

MORPHOLOGICAL NATURALIZATION OF BORROWED LEXEMES IN URDU: A CORPUS-BASED STUDY

Afshan Quyyum

Lecturer in Linguistics, Department of English

University of Okara, Pakistan, PhD Scholar in Linguistics, Minhaj University Lahore

*Emails: afshanacademia@gmail.com
2022f-mulphd-eng-003@mul.edu.pk*

ABSTRACT

This study aims to examine the morphological naturalization of borrowed English lexemes in contemporary Urdu and to identify whether this process constitutes a distinct pattern of word formation. Using a corpus-based and corpus-driven qualitative approach, data were collected from naturally occurring sources including public signage, educational texts, commercial labels, and everyday written discourse. The analysis focuses on how English-origin nouns such as ٹیبل، بک، اسکول، سسٹم، پنسل undergo systematic pluralization and grammatical integration through Urdu morphological patterns (e.g., ٹیبلیں، بکیں، اسکولوں). The findings reveal that these forms are no longer used as temporary borrowings or instances of code-mixing; rather, they function as fully integrated Urdu lexemes. The study argues that this process represents a distinct type of word formation—termed *morphological naturalization*—in which borrowed lexemes become structurally and cognitively embedded in the host language. This research is significant as it demonstrates that Urdu expands its lexicon not merely through borrowing, but through active grammatical restructuring, highlighting a dynamic and productive mechanism of lexical growth. The study is limited to written and publicly visible data and recommends further research using spoken corpora and larger datasets to extend the proposed framework.

Keywords: Morphological Naturalization (MNL), Lexical Borrowing, Word Formation, Urdu Morphology, Corpus Linguistics, Lexical Integration, Language Contact

Note: *Morphological Naturalization of Borrowed Lexemes (MNL) is proposed in this study as a new word-formation process in Urdu.*

INTRODUCTION

This study investigates how English-origin lexical items are morphologically integrated into Urdu and examines whether this integration constitutes a distinct word-formation process. The primary objective is to identify and analyze the systematic pluralization and grammatical behavior of borrowed lexemes such as ٹیبل، بک، اسکول، سسٹم، پنسل, and to determine their linguistic status within the contemporary Urdu lexicon. The study adopts a corpus-based and corpus-driven qualitative approach, drawing data from naturally occurring language used in public signage, educational materials, commercial domains, and everyday written discourse. A small specialized corpus was constructed through purposive sampling, and the data were analyzed morphologically to observe patterns of pluralization, compounding, and grammatical integration. The analysis follows an inductive framework, allowing linguistic patterns to emerge from the data rather than being imposed through pre-existing theoretical assumptions.

The study is significant in that it identifies morphological naturalization as a previously underexplored word-formation process in Urdu. While earlier studies have described borrowing as a common outcome of language contact, they have rarely examined the stage at which

borrowed lexemes become structurally embedded within the grammatical system of the recipient language. The findings of this study demonstrate that English-origin nouns in Urdu are not merely borrowed or code-mixed but undergo systematic grammatical adaptation, functioning as fully integrated lexical items.

This research is limited to:

- English-origin nouns,
- Written and publicly visible linguistic data,
- Urban Pakistani contexts.

Spoken discourse, regional dialectal variation, and diachronic development fall outside the scope of the present study and are recommended for future research.

1.1 Theoretical Orientation and Research Motivation

The present study follows an inductive, corpus-driven approach. Rather than testing a pre-established model of word formation, the analysis begins with naturally occurring linguistic data collected from public, educational, and commercial domains. Through systematic categorization and qualitative analysis, the study identifies a recurring morphological pattern in which borrowed English lexemes undergo pluralization, compounding, and syntactic integration within Urdu. This recurring pattern leads to the formulation of a new word-formation process, termed Morphological Naturalization, which has not been explicitly documented in previous studies of Urdu morphology.

Previous scholarship on language contact has primarily examined borrowing at the lexical or sociolinguistic level (Haugen, 1950; Thomason & Kaufman, 1988). However, limited attention has been given to the stage at which borrowed items undergo grammatical restructuring and begin to behave like native lexical units. In contemporary Urdu, English nouns such as table, school, system, and bottle appear not merely as loanwords but as morphologically productive forms that take Urdu plural markers (e.g., **بُوتنیں**, **اسکولوں**, **سیسمون**, **بُوتلیں**) and participate in native syntactic constructions. This shift reflects a deeper level of integration than code-switching or borrowing alone.

Drawing on insights from lexical morphology and usage-based linguistics (Bybee, 2010; Booij, 2012), this study argues that such forms represent a process of morphological naturalization, whereby borrowed lexemes are reshaped according to the grammatical rules of Urdu and become cognitively processed as part of its lexicon. This process reflects not linguistic erosion but lexical expansion driven by communicative necessity, frequency of use, and structural compatibility. By documenting this underexplored stage of lexical development, the study contributes to Urdu lexicography and provides empirical evidence that contact-induced change can result in long-term grammatical integration rather than temporary lexical borrowing.

1.2 Statement of the Problem

Although English lexical borrowing in Urdu has been widely acknowledged, existing studies largely treat such borrowings as surface-level insertions or instances of code-mixing. Little attention has been paid to how borrowed English nouns undergo morphological integration and become grammatically productive within Urdu. Consequently, the process through which borrowed lexemes acquire Urdu plural morphology and function as native nouns remains insufficiently explored. This study addresses this gap by examining morphological naturalization as a distinct and systematic word-formation process in contemporary Urdu.

1.3 Research Questions

1. How do English-origin lexical items undergo morphological naturalization in contemporary Urdu?

- What grammatical patterns indicate that borrowed lexemes have become fully integrated into Urdu morphology?

1.4 Research Objectives

- To examine the morphological behavior of English-origin nouns in Urdu, particularly with reference to pluralization and grammatical adaptation.
- To identify recurrent patterns of morphological naturalization using corpus-based evidence drawn from naturally occurring language data.

1.5 Research Gap

Previous studies on lexical borrowing in Urdu have primarily focused on *lexical adoption, code-switching, or sociolinguistic influence*, while paying limited attention to the *morphological stage of integration*. Existing literature rarely examines how borrowed English nouns undergo *systematic grammatical restructuring*, particularly through pluralization and compounding, to function as native-like lexical items in Urdu.

Moreover, most prior research relies on introspective or literary data, whereas *corpus-based evidence from public signage, educational texts, and everyday written discourse remains largely unexplored*. The absence of such data has resulted in an incomplete understanding of how borrowed forms stabilize and become part of the Urdu lexicon.

This study addresses this gap by analyzing naturally occurring corpus data and proposing *morphological naturalization* as a distinct and previously under-theorized stage in Urdu word formation.

2. LITERATURE REVIEW

The phenomenon of lexical borrowing has long been recognized as a natural outcome of language contact. Early work by Haugen (1950) conceptualized borrowing as a continuum ranging from simple lexical adoption to deeper structural integration. Subsequent scholars have emphasized that borrowing is not a superficial process but one that interacts with the morphological and grammatical systems of the recipient language (Thomason & Kaufman, 1988; Myers-Scotton, 2002). Within this framework, borrowed items may undergo phonological adjustment, semantic narrowing, and, in some cases, full morphological assimilation.

In South Asian linguistic contexts, borrowing from English has been particularly pervasive due to prolonged colonial contact, institutional dominance, and globalization (Kachru, 1986; Rahman, 2011). Urdu, in particular, has demonstrated a high degree of receptivity to English lexical items, especially in domains such as education, administration, technology, and urban life. However, most existing studies treat such borrowings primarily as instances of code-switching or lexical mixing (Baumgardner, 1993; Rahman, 2004), with limited attention to their grammatical behavior once absorbed into everyday usage.

Recent scholarship has begun to recognize that borrowed words do not remain static. Matras (2009) argues that once a borrowed lexeme becomes frequent in everyday interaction, it may undergo structural accommodation, allowing it to behave like a native item. This view aligns with the concept of **morphological naturalization**, whereby foreign lexical items adopt the inflectional and derivational rules of the host language. In Urdu, this is visible in the attachment of native plural markers such as –یں, –وں, and –ات to English-origin nouns (e.g., بونیں, پسلیں, سسٹمون). Such forms indicate not mere borrowing but grammatical incorporation.

Despite this, existing research on Urdu largely treats pluralized English nouns as peripheral or stylistic variants rather than as evidence of systematic word-formation processes. Studies on Urdu morphology (Platts, 1904; Schmidt, 1999) describe pluralization patterns

extensively but do not account for how foreign lexemes enter and participate in these paradigms. Similarly, research on code-mixing in Pakistani English and Urdu (Rasul, 2006; Mahboob, 2009) focuses on sociolinguistic motivation rather than morphological integration.

A notable gap therefore exists in the literature: *the absence of corpus-based studies that document how borrowed English nouns become morphologically productive in Urdu*. While lexical borrowing has been widely acknowledged, the stage at which borrowed items undergo pluralization, compounding, and syntactic embedding remains under-theorized. As Backus (2014) notes, this stage represents a shift from borrowing to grammatical entrenchment, where borrowed forms begin to function as native lexical units.

The present study addresses this gap by examining English-origin nouns that have entered Urdu and undergone morphological naturalization in real-world usage. Drawing on corpus data collected from public signage, educational materials, and everyday written discourse, the study moves beyond surface-level borrowing to analyze how these forms participate in Urdu plural formation and lexical expansion. In doing so, it builds on earlier work by Quyyum (2023) on lexical adaptation in Punjabi and extends the discussion to Urdu, demonstrating that morphological integration is not an exception but an emerging norm in contemporary language use.

By situating borrowed lexemes within a morphological and corpus-based framework, the study contributes to ongoing debates in contact linguistics and lexical semantics, arguing that lexical change in Urdu is not random or degradative but structurally motivated, socially grounded, and linguistically productive.

3. RESEARCH METHODOLOGY

The methodological framework of this study is designed to systematically investigate patterns of lexical change as they emerge in real-life language use. Since the focus of the research is on naturally occurring lexical phenomena rather than experimentally controlled data, a qualitative and usage-oriented methodological approach has been adopted.

3.1. Research Design

The present study adopts a qualitative, corpus-based, and corpus-driven research design to investigate patterns of lexical change as they emerge in naturally occurring language. Since the focus of the research is on real-life linguistic behavior rather than experimentally elicited data, a usage-oriented methodological approach has been employed. This design allows for the observation of language as it is actually used in social and institutional contexts, rather than as it is prescribed in normative grammars.

The study is situated within the domains of lexical semantics, morphology, and lexicography, with particular emphasis on how borrowed English lexemes undergo structural and grammatical adaptation in contemporary Urdu. Rather than treating borrowing as a marginal or peripheral phenomenon, the research views lexical change as an inherent and systematic process shaped by frequency of use, communicative necessity, and sociocultural context (Bybee, 2010; Thomason, 2001).

The research follows a corpus-driven orientation, in which analytical categories are not imposed in advance but emerge inductively from the data itself. This approach is consistent with usage-based and empirical models of linguistic analysis, which argue that grammatical patterns evolve from repeated language use rather than from abstract rules alone (Bybee, 2010; Biber et al., 1998). In this sense, the study aligns methodologically with principles associated with

grounded theory, where patterns, categories, and theoretical insights are derived from systematic observation of data rather than predetermined hypotheses (Glaser & Strauss, 1967).

The study is also descriptive and interpretive in nature, aiming to document and explain linguistic change rather than evaluate it normatively. It examines how English-origin lexical items are absorbed into Urdu through transliteration, pluralization, and morphological restructuring, particularly in domains such as education, commerce, public signage, and everyday written communication.

By combining corpus-based evidence with lexical-semantic analysis, the study provides an empirical account of how Urdu expands its lexicon through contact-induced change. This methodological framework enables the identification of morphological naturalization as a distinct and previously underexplored word-formation process in Urdu.

3.2 Data Collection

Data were collected through purposive sampling from naturally occurring written discourse. Sources included shop signs, commercial boards, institutional labels, textbooks, advertisements, and other publicly visible texts from multiple urban locations in Pakistan. These sources were selected because they represent the most active and socially influential sites of lexical innovation and language contact.

3.3 Corpus Construction

A small specialized corpus was constructed by systematically compiling English-origin lexical items written in Urdu script. The corpus was developed with attention to:

- authenticity of usage
- frequency of occurrence
- functional relevance in everyday communication
- representation across domains (education, commerce, public space)

This corpus serves as an empirical foundation for identifying recurring morphological patterns and observing the grammatical behavior of borrowed lexemes in real contexts.

3.4 Analytical Framework

The analysis is grounded in a lexical-semantic and morphological framework, focusing on how borrowed items undergo structural integration in Urdu. Particular attention is given to:

- plural formation (بن / ون / ات -)
- compounding with Urdu or English elements
- syntactic behavior of borrowed nouns
- semantic stabilization in recurrent usage

The study adopts a corpus-driven analytical strategy, allowing patterns of morphological adaptation to emerge inductively from the data. Through this process, the research identifies morphological naturalization as a distinct word-formation mechanism whereby borrowed lexemes cease to function as foreign items and instead become grammatically productive units of Urdu.

4. DATA PRESENTATION AND ANALYSIS

This section presents and analyzes the empirical data collected for the study with the aim of identifying emerging patterns of lexical change in contemporary Urdu. Unlike earlier studies that classify borrowed vocabulary primarily according to domains such as education, commerce, or technology, the present research adopts a **word-formation-oriented approach**, arguing that lexical change must be examined in terms of how borrowed items behave morphologically and semantically after entering the language.

The data reveal that English-origin lexical items in Urdu do not merely function as borrowed forms. Rather, many of them undergo **systematic morphological adaptation, semantic stabilization, and grammatical integration**, eventually behaving like native Urdu nouns. This pattern cannot be fully explained through existing notions of borrowing, code-mixing, or lexical diffusion alone. Instead, it points toward the emergence of a **distinct word-formation process**, which this study terms **Morphological Naturalization**.

Accordingly, the data are organized not by domain but by **word-formation behavior**, allowing for a clearer understanding of how English-origin lexemes become structurally embedded within Urdu.

4.1 Scheme of Data Classification

Based on recurring patterns observed in the corpus, the data have been classified into three major analytical categories, each representing a different stage in the process of lexical integration.

Category 1: Lexical Borrowing with Orthographic Integration

This category includes English-origin words that are written in Urdu script and widely used in public and institutional contexts, yet retain their original morphological structure. These forms are visually and cognitively naturalized but do not undergo grammatical modification.

Examples include:

- پین (pen)
- پینسل (pencil)
- کالج (college)
- بیکری (bakery)
- بسکٹ (biscuit)
- ماسک (mask)
- چاکلیٹ (chocolate)

These items:

- have no widely used Urdu equivalents,
- appear frequently in signage, textbooks, and daily speech,
- are processed by speakers as Urdu words despite their foreign origin.

This category reflects *lexical adoption without structural change*.

Category 2: Borrowed Lexemes with Functional Stabilization

This category includes borrowed forms that have become socially normalized and semantically fixed, even where Urdu equivalents exist. Their continued use reflects communicative efficiency and conventionalization rather than lexical necessity.

Examples include:

- کچن (kitchen) → کچن
- کزن (cousin) → کزن
- جیولری (jewellery) → جیولری
- میک اپ (makeup) → میک اپ
- بیکری (bakery) → بیکری

- کاسمیٹک شاپ (cosmetic shop)

Although alternatives such as باورچی خانہ exist, these English-origin forms dominate actual usage. Their persistence demonstrates that lexical choice in Urdu is governed more by **frequency, social practice, and communicative convenience** than by linguistic purism.

Category 3: Morphological Naturalization of Borrowed Lexemes (Proposed Category)

This category represents the **core contribution of the present study**.

Here, English-origin nouns are not merely borrowed but are **fully integrated into the morphological system of Urdu**, particularly through:

- Urdu plural markers (یں / سوں -)
- hybrid plural formations
- syntactic embedding in Urdu sentence structures

Examples include:

- پین → پینوں
- پینسل → پینسلیوں
- کالج → کالجوں
- کار → کاریں
- سوفا → سوفرے / سوفوں
- ٹیکسی → ٹیکسیوں
- پرنسٹر → پرنسٹروں
- پیلشر → پیلشروں

These forms clearly demonstrate that borrowed lexemes:

- follow Urdu grammatical rules,
- take native plural suffixes,
- behave like indigenous nouns,
- Participate in further word formation.

This phenomenon cannot be adequately described as borrowing or code-mixing. Instead, it represents a **new stage of lexical integration**, where borrowed words become morphologically productive within the host language.

4.2 Rationale for a New Word-Formation Category

The data strongly support the recognition of *Morphological Naturalization* as a distinct word-formation process characterized by:

1. Lexical borrowing
2. Orthographic adaptation
3. Morphological integration
4. Semantic stabilization
5. Productive plural formation

Unlike earlier studies that treat borrowed forms as temporary or peripheral, the present findings demonstrate that such items enter the **core grammatical system of Urdu**.

Forms such as:

- پین → پینوں

- كالج → كالجون
- بسکٹ → بسکتون
- صوفه → صوفون

Show that Urdu is not merely borrowing vocabulary but **actively restructuring it through native grammatical mechanisms**.

4.3 Significance of the Categorization

This categorization is significant because it:

- Moves beyond domain-based classification
- Identifies a previously under-theorized word-formation process
- Demonstrates the role of public usage and corpus evidence in lexical change
- Shows that Urdu is a productive, adaptive linguistic system

Most importantly, the study establishes that lexical change in Urdu is **systematic rather than accidental**, driven by communicative necessity and morphological compatibility.

Although the corpus used in this study is not machine-processed, it fulfills the essential criteria of a linguistic corpus

- Authenticity
- systematic collection
- natural usage
- analytical relevance

Thus, the findings provide empirical evidence that *morphological naturalization constitutes a genuine word-formation process in contemporary Urdu*.

Category 3: Morphological Naturalization of Borrowed Lexemes in Urdu

(*Pluralization, Grammatical Integration, and Lexicographical Stabilization*)

While Categories 1 and 2 document large-scale lexical borrowing through transliteration and public usage, the corpus reveals a *structurally deeper and linguistically more significant pattern* in a subset of English-origin nouns. These items are not merely written in Urdu script or used in everyday communication; rather, they undergo *systematic morphological adaptation*, particularly through pluralization and syntactic integration.

This pattern indicates a shift from surface-level borrowing to *grammatical assimilation*, where borrowed lexemes begin to behave like native Urdu nouns. Forms such as اسکولون پاسٹل → اسکولون پاسٹل as borrowed nouns begin to behave like native Urdu nouns. Forms such as بوتلیں → بوتلیں پاسٹلون, سسٹم → سسٹمون → سسٹمون demonstrate that Urdu does not simply accommodate foreign vocabulary but actively *restructures it through its own morphological system*.

This process goes beyond code-mixing or lexical borrowing and reflects a deeper level of linguistic incorporation. The data therefore point toward a distinct word-formation mechanism, here termed *Morphological Naturalization*, in which borrowed lexemes become grammatically productive and lexicographically stable within Urdu.

Category 3.1: Pluralization of Borrowed Nouns

(*Morphological Naturalization through Urdu Plurals: -ات / -ون / -ین*)

Table 3.1

Borrowed Nouns with Urdu Plural Morphology

#	Singular (Urdu)	English Base	Plural (Urdu)	Morphological Process	Status
1	ٹیبل	Table	ٹیبلیں	Borrowing + Urdu plural	Fully naturalized
2	بک	Book	بکیں	Borrowing + Urdu plural	Fully naturalized
3	اسکول	School	اسکولوں	Borrowing + Urdu plural	Established
4	ہاسٹل	Hostel	ہاسٹلوں	Borrowing + Urdu plural	Established
5	سسٹم	System	سسٹموں	Borrowing + Urdu plural	Established
6	بوتل	Bottle	بوتلیں	Borrowing + Urdu plural	Fully naturalized
7	پنسل	Pencil	پنسلیں	Borrowing + Urdu plural	Fully naturalized
8	کار	Car	کاریں	Borrowing + Urdu plural	Fully naturalized
9	پلیٹ	Plate	پلیٹیں	Borrowing + Urdu plural	Fully naturalized
10	کاپی	Copy	کاپیاں	Borrowing + Urdu plural	Fully naturalized

3.1 Morphological Analysis of Borrowed Lexemes in Urdu

The data in Table 3.1 clearly demonstrate that borrowing in Urdu extends beyond lexical adoption and enters the domain of **morphological integration**. The pluralization of English-origin nouns through native Urdu suffixes (-یں / -وں) shows that these forms are no longer treated as foreign insertions but are processed according to Urdu grammatical rules.

This can be observed in naturally occurring constructions such as:

- دو اسکولوں میں مقابلہ ہوا
- میز پر تین بوتلیں بیس

In these examples, the borrowed nouns function identically to native Urdu nouns in terms of agreement, case marking, and syntactic positioning. Such usage cannot be explained through code-mixing, as code-mixed items typically retain their original grammatical structure and do not undergo full morphological adaptation.

Instead, the data indicate a process of **morphological domestication**, whereby borrowed lexemes are:

- integrated into Urdu noun classes,
- assigned native plural markers,
- used productively in syntax, and
- Cognitively processed as Urdu words by speakers.

This level of grammatical behavior confirms that the borrowed forms have crossed the threshold from lexical borrowing to **structural incorporation**.

4.4 Analytical Implication

The evidence presented in this category supports the central claim of the study: that Urdu exhibits a **distinct word-formation process** in which borrowed lexemes undergo systematic morphological naturalization.

Unlike traditional borrowing, which involves surface-level adoption, this process includes:

1. Orthographic adaptation
2. Morphological integration

3. Semantic stabilization
4. Productive grammatical behavior

This finding challenges earlier assumptions that borrowed words remain peripheral to the grammatical system and instead demonstrates that Urdu actively restructures foreign lexical material in accordance with its morphological rules.

The analysis presented in this section directly addresses the research questions of the study by demonstrating how English-origin nouns undergo systematic morphological naturalization in Urdu. Through corpus-based evidence, the study identifies recurring grammatical patterns—particularly pluralization and syntactic integration—that confirm the emergence of a distinct word-formation process. These findings fulfill the research objectives by empirically establishing morphological naturalization as a productive and linguistically significant mechanism in contemporary Urdu.

Category 3.2: Hybrid Pluralization in Borrowed Lexemes

(Urdu-English Morphological Alternation)

This category examines a distinct pattern of plural formation in which English-origin nouns exhibit **hybrid pluralization**, alternating between English plural markers (-s / -z) and Urdu plural suffixes (-وں / -ین). Unlike simple borrowing, these forms display **dual morphological behavior**, reflecting partial integration into Urdu grammar while retaining traces of their English origin.

This phenomenon is particularly significant because it demonstrates **transitional morphology**, where borrowed lexemes fluctuate between two grammatical systems before achieving full naturalization. Such alternation indicates that the lexemes are no longer peripheral borrowings but have entered the productive morphological system of Urdu.

Table 3.2

Hybrid Plural Formation in Borrowed Lexemes

#	Singular (Urdu)	English Base	Plural Form(s)	Morphological Process	Status
1	چیئر	Chair	چیئرز / چیئرین	English + Urdu plural	Established
2	موٹر	Motor	موٹرز / موٹرین	Hybrid pluralization	Established
3	ڈاکٹر	Doctor	ڈاکٹرز / ڈاکٹروں	Hybrid + Urdu plural	Fully naturalized
4	نرس	Nurse	نرسز / نرسوں	Hybrid + Urdu plural	Fully naturalized
5	پروفیسر	Professor	پروفیسرز / پروفیسروں	Hybrid pluralization	Emerging
6	ٹیچر	Teacher	ٹیچرز / ٹیچروں	Hybrid pluralization	Established
7	ورکر	Worker	ورکرز / ورکروں	Hybrid pluralization	Established
8	فریج	Fridge	فریجز / فریجوں	Hybrid pluralization	Emerging
9	مشین	Machine	مشینیں / مشینز	Hybrid pluralization	Fully naturalized
10	ممبر	Member	ممبرز / ممبروں	Hybrid pluralization	Established
11	منسٹر	Minister	منسٹرز / منسٹروں	Hybrid pluralization	Institutionalized

#	Singular (Urdu)	English Base	Plural Form(s)	Morphological Process	Status
12	یونیورسٹی	University	یونیورسٹیاں / یونیورسٹیز	Hybrid pluralization	Established
13	برانڈ	Brand	برانڈز / برانڈوں	Hybrid pluralization	Established
14	کیٹیگری	Category	کیٹیگریز / کیٹیگریاں	Hybrid pluralization	Emerging
15	سیگمنٹ	Segment	سیگمنٹس / سیگمنٹوں	Hybrid pluralization	Emerging
16	سولر پنل	Solar Panel	سولر پنلز / سولر پنلوں	Compound + Hybrid	Institutionalized
17	بورڈ	Board	بورڈز / بورڈوں	Hybrid pluralization	Fully naturalized
18	مشینری	Machinery	مشینریز / مشینریاں	Hybrid pluralization	Emerging

Morphological Interpretation (3.2)

The data presented in Table 3.2 demonstrate that a large number of English-origin nouns in Urdu exhibit **hybrid plural behavior**, whereby speakers alternate between:

- **English plural marking** (-s / -z), and
- **Urdu plural morphology** (-وں / -یں)

This alternation is not random. Instead, it reflects a **transitional stage of lexical integration**, where borrowed forms gradually shift from foreign lexical status toward full grammatical incorporation.

For example:

- ڈاکٹر → ڈاکٹرز / ڈاکٹروں
- ٹیچر → ٹیچرز / ٹیچروں
- مشین → مشینیں / مشینز

Such variation confirms that speakers treat these lexemes as **grammatically flexible units**, capable of adapting to Urdu inflectional patterns while retaining traces of English morphology.

Theoretical Interpretation

This hybrid behavior provides strong evidence for **Morphological Naturalization**, a process that lies between:

1. **Simple Borrowing** (lexical import only), and
2. **Full Morphological Integration** (complete Urdu inflection)

The data indicate that borrowed lexemes pass through an **intermediate hybrid stage**, during which:

- both plural systems coexist,
- usage frequency determines stabilization,
- Grammatical choices depend on context and speaker competence.

This aligns with usage-based theories of language change, which argue that grammar emerges from repeated usage rather than pre-fixed rules (Bybee, 2010).

Category 3.3: Compound + Plural Integration (Morphological Naturalization through Compound Formation)

This category represents the **most advanced stage of morphological naturalization** observed in the corpus. Unlike simple borrowed nouns or hybrid plural forms, the items in this category involve **compound lexical units of English origin that behave as single grammatical nouns in Urdu and undergo Urdu pluralization as whole units**.

In this pattern, **both elements of the compound are borrowed**, yet the compound functions as a **single morphological unit**, taking Urdu plural markers such as **-ں** and **-ں**. This indicates a high degree of grammatical assimilation and confirms that morphological naturalization in Urdu extends beyond individual lexemes to **multi-word constructions**.

Table 3.3

Compound Borrowings with Urdu Morphology

#	Singular (Urdu)	English Base	Plural (Urdu)	Word-Formation Process	Status
1	انک پاٹ	Ink Pot	انک پاٹیں	Compounding + Urdu plural	Integrated
2	پنسل باکس	Pencil Box	پنسل باکسیں	Compounding + Urdu plural	Integrated
3	کار پارک	Car Park	کار پارکیں	Compounding + Urdu plural	Integrated
4	بوتل شاپ	Bottle Shop	بوتل شاپیں	Compounding + Urdu plural	Integrated
5	بس اسٹاپ	Bus Stop	بس اسٹاپیں	Compounding + Urdu plural	Established
6	ٹیوب ویل	Tube Well	ٹیوب ویلیں	Compounding + Urdu plural	Integrated
7	سولر پینل	Solar Panel	سولر پینلیں	Compounding + Urdu plural	Emerging
8	فائل کور	File Cover	فائل کورز / فائل کوریں	Hybrid plural	Emerging

Qualitative Analysis (3.3)

The data in Table 3.3 demonstrate that **compound borrowings in Urdu are no longer treated as foreign lexical insertions**, but rather as **grammatically integrated noun units**. The entire compound—rather than an individual component—receives Urdu plural morphology, as shown in:

- انک پاٹ → انک پاٹیں
- پنسل باکس → پنسل باکسیں
- کار پارک → کار پارکیں

This pattern confirms that **Urdu morphology operates over borrowed compound structures**, not merely over isolated lexical items. Such behavior cannot be explained through conventional borrowing models, which assume that borrowed forms remain morphologically inert (Haugen, 1950; Thomason & Kaufman, 1988).

Instead of functioning as mere lexical insertions, the forms analyzed in this category reflect a process of **structural incorporation** into Urdu. In this process, a borrowed compound first enters the language through contact, gradually attains semantic stability, and subsequently undergoes pluralization through native Urdu morphological markers. Over time, such forms begin to function syntactically like indigenous nouns. This pattern is evident in examples such as **کار پارک → کار پارکیں** and **انک پاٹ → انک پاٹیں**, where the entire compound—not merely one lexical element—receives Urdu plural morphology. This behavior demonstrates that the borrowed items are no longer treated as foreign expressions but are cognitively processed as part of the Urdu grammatical system. Such developments align with usage-based models of morphology, which

propose that grammatical structure emerges through repeated use and frequency rather than through pre-established formal rules (Bybee, 2010).

The findings of Category 3.3 strongly support the argument that **morphological naturalization constitutes a distinct word-formation process**. Unlike simple lexical borrowing, this process involves compound-level integration, productive pluralization, and full grammatical compatibility with Urdu morphology. The repeated and systematic use of plural markers with borrowed compounds demonstrates that these forms have crossed the threshold from lexical borrowing to grammatical incorporation. Moreover, the consistency of these patterns across different lexical items indicates that this is not an isolated or stylistic phenomenon but a productive and rule-governed process within contemporary Urdu.

From a theoretical perspective, this pattern challenges traditional classifications of word formation. Classical models typically distinguish borrowing as a lexical process, compounding as a word-formation strategy, and inflection as a grammatical operation. However, the present data reveal a hybrid mechanism in which borrowed compounds undergo morphological adaptation and function as productive nouns within Urdu. This intermediate status—situated between lexical borrowing and grammatical morphology—has not been adequately described in earlier models of word formation (Booij, 2012; Bybee, 2010). The evidence therefore supports the proposal that **Morphological Naturalization** operates at the interface of borrowing, compounding, and inflection.

Overall, the analysis demonstrates that Urdu does not merely import foreign lexical material but actively restructures it according to its grammatical system. Borrowed forms are not retained as peripheral or marked items; instead, they become morphologically productive and syntactically integrated. This confirms that lexical change in Urdu is not random or degenerative but systematic and rule-governed. Consequently, morphological naturalization can be considered a distinct and previously under-theorized word-formation process, separate from borrowing, code-mixing, and surface-level lexical diffusion.

Category 3.4: Morphological Naturalization of Borrowed Lexemes

(Borrowed nouns with Urdu pluralization and no functional Urdu alternatives)

Table 3.4

Morphologically Naturalized Borrowed Nouns in Urdu

#	Singular (Urdu)	English Base	Plural (Urdu)	Word-Formation Process	Urdu Alternate?	Status
1	اسکول	School	اسکولوں	Borrowing + Urdu plural	✗	Fully naturalized
2	کالج	College	کالجوں	Borrowing + Urdu plural	✗	Fully naturalized
3	فیکٹری	Factory	فیکٹریوں	Borrowing + Urdu plural	✗	Fully naturalized
4	بوتل	Bottle	بوتلیں	Borrowing + Urdu plural	✗	Fully naturalized
5	سیلوں	Salon	سیلوںوں	Borrowing + Urdu plural	✗	Fully naturalized
6	پارلر	Parlour	پارلروں	Borrowing + Urdu	✗	Fully

#	Singular (Urdu)	English Base	Plural (Urdu)	Word-Formation Process	Urdu Alternate?	Status
				plural		naturalized
7	ریلنگ	Railing	ریلنگوں	Borrowing + Urdu plural	✗	Integrated
8	لیپ ٹاپ	Laptop	لیپ ٹاپوں	Borrowing + Urdu plural	✗	Fully naturalized
9	آئٹم	Item	آئٹمز / آئٹموں	Hybrid plural	✗	Fully naturalized
10	نمبر	Number	نمبروں	Borrowing + Urdu plural	✗	Fully naturalized
11	کلک	Click	کلکوں	Borrowing + Urdu plural	✗	Integrated
12	شیپر	Shaper	شیپروں	Borrowing + Urdu plural	✗	Integrated
13	فلٹر	Filter	فلٹروں	Borrowing + Urdu plural	✗	Fully naturalized
14	آئس کریم	Ice cream	آئس کریمیں	Borrowing + Urdu plural	✗	Fully naturalized
15	کینڈی	Candy	کینڈیاں	Borrowing + Urdu plural	✗	Fully naturalized
16	ٹافی	Toffee	ٹافیاں	Borrowing + Urdu plural	✗	Fully naturalized
17	بسکٹ	Biscuit	بسکٹوں	Borrowing + Urdu plural	✗	Fully naturalized
18	چاکلیٹ	Chocolate	چاکلیٹیں	Borrowing + Urdu plural	✗	Fully naturalized
19	فائبر	Fiber	فائبرز / فائبرے	Hybrid plural	✗	Integrated
20	اپریشن	Operation	اپریشنز	Borrowing + Hybrid plural	✗	Fully naturalized
21	تھیٹر	Theatre	تھیٹروں	Borrowing + Urdu plural	✗	Fully naturalized
22	کیٹیگری	Category	کیٹیگریاں	Borrowing + Urdu plural	✗	Fully naturalized

The lexical items presented above clearly demonstrate that **Urdu does not merely borrow vocabulary but structurally absorbs it**. These forms show consistent application of Urdu plural markers such as:

- -یں-
- -وں-

- **یاں** –
- Hybrid formations (ز / وون –)

This behavior confirms that these lexemes are no longer foreign insertions but have become morphologically productive units within Urdu grammar.

For example:

- اسکول → اسکولوں
- بوتل → بوتلیں
- ٹافی → ٹافیاں
- فیکٹری → فیکٹریوں
- چاکلیٹ → چاکلیٹیں

These forms:

- follow Urdu inflectional rules
- occur naturally in spoken and written discourse
- have no commonly used Urdu alternatives
- are cognitively processed as Urdu words

Why These Belong to Category 3 (Not Category 1 or 2)

Not Category 1 → because they are not just written in Urdu script

Not Category 2 → because they are not merely stylistic choices

Category 3 → because they show morphological productivity

This confirms that: **Morphological Naturalization is a distinct word-formation process**, not reducible to borrowing or code-mixing.

Theoretical Implication

The data strongly supports the claim that:

- Borrowing → Integration → Morphological Productivity
- Lexical change is **systematic**, not random
- Urdu expands through **grammatical accommodation**, not erosion

This aligns with usage-based morphology (Bybee, 2010) and lexical entrenchment theory.

This category highlights a crucial stage in the process of lexical integration: **institutional and technical lexicalization**. Unlike earlier categories where borrowing occurs through general usage or morphological adaptation, the items in this group gain stability primarily through **institutional repetition and functional necessity**. Words such as سسٹم and بفی سسٹم appear consistently in formal domains including hotels, administrative discourse, service industries, and commercial documentation.

The lexical item بفی سسٹم (buffet) provides a particularly strong example of this phenomenon. Urdu lacks a concise and functionally equivalent term for *buffet*; as a result, the borrowed form has become entrenched in everyday usage. The compound بفی سسٹم is now processed as a single lexical unit rather than as a code-switched phrase. Its frequent occurrence in menus, advertisements, and spoken discourse has led to **semantic stabilization** and **institutional normalization**, two key indicators of lexicalization.

Similarly, گھوٹا demonstrates how technical vocabulary enters Urdu through necessity and becomes grammatically productive. Its plural form گھوٹاں follows Urdu morphological rules, indicating full grammatical integration. These forms are no longer perceived as foreign or stylistically marked; rather, they function as default lexical choices in professional and public contexts.

This category strengthens the argument that *morphological naturalization operates beyond individual word usage and extends into institutional language practices*. The data show that:

- Borrowed lexemes become fixed through repeated institutional use
- Lexicalization is driven by communicative efficiency rather than prestige
- Absence of Urdu equivalents accelerates permanent adoption
- Institutional frequency leads to dictionary-level legitimacy

Unlike casual borrowing or code-switching, these forms exhibit *high functional load and semantic indispensability*, making them resistant to replacement.

Traditional models treat borrowing as a peripheral lexical process. However, the evidence from Category 3.4 demonstrates a more advanced mechanism:

Borrowing → Institutional Usage → Morphological Integration →

Lexicalization

This process supports usage-based and constructionist views of morphology (Bybee, 2010; Booij, 2012), which argue that grammatical structure emerges through repeated use rather than prescriptive rules. The findings confirm that institutional discourse plays a decisive role in transforming borrowed items into permanent components of Urdu vocabulary. Forms such as سسٹم and بفی are not transitional borrowings but **lexically stabilized units**, embedded in both grammar and cognition. This category therefore provides strong empirical support for the study's central claim:

Morphological Naturalization is a distinct and productive word-formation process in contemporary Urdu.

The analysis demonstrates that English-origin lexemes in Urdu undergo a process of morphological naturalization whereby they acquire Urdu plural morphology, participate in syntactic structures, and function as fully integrated lexical items. This process constitutes a distinct word-formation mechanism not previously identified in Urdu linguistic scholarship.

4.6 Position of Morphological Naturalization in Word-Formation Theory

In traditional morphological theory, word formation is broadly divided into **lexical processes** (such as borrowing and compounding) and **grammatical processes** (such as inflection and derivation). However, the empirical evidence presented in the present study reveals a pattern that does not fully conform to either of these established categories.

The phenomenon identified here—termed **Morphological Naturalization**—occupies an intermediate and previously under-theorized position between lexical borrowing and grammatical morphology. While the lexical base of the items originates from English, their subsequent linguistic behavior is governed by the grammatical rules of Urdu, particularly in terms of plural formation, syntactic embedding, and morphological productivity.

Unlike ordinary borrowing, these lexemes do not remain structurally inert or stylistically marked. Nor can they be treated as simple cases of inflection, since the lexical roots themselves are foreign in origin. Instead, they undergo a process through which borrowed forms are

systematically absorbed into the morphological system of Urdu and begin to function as productive members of the lexicon.

This process involves multiple stages, including:

- lexical adoption from a donor language,
- morphological restructuring according to Urdu grammar,
- grammatical productivity through pluralization and syntactic integration, and
- Cognitive acceptance as native lexical items by speakers.

Table 3.5

Classification Summary

Level	Traditional Category	Status in Present Study
Lexical	Borrowing	✓ Present
	Morphological Inflection	✓ Present
Derivational	New word formation	✓ Partially
	New Category Not previously defined	✓ Morphological Naturalization

4.7 Theoretical Implication

Morphological Naturalization represents a **previously unrecognized word-formation process** in Urdu. It differs from simple borrowing in that borrowed items do not remain static or peripheral. Instead, they undergo grammatical integration and participate actively in Urdu morphological patterns.

This places morphological naturalization at the *interface of lexical borrowing and grammatical morphology*, supporting usage-based and constructionist models of language change (Bybee, 2010; Booij, 2012). The findings demonstrate that word-formation is not limited to derivation or compounding alone but may also emerge through sustained usage, institutional reinforcement, and morphological accommodation.

Consequently, the study argues that morphological naturalization should be recognized as an independent word-formation mechanism in Urdu, distinct from:

- borrowing,
- code-mixing,
- And surface-level lexical diffusion.

Before presenting the consolidated model of Morphological Naturalization, it is necessary to synthesize the analytical insights gained from Categories 3.1 to 3.4. The preceding sections demonstrated that English-origin lexical items in Urdu do not remain confined to surface-level borrowing or stylistic usage. Instead, the data revealed a gradual yet systematic process through which borrowed forms undergo orthographic adaptation, morphological integration, and grammatical stabilization. Each category highlighted a different stage of this process—from simple pluralization to hybrid formation and institutional lexicalization—showing that these changes are neither random nor exceptional.

Taken together, the findings indicate that lexical borrowing in Urdu follows a structured developmental trajectory rather than an isolated or unsystematic pattern. The recurrence of plural marking, syntactic embedding, and productive usage across multiple domains suggests the

existence of an underlying word-formation mechanism operating beyond traditional models of borrowing and code-mixing. In order to capture this progression more clearly, the following table summarizes the sequential stages through which borrowed lexemes evolve into fully naturalized elements of the Urdu lexicon.

Table 3.5 therefore presents a synthesized model of this process, illustrating how borrowed English lexemes move from initial contact to full lexical integration. This model serves as a conceptual bridge between the empirical data and the theoretical claim advanced in this study—namely, that Urdu exhibits a distinct and previously unrecognized word-formation process termed *Morphological Naturalization*.

Table 3.6

Summary of the Proposed Morphological Naturalization Process

Stage	Description
Borrowing	English lexeme enters Urdu discourse
Orthographic Adaptation	Lexeme is represented in Urdu script
Morphological Integration	Urdu plural markers applied (—وں / یں)
Productivity	Lexeme participates in compounding and plural formation
Stabilization	Frequent use in public and institutional contexts
Lexicalization	Lexeme becomes part of Urdu lexical inventory

4.8. Grammatical Domestication of Loanwords (GDL)

Grammatical Domestication of Loanwords refers to a word-formation process in which borrowed lexical items are not only adopted into Urdu but are also **morphologically integrated**, **grammatically productive**, and **lexically stabilized** through native inflectional and syntactic mechanisms.

In this process, borrowed lexemes undergo:

- Urdu pluralization (—وں / یں),
- syntactic embedding in native constructions,
- and repeated institutional and social usage,

Resulting in their treatment as *fully functional Urdu nouns* rather than foreign insertions

4.9 Corpus-Based Evidence

Taken together, **Tables 3.1–3.4** constitute a structured, empirically grounded dataset representing naturally occurring language use. Although manually compiled, the dataset fulfills key criteria of a corpus in corpus linguistics:

- authenticity of data,
- systematic collection,
- functional consistency,
- And analytical reproducibility.

If expanded to a larger scale (e.g., 1,000+ entries), this dataset would constitute a **specialized morphological corpus**, suitable for corpus-driven analysis of loanword integration in Urdu.

Figure 4.1

Morphological Naturalization of Borrowed Lexemes in Urdu

Morphological Naturalization of Borrowed Lexemes in

ENGLISH LEXEME

(Table, Bottle, System, Pencil)

LEXICAL BORROWING

- Entry into Urdu
- Transliteration
- Public usage

MORPHOLOGICAL INTEGRATION

- Urdu plural markers (بن / ون)
- Phonological adaptation
- Syntactic compatibility

Examples: بونل → بونلیں | سسٹم → سسٹمون | بنسٹل → بنسٹلیں

MORPHOLOGICAL NATURALIZATION

- Productive morphology
- Native-like behavior
- Cognitive acceptance

LEXICALIZATION IN URDU

- Institutional usage
- Textbooks
- Dictionary-level stability

Note. This model illustrates the gradual transition of English-origin lexical items from surface borrowing to full lexicalization in Urdu. The process shows how borrowed forms move through stages of orthographic adaptation, morphological integration, and grammatical productivity, ultimately becoming indistinguishable from native Urdu nouns. The proposed model demonstrates that borrowed English lexemes in Urdu do not remain at the level of superficial borrowing. Instead, they undergo a structured process of morphological integration in which:

- Urdu plural markers are applied,
- syntactic behavior aligns with native noun patterns,
- semantic stability emerges through repeated use,
- Institutional acceptance reinforces permanence.

This process results in **morphological naturalization**, whereby borrowed forms are cognitively processed and grammatically treated as Urdu lexemes.

4.10 Theoretical Contribution

The findings of this study extend traditional word-formation theory by identifying a **third pathway** of lexical development:

Table 3.7

This study proposes Morphological Naturalization as an additional word-formation process that complements traditional mechanisms such as borrowing, inflection, and derivation

Traditional Processes	Proposed Addition
Borrowing	✓
Inflection	✓
Derivation	✓
Morphological Naturalization	✓ Newly Identified

Unlike simple borrowing or code-switching, morphological naturalization involves:

- grammatical restructuring,
- morphological productivity,
- and long-term lexical entrenchment.

This aligns with usage-based and constructionist views of morphology, which argue that grammatical structure emerges from repeated use and communicative necessity (Bybee, 2010; Booij, 2012).

The present study demonstrates that English-origin lexemes in Urdu undergo a systematic process of **morphological naturalization**, whereby they are absorbed into the grammatical system of Urdu and function as fully integrated lexical items. This process represents a distinct and previously undocumented word-formation mechanism in Urdu.

By identifying and empirically validating this process, the study contributes to:

- word-formation theory,
- lexical semantics,
- contact linguistics,
- and Urdu lexicography.

It further establishes that lexical change in Urdu is not random or degenerative, but structurally governed, usage-driven, and linguistically productive.

4.11 Naturalization of Borrowed Lexemes (MNBL)

Table 3.8

A Newly Identified Word-Formation Process in Urdu, Morphological Naturalization of Borrowed Lexemes in Urdu

Stage	Description	Linguistic Evidence	Example
1. Lexical Borrowing	English lexeme enters Urdu via contact	Loanword adoption	Table → ٹیبل
2. Orthographic Adaptation	Word written in Urdu script	Transliteration	Bottle → بوتل
3. Morphological Integration	Urdu plural / inflection applied	بن / سون / ات-	بوتل → بوتلیں

Stage	Description	Linguistic Evidence	Example
4. Syntactic Embedding	Word behaves like Urdu noun	Case, number agreement	دو بوتیں بیں
5. Productivity	Word allows further derivation	Compound & plural forms	بوتیں شاپ، بوتیں
6. Lexical Stabilization	Word accepted as Urdu lexeme	Dictionary-level usage	Textbooks, signage
7. Cognitive Naturalization	Speakers process it as Urdu	No perception of “English”	Native-like usage

Why This Is a New Word-Formation Process

1. What Is a Word and a Lexeme?

A **word** is a linguistic form used in speech or writing, while a **lexeme** is an abstract vocabulary unit that includes all its inflected forms (Crystal, 2008).

For example:

- **Lexeme:** BOOK
- **Word forms:** book, books

In Urdu:

- **Lexeme:** بوتیں
- **Word forms:** بوتیں، بوتیں

Once a borrowed word behaves like this, it ceases to be foreign.

2. Traditional Word-Formation Processes (Established in Linguistics)

Table 3.9

Traditional Word-Formation Processes (Established in Linguistics)

Process	Example	Nature
Borrowing	taxi, radio	Lexical import
Derivation	happy → happiness	Affix-based
Inflection	book → books	Grammatical
Compounding	toothpaste	Lexical
Blending	brunch	Structural
Clipping	lab	Shortening

Problem None of these fully explain what happens in Urdu with words like:

- بوتیں → بوتیں
- سسٹم → سسٹم
- پنسل → پنسل
- اسکول → اسکول

These are **not simple borrowings**, because:

- they take Urdu morphology
- they follow Urdu grammar
- they behave like native nouns
- they appear in textbooks and signage

3. Why This Is a New Word-Formation Process

Current study's data shows a **multi-stage linguistic phenomenon**, not covered in classical word-formation theory:

Step 1: Borrowing

English word enters Urdu.

Step 2: Morphological Adaptation

Urdu plural markers attach:

- \rightarrow بے
- \rightarrow وے
- \rightarrow اتے

Step 3: Grammatical Productivity

The word:

- pluralizes
- combines with adjectives
- participates in syntax

Step 4: Lexical Stabilization

It becomes:

- cognitively Urdu
- socially normalized
- lexicographically valid

This process is neither borrowing nor derivation alone. So the *Proposed Term with the analysis and data of the current study*

4.12. Morphological Naturalization of Borrowed Lexemes (MNBL)

Definition (Proposed):

Morphological Naturalization of Borrowed Lexemes refers to a word-formation process in which foreign lexical items, after being borrowed into a language, undergo systematic morphological integration and begin to behave as native lexical units, exhibiting pluralization, syntactic compatibility, and semantic stability within the host language.

Table 3.10

Showing Why This Process Is Unique

Feature	Borrowing	Code-Mixing	MNBL (current Study)
Temporary use	✓	✓	✗
Morphological marking	✗	✗	✓
Productivity	✗	✗	✓
Cognitive acceptance	✗	✗	✓
Dictionary potential	✗	✗	✓
Institutional usage	✗	✗	✓

Table 3.11

Examples Proving MNBL

English Urdu Singular Urdu Plural

Bottle	بottle	بottle
Pencil	پنسل	پنسلیں
System	سسٹم	سسٹمیں
Car	کار	کاریں
Plate	پلیٹ	پلیٹیں
Copy	کاپی	کاپیاں

These forms follow **Urdu morphology**, not English grammar.

Why This Has Not Been Identified Earlier

1. Previous studies focused on:
 - o Code-switching
 - o Borrowing
 - o Loanwords
2. No study:
 - o Treated pluralization as evidence of word-formation
 - o Used corpus-based street + textbook data
 - o Examined morphological productivity
3. Your study:
 - o Uses **natural corpus**
 - o Tracks **structural behavior**
 - o Proposes **a new theoretical category**

Note. *This study proposes Morphological Naturalization of Borrowed Lexemes as a distinct word-formation process in Urdu, whereby borrowed English nouns undergo systematic morphological integration and attain full lexical status. This process extends existing models of word formation and demonstrates that lexical change in Urdu is not merely a result of borrowing, but of structural and grammatical assimilation.*

4.13 Morphological Naturalization as a Newly Identified Word-Formation Process

The present study proposes **Morphological Naturalization** as a distinct and previously underexplored word-formation process in Urdu. In order to establish the theoretical validity of this claim, it is necessary to first clarify the linguistic concepts of *word* and *lexeme*. A word refers to a concrete linguistic form used in speech or writing, whereas a lexeme represents an abstract lexical unit that may appear in different inflected or morphologically modified forms (Cruse, 2011). For instance, *TABLE* functions as a lexeme, while *table*, *tables*, and its Urdu-adapted form ٹیبل represent its realized word forms. This distinction is crucial because the present study does not merely examine surface-level borrowing, but rather the grammatical transformation of lexemes within the Urdu morphological system.

Traditional models of word formation recognize a limited set of mechanisms, including borrowing, derivation, compounding, inflection, conversion, and blending. While these categories account for many structural changes in language, they fail to fully explain the process through which English lexical items enter Urdu, acquire Urdu plural morphology, behave syntactically like native nouns, and ultimately gain social and cognitive acceptance. Borrowing alone cannot explain this phenomenon, as borrowed words are typically treated as foreign elements. Likewise, inflection fails to capture the lexical permanence achieved by such forms,

and derivation does not apply because no category change or semantic expansion occurs. This theoretical gap necessitates the identification of a new process.

The present study introduces **Morphological Naturalization** as a distinct word-formation mechanism. This process explains how borrowed English lexemes become fully integrated into Urdu through a series of gradual and systematic stages. Initially, an English lexeme enters Urdu through lexical borrowing, primarily via public usage, education, media, and institutional discourse. At this stage, the word retains its foreign identity but becomes recognizable within Urdu speech. Subsequently, the borrowed item undergoes morphological integration, during which it adopts Urdu grammatical markers, particularly plural suffixes such as **-پں** and **-وں**, and conforms to Urdu syntactic patterns. Over time, this repeated and rule-governed usage leads to morphological naturalization, where the word begins to behave like a native lexical item. The final stage is lexicalization, in which the word gains institutional acceptance, appears in textbooks and formal registers, and eventually qualifies for dictionary inclusion.

This process differs fundamentally from traditional borrowing because it involves structural assimilation rather than surface adoption. The borrowed lexeme does not remain foreign; instead, it becomes morphologically productive and grammatically indistinguishable from native words. For example, native nouns such as **بُوتنیں** → **بُوٹنے** behave in the same way as borrowed forms like **بُوٹنیں** → **بُوٹنے** or **بُوٹنیں** → **بُوٹنے**. The parallel behavior of native and borrowed nouns demonstrates that the latter have undergone full grammatical incorporation rather than mere inflection.

Importantly, this process cannot be categorized as derivational morphology either. Derivation typically results in a new lexical item, often involving a change in word class or semantic scope. In contrast, morphological naturalization does not alter the semantic core or grammatical category of the borrowed word. Instead, it allows the lexeme to participate in the native morphological system without altering its meaning. For instance, *bottle* → **بُوٹل** → **بُوٹنیں** → **بُوٹنے** shows neither semantic shift nor category change, yet the pluralized form functions identically to native Urdu nouns. This confirms that morphological naturalization represents a unique word-formation process rather than a subtype of derivation or inflection.

The empirical data further demonstrate that this process is productive and systematic. Borrowed nouns regularly appear with Urdu plural markers, participate in compounding, and occur in institutional registers such as education, administration, and print media. Examples such as **اسکولوں**, **پسٹمیں**, and **بُوٹنیں** illustrate how these forms have become grammatically stable and socially accepted. Their frequent occurrence in textbooks and formal discourse indicates that they are no longer perceived as foreign elements but as integral components of the Urdu lexicon.

This study therefore proposes *Morphological Naturalization* as a new category within word-formation theory. The process occupies an intermediate position between borrowing and lexicalization, representing a gradual transition from foreign lexical item to fully naturalized word. Unlike traditional borrowing, morphological naturalization involves rule-based grammatical adaptation. Unlike derivation, it does not create new meanings or word classes. Instead, it reflects the dynamic capacity of Urdu to absorb foreign lexemes and restructure them according to its morphological system.

From a theoretical perspective, this finding contributes to word-formation studies by introducing a mechanism that has not been adequately addressed in existing models. It challenges the rigid classification of word formation into inflection and derivation and highlights the need for an expanded framework that accounts for contact-induced morphological change.

The proposed model also holds significance for lexicography, as it provides criteria for determining when a borrowed word qualifies for dictionary inclusion. Furthermore, it offers valuable insights for corpus linguistics by enabling systematic tracking of lexical integration across time and domains.

In applied contexts, the Morphological Naturalization Framework can be used in language teaching, curriculum development, and lexicographical compilation. It explains why forms such as *اسکولیں سسٹمون* *پوٹیں* are not erroneous or informal but represent grammatically legitimate outcomes of language contact. The framework also offers a methodological tool for analyzing similar processes in other contact languages such as Punjabi, Hindi, Pashto, and Bengali.

In conclusion, the findings of this study demonstrate that morphological naturalization is neither accidental nor marginal. It represents a systematic, productive, and rule-governed word-formation process through which Urdu expands its lexicon in response to social, technological, and cultural change. By identifying and theorizing this process, the study contributes a novel analytical category to morphological theory and provides a robust framework for future research on lexical change and language contact.

5. Suggestions and Recommendations

The present study has introduced *Morphological Naturalization of Borrowed Lexemes (MNBL)* as a previously unrecognized word-formation process in Urdu. Based on the empirical findings and theoretical insights of the study, several important suggestions and recommendations are proposed for future research, corpus development, and theoretical advancement in linguistics.

5.1 Implications for Future Linguistic Research

This study demonstrates that word-formation processes in Urdu cannot be fully explained through traditional categories such as borrowing, inflection, or derivation alone. The identification of *Morphological Naturalization* suggests that language contact leads not only to lexical adoption but also to *systematic grammatical integration*.

Future researchers are therefore encouraged to:

- Investigate **morphological naturalization in other varieties of Punjabi**, such as Majhi, Potohari, Siraiki, and Hindko.
- Examine whether similar processes occur in **other contact languages**, including Hindi, Pashto, Bengali, and Sindhi.
- Explore how borrowed lexemes behave morphologically across **different dialects of the same language**, as dialectal variation may reveal different degrees of grammatical integration.

This will help establish whether morphological naturalization is a *language-specific phenomenon or a cross-linguistic pattern of contact-induced change*.

This aligns with usage-based and contact linguistics perspectives, which view grammatical structure as emerging from repeated language use (Bybee, 2010; Thomason, 2001).

5.2 Recommendations for Corpus-Based Research

One of the major contributions of this study is the construction of a **small, specialized morphological corpus** based on naturally occurring Urdu data. However, future studies can extend this work by:

- Developing **large-scale digital corpora** focusing specifically on borrowed and naturalized lexemes.
- Tagging data for:
 - morphological behavior

- plural formation
- syntactic position
- frequency of occurrence
- Comparing spoken vs. written corpora to observe differences in naturalization patterns.
- Creating diachronic **corpora** to trace how borrowed forms evolve over time.

Such corpus-based expansion would allow quantitative validation of Morphological Naturalization and contribute to computational linguistics and lexicography.

Corpus-driven approaches are particularly suited for identifying emergent grammatical patterns (Biber, Conrad & Reppen, 1998; McEnery & Hardie, 2012).

5.3 Pedagogical and Lexicographical Implications

The findings of this study have direct implications for:

Language Teaching

- Urdu language textbooks can include naturalized forms such as:

○ بوئیں
 ○ سسٹمون
 ○ پنسلیں

- These forms should no longer be marked as “incorrect” or “informal”.
- Teachers can explain them as examples of **productive morphological adaptation**, not language decay.

5.4. Dictionary Compilation

Lexicographers may use the MNBL framework to:

- Decide when a borrowed word qualifies for dictionary inclusion
- Distinguish between temporary borrowings and stabilized lexemes
- Record plural forms as legitimate entries

This supports the view that **lexicography must reflect actual language use**, not prescriptive ideals (Booij, 2012).

5.5 Theoretical Significance and Model Applicability

The proposed theory of **Morphological Naturalization** offers a new analytical lens for word-formation studies. It can be applied in:

- **Morphology – to identify non-traditional formation processes**
- **Sociolinguistics – to study language contact and change**
- **Corpus Linguistics – to trace frequency-based grammaticalization.**
- **Cognitive Linguistics – to explain lexical entrenchment**
- **Language Planning & Policy – to understand natural lexical evolution**

Unlike classical models that restrict word formation to inflection or derivation, MNBL recognizes a **third pathway**, where borrowed items become grammatically productive without changing category or meaning.

This aligns with construction-based and usage-based models of grammar (Bybee, 2010; Booij, 2012), which argue that grammar emerges from repeated usage patterns rather than rigid rules.

5.6 Directions for Future Research

Future studies may extend this framework by:

1. Applying MNBL to **spoken corpora**
2. Conducting **comparative studies** across South Asian languages
3. Investigating **phonological adaptation** alongside morphology
4. Exploring **children's acquisition** of naturalized forms
5. Studying **diachronic change** to observe long-term lexical stabilization

6. Developing **computational tools** for detecting naturalized forms automatically. Such studies will strengthen the empirical grounding of Morphological Naturalization and establish it as a robust analytical category in linguistics.

This study recommends the recognition of *Morphological Naturalization of Borrowed Lexemes (MNLB) as a distinct word-formation process* operating at the interface of borrowing and grammatical morphology. The process demonstrates that Urdu does not merely absorb foreign vocabulary but actively restructures it according to its grammatical system.

By proposing this framework, the study contributes:

- a new theoretical category,
- a corpus-based analytical method,
- And a replicable model for future linguistic research.

The findings strongly suggest that lexical change in Urdu is not accidental or chaotic but systematic, rule-governed, and cognitively grounded.

REFERENCES

Backus, A. (2014). A usage-based approach to code-switching: The role of frequency and entrenchment. *Journal of Language Contact*, 7(1), 27–63. <https://doi.org/10.1163/19552629-00701002>

Backus, A. (2014). The role of entrenchment in contact-induced language change. In J. Besters-Dilger, C. Dermarkar, S. Pfänder, & A. Rabus (Eds.), *Congruence in contact-induced language change* (pp. 51–68). De Gruyter Mouton.

Baumgardner, R. J. (1993). The indigenization of English in Pakistan. *World Englishes*, 12(1), 1–13. <https://doi.org/10.1111/j.1467-971X.1993.tb00003.x>

Biber, D., Conrad, S., & Reppen, R. (1998). *Corpus linguistics: Investigating language structure and use*. Cambridge University Press.

Booij, G. (2012). *The grammar of words: An introduction to linguistic morphology* (3rd ed.). Oxford University Press.

Bybee, J. (2010). *Language, usage and cognition*. Cambridge University Press. <https://doi.org/10.1017/CBO9780511750526>

Croft, W., & Cruse, D. A. (2004). *Cognitive linguistics*. Cambridge University Press.

Crystal, D. (2008). *A dictionary of linguistics and phonetics* (6th ed.). Blackwell.

Glaser, B. G., & Strauss, A. L. (1967). *The discovery of grounded theory: Strategies for qualitative research*. Aldine.

Haspelmath, M. (2009). Lexical borrowing: Concepts and issues. In M. Haspelmath & U. Tadmor (Eds.), *Loanwords in the world's languages* (pp. 35–54). De Gruyter.

Haugen, E. (1950). The analysis of linguistic borrowing. *Language*, 26(2), 210–231.
<https://doi.org/10.2307/410058>

Kachru, B. B. (1986). *The alchemy of English: The spread, functions, and models of non-native Englishes*. Pergamon Press.

Kachru, B. B. (1994). Englishization and contact linguistics. *World Englishes*, 13(2), 135–154.
<https://doi.org/10.1111/j.1467-971X.1994.tb00287.x>

Mahboob, A. (2009). English as an Islamic language: A case study of Pakistani English. *World Englishes*, 28(2), 175–189. <https://doi.org/10.1111/j.1467-971X.2009.01583.x>

Matras, Y. (2009). *Language contact*. Cambridge University Press.

McEnery, T., & Hardie, A. (2012). *Corpus linguistics: Method, theory and practice*. Cambridge University Press.

Myers-Scotton, C. (2002). *Contact linguistics: Bilingual encounters and grammatical outcomes*. Oxford University Press.

Platts, J. T. (1904). *A dictionary of Urdu, classical Hindi, and English*. Oxford University Press.

Quyyum, A. (2023). *A corpus-based contrastive analysis of lexical reduplication in Punjabi varieties* (Unpublished doctoral dissertation).

Rahman, T. (2004). Language policy and localization in Pakistan. *Orientations*, 35(1), 1–26.

Rahman, T. (2004). Language policy and localization in Pakistan: Proposal for a paradigm shift. *Current Issues in Language Planning*, 5(3), 245–264.
<https://doi.org/10.1080/14664200408668248>

Rahman, T. (2011). *From Hindi to Urdu: A social and political history*. Oxford University Press.

Schmidt, R. (1999). Lexical development in second language acquisition. *Applied Linguistics*, 20(3), 357–386.

Thomason, S. G. (2001). *Language contact: An introduction*. Edinburgh University Press.

Thomason, S. G., & Kaufman, T. (1988). *Language contact, creolization, and genetic linguistics*. University of California Press.