.0901.305>
- Naika, O. J. (2021). *A morphophonological analysis of Oluwanga loanwords: an optimality theory account* [Kenyatta University]. <https://ir-library.ku.ac.ke/server/api/core/bitstreams/7f2c8dca-dd8e-4abb-af02-d7da7302e493/content>
- Nam, U. V., & Duc, N. M. (2021). Population ageing and older persons in Viet Nam. *General Statistics Office*. https://vietnam.unfpa.org/sites/default/files/pub-pdf/ageing_report_from_census_2019_eng_final20082021.pdf
- Nguyễn, N. S. (2003). *Tìm hiểu tiếng Việt lịch sử*. Nhà xuất bản Đại học sư phạm. <https://books.google.com.pk/books?id=pHtkAAAAAMAAJ>
- Nguyen, T. G. (2010). *Từ vựng học tiếng Việt [Vietnamese lexicology]*. NXB Giáo dục Việt Nam [Vietnam Education Publishing House]. <https://www.scribd.com/document/671988521>
- Nguyễn, T. M. H., Romary, L., Rossignol, M., & Vũ, X. L. (2006). A lexicon for Vietnamese language processing. *Language Resources and Evaluation*, 40(3), 291-309. <https://doi.org/10.1007/s10579-007-9034-8>
- Person, S. L. (2022). *Li Xei: A phonology of an understudied Bahnaric language in central Vietnam*. The

- University of North Dakota. <https://commons.und.edu/theses/4288>
- Pham, A. H. (2019). Vietnamese dialects. *Interdisciplinary Perspectives on Vietnamese Linguistics*, 31. <https://www.torrossa.com/en/resources/an/5000491#page=40>
- Phong, L. H. (2005). Từ vựng đối chiếu Việt – Ba Na, Ba Na – Việt [Vietnamese–Bahnar, Bahnar–Vietnamese comparative vocabulary]. NXB Văn hóa Dân tộc [Ethnic Culture Publishing House]. <https://cir.nii.ac.jp/crid/1971712334760286080?lang=en>
- Phu, V. N., Chau, V. T. N., Tran, V. T. N., & Dat, N. D. (2018). A Vietnamese adjective emotion dictionary based on exploitation of Vietnamese language characteristics. *Artificial Intelligence Review*, 50(1), 93-159. <https://doi.org/10.1007/s10462-017-9538-6>
- Prince, A., & Smolensky, P. (2004). Optimality theory: Constraint interaction in generative grammar. *University, New Brunswick, and University of Colorado*. <https://doi.org/10.1002/9780470759400>
- Scholvin, V. (2022). Syllabification in Language Contact between French and Vietnamese. <https://www.researchgate.net/publication/369596712>
- Scholvin, V., & Meinschaefer, J. (2018). The integration of French loanwords into Vietnamese: A corpus-based analysis of tonal, syllabic and segmental aspects. *Journal of the Southeast Asian Linguistics Society*(Special Publication 3), 157-173. <https://www.researchgate.net/publication/342589329>
- Shafi, S. (2017). *A phonological analysis of English loanwords in Mirpur Pahari: Exploring variable adaptation in Optimality Theory* University of York]. https://etheses.whiterose.ac.uk/id/oai_id/oai:etheses.whiterose.ac.uk:24292
- Sidwell, P. (2024). 500 Proto Austroasiatic Etyma: Version 1.0. *Journal of the Southeast Asian Linguistics Society*, 17(1). <http://hdl.handle.net/10524/52519>
- Trang, N. H. (2021). Segmental Adaptation of French Loans in the Vietnamese Language: The Use of Optimality-Theory. *Journal of Universal Language*, 22(2), 105-143. <https://doi.org/10.22425/jul.2021.22.2.105>
- Thomas, D. (1966). Mon-Khmer subgroupings in Vietnam. *Studies in comparative Austroasiatic linguistics*, 194-202. <http://sealang.net/sala/archives/pdf8/thomas1966mon.pdf>
- Thompson, L. C. (1963). The problem of the word in Vietnamese. *Word*, 19(1), 39-52. <https://doi.org/10.1080/00437956.1963.11659787>
- Thomason, S. G., & Kaufman, T. (2023). *Language Contact, Creolization, and Genetic Linguistics*. University of California Press. https://books.google.com.pk/books?id=d0_hEAAAQBAJ
- Van Van, H. (2020). The roles and status of English in present-day Vietnam: A socio-cultural analysis. *VNU Journal of Foreign Studies*, 36(1). <https://doi.org/10.25073/2525-2445/vnufs.4495>
- Verdonschot, R. G., Phuong, H. T. L., & Tamaoka, K. (2022). Phonological encoding in Vietnamese: An experimental investigation. *Quarterly Journal of Experimental Psychology*, 75(7), 1355-1366. <https://doi.org/10.1177/17470218211053244>
- William, S.-Y. W. (1969). Competing changes as a cause of residue. *Language*, 45(1), 9. <https://doi.org/10.2307/411748>
- Winford, D. (2003). *An introduction to contact linguistics*. Wiley-Blackwell. <https://www.researchgate.net/publication/273003007>
- Yip, M. (2002). *Tone*. Cambridge University Press. <https://doi.org/10.1017/CBO9781139164559>
- Yule, G. (2006). *The study of language* (3rd ed.). Cambridge University Press. <https://doi.org/10.1017/CBO9780511819742>