

## Bridging Theory to Empirical Insights: Revisiting Thai Caused-Separation Verbs<sup>1</sup>

*Nitipong Pichetpan*

Faculty of Liberal Arts, Thammasat University, Thailand

Email: pnitipon@tu.ac.th

Received 13 May 2025; revised 30 July 2025; accepted 31 July 2025

### Abstract

Studies of Thai caused-separation verbs elucidate how speakers conceptualize and articulate events involving physical separation. While prior research has examined this verbal domain through both rule-based (henceforth, *prescriptive*) and usage-driven (*empirical*) approaches, a systematic comparison between the two frameworks remains lacking. This study aims to bridge the gap by offering the first comprehensive, comparative analysis of Thai caused-separation verbs across prescriptive taxonomies and empirical usage data. Methodologically, it proceeds in two stages: (1) reviewing both prescriptive and empirical accounts of Thai caused-separation verbs, and (2) assessing the degree of convergence or divergence between them in terms of syntactic realization, semantic categorization, and lexical selection. Findings reveal that, although prescriptive models capture certain structural generalizations, they often overlook the contextual and pragmatic variability observed in usage. Thai caused-separation verbs exhibit notable syntactic flexibility—particularly through serial verb constructions—and often display overlapping or fluid categorization shaped by instrument type, object properties, and communicative intent. By comparing theoretical constructs with real-world language usage, this study offers a novel analytical lens, reinforcing the need for a hybrid, probabilistic model of verb categorization that more accurately reflects how Thai speakers select and organize verbs. The results have broader implications for linguistic typology, cognitive semantics, and corpus-based language modeling of syntactic variation and context-sensitive verb use, with applications to Thai lexical resources for language education, lexicography, and natural language processing.

**Keywords:** event categorization, lexical semantics, syntactic variability, Thai language, caused-separation verbs

---

<sup>1</sup> This article draws in part on the author's PhD dissertation, *Verbs of Caused-Separation in Thai and Khmer: Lexical Semantics and Language Convergence in Mainland Southeast Asia* (University of Sydney, 2021), which serves as empirical grounding for this study's literature review, analysis, and discussion.

## 1. Introduction

### 1.1 Theoretical Background: Semantic Categorization and Caused-Separation Events

Semantic categorization—how languages group concepts by shared features—has long been central to linguistics, anthropology, and psychology (Du et al., 2023; Murphy, 2002). Foundational studies in domains such as color (Berlin & Kay, 1969), emotion (Ekman, 1972), and kinship (Goodenough, 1965; Wallace & Atkins, 1960) have revealed both universal and language-specific patterns, prompting cross-linguistic investigations into conceptual structure (Majid et al., 2008). More recently, research has shifted toward the categorization of dynamic events—such as motion or state change—focusing on features like trajectory, instrument, and outcome (Du et al., 2023; Pye et al., 1995). Cognitive linguistic approaches underscore that language-specific grammatical structures and cultural framing systematically shape event construal, leading to cross-linguistic variation (Papafragou & Selimis, 2010; Talmy, 1985).

Among such events, *separation*—in which the physical integrity of an object is altered—serves as a compelling test case for examining the interface between meaning, structure, and categorization. These events vary in agency and reversibility, ranging from *spontaneous* (e.g., *cracking*) and *reversible* (e.g., *opening a container*) to *caused separation* (e.g., *cutting* or *smashing* by deliberate external agency) (Hale & Keyser, 1987; Pichetpan, 2021; Pye et al., 1995).

Caused-separation events reveal cross-linguistic variation in how instrumentality, force, and outcome are encoded, with languages adopting diverse lexical strategies (Du et al., 2023; Majid et al., 2007a). English differentiates “cut” and “break” by instrument and result, while Germanic languages emphasize separation predictability (Majid et al., 2007b). Other languages highlight different features—Mandarin encodes blade types (Hsiao, 2015), Japanese generalizes cutting under *kiru* (Saji et al., 2024).

Thai, meanwhile, exhibits some cross-linguistic patterns but stands out for its fine lexical distinctions and syntactic flexibility.

A hallmark of Thai is its use of serial verb constructions (SVCs) to encode combinations of instrument, action, and outcome (Sudmuk, 2005; Thepchuaysuk, 2016; Thepkanjana & Uehara, 2009). These constructions enable fine-grained distinctions between realized and unrealized outcomes—e.g., *tàt p<sup>h</sup>â. khà.ìt* ('cut cloth torn') vs. *tàt p<sup>h</sup>â. mâj khà.ìt* ('cut cloth not torn'). In contrast, monoverbal clauses like *tàt p<sup>h</sup>â.* ('cut cloth') often leave results pragmatically inferred (Pichetpan, 2021). Core Thai verbs like *tàt* 'cut,' *t<sup>h</sup>úp* 'smash,' and *hàk* 'snap' reflect semantic distinctions grounded in tool use, object properties, and result predictability (Pichetpan, 2021).

## 1.2 Review of Prescriptive and Empirical Approaches in Thai

While both prescriptive (e.g., Phanthumetha, 2016; Premsrirat, 1987) and empirical (e.g., Pichetpan, 2021; Thepchuaysuk, 2016) studies have addressed caused-separation verbs in Thai, they often proceed independently. Prescriptive work tends to focus on idealized semantic hierarchies and static features, whereas empirical research relies on usage data and speaker intuition. However, no systematic analysis yet compares how prescriptive classifications align with actual usage, contextual variation, and syntactic diversity—leaving a critical gap in understanding how caused-separation events in Thai are lexically and syntactically represented.

Early Thai prescriptive accounts often relied on idealized semantic classifications with little syntactic or contextual detail. The Royal Institute of Thailand (RIT) dictionaries (1950; 1982; 2011) exemplify this limitation, offering minimal elaboration and thus falling outside the scope of the present analysis.

Conversely, other prescriptive models provide more structured—though still limited—approaches to Thai verb classification. These include componential taxonomies (e.g., Premsrirat, 1987), cognitive-semantic models with entailed results (e.g., Thepkanjana & Uehara, 2009), and broader class-based systems

(e.g., Phanthumetha, 2016). While theoretically valuable, such models often lack empirical validation and overlook pragmatic flexibility. For instance, Thepkanjana and Uehara’s (2009) entailed-result model assumes that verbs like *tàt* ‘cut’ or *hàk* ‘break’ lexically encode outcome states. However, their own data (2009, p. 605) present examples in which the intended result does not occur (see Examples 1-2).

- |     |  |            |                        |            |                         |
|-----|--|------------|------------------------|------------|-------------------------|
| (1) | <i>sǒmtɛ̃<sup>h</sup>a:j</i>                                 | <i>hàk</i> | <i>kinmáj</i>          | <i>máj</i> | <i>?ò:k</i>             |
|     | Somchaay   | break      | branch                 | NEG        | exit                    |
|     | ‘Somchaay tried to break the branch, but it was not broken.’ |            |                        |            |                         |
|     |  |            |                        |            |                         |
| (2) | <i>sǒmtɛ̃<sup>h</sup>a:j</i>                                 | <i>tàt</i> | <i>p<sup>h</sup>aː</i> | <i>máj</i> | <i>k<sup>h</sup>à:t</i> |
|     | Somchaay   | cut        | cloth                  | NEG        | torn                    |
|     | ‘Somchaay tried to cut the cloth, but it was not torn.’      |            |                        |            |                         |

These examples show that the result states are not always entailed, as seen in negated constructions where separation fails—challenging lexical entailment and underscoring the role of pragmatic inference.

Along similar lines, Phanthumetha (2016) offers a more extensive lexicographic classification with twelve subclasses (e.g., ‘breaking,’ ‘removing by force’) but similarly lacks empirical grounding or cognitive-pragmatic validation. These limitations are summarized in Table 1, which contrasts major prescriptive models in terms of scope, methodological approach, and theoretical gaps.

**Table 1**
*Contrasting Prescriptive Models of Thai Caused-Separation Verbs*

Study	Scope	Methodological Approach	Limitations
Premsrirat (1987)	Cutting verbs	Componential analysis using features (e.g., instrument, action)	Based heavily on dictionary data; limited speaker-based input
Thepkanjana and Uehara (2007)	<i>hàk</i> ‘break’ (literal + metaphorical senses)	Cognitive semantics; contrastive analysis with Japanese	Narrow lexical scope; metaphorical focus lacks systematic extension
Thepkanjana and Uehara (2009)	Verbs like <i>tàt</i> ‘cut,’ <i>tə<sup>h</sup>.i.k</i> ‘tear,’ <i>hàk</i> ‘break’	Cognitive-semantic approach; entailed-result model	Limited empirical testing; pragmatics underexplored
Phanthumetha (2016)	Broad coverage across 12 subclasses (e.g., breaking, tearing)	Lexicographic taxonomy	No empirical validation; lacks cognitive/pragmatic grounding

In contrast to static, idealized models, recent empirical studies adopt usage-based and experimental approaches. Pichetpan (2021), for example, draws on video-based elicitation with native Thai speakers to show that while prototypical contexts (e.g., using *tàt* for knife-cutting) yield consistent verb choices, less typical scenarios invite lexical variation. Lexical selection in such contexts is shaped by instrument-object affordances, contextual prominence, and prototype effects. Similarly, Thepchuaysuk (2016), and Thepchuaysuk and Thepkanjana (2017) examine syntactic and semantic patterns drawing on structured elicitation data. They classify verbs by transitivity, result expression, and alternation behavior, though their models

remain largely independent of prototype theory and underrepresent pragmatic modulation.

Together, these findings point to the need for an integrated framework that reconciles the structural insights of prescriptive models with the contextual flexibility revealed through empirical data, offering a clearer picture of Thai caused-separation verb usage.

### **1.3 Research Objectives and Questions**

This study compares prescriptive and empirical accounts of Thai caused-separation verbs across three dimensions: (1) syntactic realization, (2) semantic categorization, and (3) verb selection. It addresses two central research questions:

1. How do prescriptive classifications of Thai caused-separation verbs correspond with empirical syntactic structures and semantic distinctions?
2. How do prescriptive patterns align with empirical findings on verb choice and lexical variability?

The study hypothesizes that prescriptive models underrepresent the syntactic and contextual flexibility of Thai caused-separation verbs. Empirical findings are expected to reveal usage-driven patterns marked by lexical overlap, syntactic alternation, and adaptive categorization thereby contributing to semantic modeling and applications in typology, lexicography, and language technology. “Context” here refers to specific featural conditions (e.g., instrument/object affordances) or syntactic requirements that modulate lexical categorization and selection. This framing allows for a clear and testable account of the role of context in structuring verb organization.

### **1.4 Data and Methodological Overview**

To address the research questions, this study draws on both prescriptive and empirical sources. The prescriptive component includes key accounts that define semantic categories and entailment relations (e.g., Premsrirat, 1987; Thepkanjana & Uehara, 2007, 2009; Phanthumetha, 2016), though these remain underexamined against actual usage and offer limited pragmatic

or syntactic insight. The empirical core builds on Pichetpan (2021), which analyzes over 300 verbal responses to video stimuli of caused-separation events (e.g., cutting, smashing), interpreted through prototype theory and statistically clustered to identify usage patterns. Complementary syntactic insights come from Thepchuaysuk (2016), and Thepchuaysuk and Thepkanjana (2017), whose work classifies verbs by argument structure, result expression, and alternation behavior.

This research employs a comparative framework to examine how prescriptive classifications correspond to actual usage, structured across three dimensions (see Table 2).

**Table 2**

*Comparative Dimensions of Prescriptive and Empirical Accounts*

<b>Dimension</b>	<b>Focus</b>	<b>Summary</b>
<b>1. Syntactic Realization</b>	Argument structure and patterns	Examines verb constructions and argument roles in both prescriptive and empirical models.
<b>2. Semantic Categorization</b>	Hierarchy and adaptability	Compares taxonomies with usage-based categorization; considers context, lexical density, and cognitive-cultural influences.
<b>3. Lexical Selection and Usage</b>	Parameters and variation	Analyzes semantics, syntax, and pragmatics in verb choice; contrasts stability and flexibility in prescriptive and empirical frameworks.

These dimensions serve as the basis for evaluating the convergence and divergence between normative classifications and usage-driven verbal behavior in Thai caused-separation events.

Before proceeding, two methodological limitations should be noted. First, the prescriptive sources examined represent a selective sample of Thai verb scholarship and may not capture the full theoretical range. Second, the empirical findings, while based on a robust dataset (Pichetpan, 2021), rely on a single elicitation method and may not reflect the full extent of sociolinguistic variation. Accordingly, the scope of generalization should be understood in light of these constraints.

## 2. Empirical Usage vs. Prescriptive Models in Thai Caused-Separation Verbs

### 2.1 Overview of Empirical Insights

This subsection presents evidence-based insights into the semantic and syntactic properties of Thai caused-separation verbs, primarily from Pichetpan's (2021) work, supplemented by Thepchuaysuk's (2016) findings. Both relied on video-elicited responses from native speakers, though with distinct emphases: Thepchuaysuk focused on constructional variation, while Pichetpan applied statistical tools, such as hierarchical clustering and the Gini-Simpson Index, to model semantic categorization.

Pichetpan (2021) investigated how Thai speakers describe caused-separation events (e.g., cutting, smashing, and tearing) using 43 video scenes adapted from Bohmeyer et al. (2001). Depicting irreversible separation, the stimuli enable controlled yet naturalistic elicitation. The study contributes to cross-linguistic discussions on event typology and lexico-semantic convergence in the Mainland Southeast Asia (MSEA) linguistic area.

Across these scenes, Thai speakers used 24 distinct caused-separation verbs. Nine high-frequency verbs, including *tàt* 'cut,' *t<sup>h</sup>úp* 'smash,' and *hàk* 'snap,' accounted for over 80% of the responses (Pichetpan, 2021, p. 328). Cluster analysis revealed patterned verb usage across scenes (pp. 146, 251), with *tàt* linked to high-predictability separations (e.g., slicing fabric) and *hàk* to lower-predictability actions like snapping twigs (pp. 169, 180).

Four key patterns emerge from Pichetpan’s (2021) findings. First, most scenes elicited only one to three verbs, reflecting both lexical consistency and selective variation (p. 133). Second, six major verb semantic categories were identified: *tàt* ‘cut,’ *hàk* ‘snap,’ *t<sup>h</sup>úp* ‘smash,’ *tɛ<sup>h</sup>i:k* ‘tear,’ *t<sup>h</sup>im* ‘pierce,’ and *kri.t* ‘slit,’ each tied to typical combinations of instrument type, object texture, and motion dynamics (pp. 152, 193-207).

Third, verb selection was guided by instrument type, force intensity, and predictability of separation (pp. 210, 214-215). For instance, sharp-bladed tools with high predictability triggered *tàt* ‘cut,’ while blunt tools and low predictability favored *t<sup>h</sup>úp* ‘smash’ (pp. 73, 171-172, 205). Distinctions between verbs like *fán* ‘hack’ and *hàn* ‘slice’ reflected distinctions in blade orientation and strike angle (pp. 195, 217).

Fourth, syntactically, most utterances conformed to a [SBJ VB OBJ] pattern, aligning with Thepchuaysuk’s (2016, pp. 243-251) agentive/causative type. Semantic nuance was often conveyed through verb choice rather than syntactic variation (Pichetpan, 2021, pp. 135-137). Notably, Thai allows for serial verb constructions to express instrument use, causation, and result. Examples from Pichetpan (2021, p. 136) illustrate the structures:

(3)	<i>p<sup>h</sup>u.jiŋ</i>	<i>tɛ<sup>h</sup>i:k</i>	<i>p<sup>h</sup>a:</i>	<i>ʔɔ:k</i>	<i>tɛà:k</i>	<i>kan</i>
	woman	tear	cloth	exit	ABL	RECIP
	SBJ	VB1	OBJ1	VB2	PREP	PRON
	‘The woman tore the cloth apart.’					

→ Serial verb with directional (*ʔɔ:k* ‘exit’) and reciprocal markers (*tɛà:k kan* ‘from each other’).

(4)	<i>p<sup>h</sup>u.jiŋ</i>	<i>tɛ<sup>h</sup>áj</i>	<i>k<sup>h</sup>ɔ:n</i>	<i>t<sup>h</sup>úp</i>	<i>tɛa:n</i>	<i>tɛ:k</i>
	woman	use	hammer	smash	plate	break
	SBJ	VB1	OBJ1	VB2	OBJ2	VB3
	‘The woman used a hammer to smash the plate.’					

→ Three-verb chain: manipulation (*tɛ<sup>h</sup>áj*), action (*t<sup>h</sup>úp*), result (*tɛ:k*).

(5)	<i>p<sup>h</sup>u.jiŋ</i>	<i>t<sup>h</sup>úp</i>	<i>tɛ<sup>h</sup>ú:ak</i>	<i>háj</i>	<i>k<sup>h</sup>à:t</i>
	woman	smash	rope	CAUS.PURP:GIVE	be.torn
	SBJ	VB1	OBJ	CAUS	VB2

‘The woman struck the rope to break it.’

→ **Resultative construction with *háj* ‘CAUS.PURP:GIVE’ introducing the intended result.**

Such constructions enable expression of intent, instrumentality, and consequence. While simple transitive clauses express action, serial constructions offer precision, especially in less prototypical scenes. Section 2.2.1 explores this further.

Pichetpan (2021) underscores how Thai balances lexical economy with semantic granularity (pp. 133, 337, 344). Core verbs like *tàt* ‘cut’ are commonly used, but context-specific verbs like *tɔ̀w?* ‘puncture’ are preferred when action and instrument properties diverge (p. 343).

In summary, empirical findings reveal that Thai caused-separation verbs reflect instrument, action dynamics, and event structure, shaped by pragmatic reasoning, cognitive salience, and syntactic adaptability. The next subsection evaluates how prescriptive accounts correspond to this usage-driven complexity.

## 2.2 Comparison with Prescriptive Frameworks

This subsection contrasts prescriptive linguistic models with empirical findings on Thai caused-separation verbs. While prescriptive frameworks impose structured classifications and rigid hierarchies, usage-based data suggests a more flexible, context-sensitive system shaped by pragmatic and cognitive factors.

Section 2.2.1 investigates syntactic realization, emphasizing argument structure and alternation patterns. Section 2.2.2 assesses semantic categorization, comparing hierarchical models from Premsrirat (1987) and Phanthumetha (2016)

with usage-based evidence from Pichetpan (2021). Section 2.2.3 examines how grammatical architecture, and contextual dynamics shape lexical selection.

### 2.2.1 Syntactic Realization: Construction Types and Verb Alternation

Traditional Thai prescriptive accounts prioritize lexical semantics over syntax, though Thepkanjana and Uehara (2009) illuminate resultative constructions that link action to outcome, assuming caused-separation verbs entail specific results. Yet empirical work (Pichetpan, 2021) shows these meanings are often pragmatically inferred rather than lexically encoded. Thai causative actions may occur without guaranteed results, as interpretation depends on context and speaker construal. For example, *tàt* ‘cut’ typically implies full separation unless explicitly negated, while *t<sup>h</sup>úp* ‘smash’ varies with object fragility. Such variation reflects Langacker’s (2008) concept of “accommodation,” where meaning arises from context-form interplay (cf. Thepkanjana & Uehara, 2007).

Empirical studies, particularly Pichetpan (2021) and Thepchuaysuk (2016), also underscore Thai’s syntactic richness, with speakers employing both single- and multi-verb constructions. Some of these patterns overlap with those in Thepkanjana and Uehara (2009), as shown below.

#### (6) Single-Verb Construction

<i>som<sup>ǂ</sup>tɛ<sup>h</sup>a:j</i>	<i>tɛ<sup>h</sup>ɪ:k</i>	<i>kradà:t</i>	<i>(p<sup>h</sup>a:jnɔj</i>	<i>nùŋ</i>	<i>tɛ<sup>h</sup>úamo:ŋ)</i>
Somchaay	tear	paper	(in	one	hour)
SBJ/Causer	VB/Verb	OBJ/	Adjunct (optional)		
Undergoer					

‘Somchaay tore the paper (in one hour).’

→ **Single transitive clause with optional temporal adjunct.**

(Source: Adapted from Thepkanjana & Uehara, 2009, p. 596; constructional analysis following Pichetpan, 2021, p. 139 and Thepchuaysuk, 2016, p. 149.)

(7) **Multi-Verb Construction**

sǒmtɕ <sup>h</sup> aj	tɕ <sup>h</sup> i:k	p <sup>h</sup> ǎ:	k <sup>h</sup> ǎ:t
Somchaay	tear	cloth	torn
SBJ/Causer	VB1/Action	OBJ/Undergoer	VB2/Resultative Verb
	Verb		

‘Somchaay tore the cloth (and it was torn).’

→ **Serial verb construction encoding both action and result.**

(Source: Adapted from Thepkanjana & Uehara, 2009, p. 604; analysis following Pichetpan, 2021, p. 140 and Thepchuaysuk, 2016, p. 153.)

Further, Pichetpan identifies two serialization types that encode complex event relations through syntax. *Sequential serialization* links an action verb with a path verb to indicate movement or spatial trajectory, forming a multi-verb construction without subordination (e.g., tɕ<sup>h</sup>i:k ‘tear’ + ɔ̀:k ‘exit’; see Example 3 in Section 2.1). *Causative serialization*, by contrast, uses *hāj* to connect an initiating action to its intended result (e.g., t<sup>h</sup>úp ‘smash’ + *hāj* k<sup>h</sup>ǎ:t ‘cause to tear’; see Example 5 in Section 2.1), with *hāj* functioning as a predicate bridging agent and outcome.

These constructions show how Thai encodes event relations through verb order and functional marker within a multi-verb structure.

Extending this empirical perspective on syntax, Thepchuaysuk (2016; further developed in Thepchuaysuk & Thepkanjana, 2017) classifies caused-separation verbs into four syntactic subtypes, based on argument structure and constructional behavior:

1. **Transitive action verbs** (e.g., *tàt* ‘cut,’ or *hàn* ‘slice’) explicitly encode instrumental actions and appear in single-verb constructions (see Thepchuaysuk, 2016, p. 149).

2. **Non-causative result verbs** (e.g., *tɛ:k* ‘shatter,’ or *kʰà.t* ‘tear’) denote outcomes without causation and occur in biclausal or passive constructions, where result and action are syntactically distinct (see Thepchuaysuk, 2016, pp. 155, 178).

3. **Causative/inchoative alternators** (e.g., *hàk* ‘break’) shift between transitive and intransitive forms depending on agent expression (see Thepchuaysuk, 2016, p. 197).

4. **Action-result serial verbs** pair two verbs in periphrastic constructions to separate action and result, often marking volitional causation (see Thepchuaysuk, 2016, p. 157).

These subtypes show that argument structure and syntactic alternation—not lexical meaning alone—govern verb behavior, challenging the adequacy of purely semantic classifications.

Overall, empirical evidence underscores syntactic flexibility in modeling Thai caused-separation verbs, supporting a usage-based account grounded in real-world patterns.

### 2.2.2 Semantic Categorization: Prescriptive Hierarchies vs. Prototypical Gradience

Thai caused-separation verbs are categorized differently across frameworks. Prescriptive models, such as Premsrirat (1987) and Phanthumetha (2016), define hierarchical taxonomies using componential features like motion, instrumentality, and affectedness, assuming stable semantic boundaries. Premsrirat, for example, uses 89 features to classify Thai cutting verbs into four main classes (pp. 153-159), assigning verbs like *hàn* ‘slice’ to vertical blade and food contexts.

Yet empirical findings challenge such rigidity. Based on video-elicited responses, Pichetpan (2021) shows speakers flexibly categorize verbs based on tool type, force, and object properties. Verb overlaps suggest

fluid boundaries and transitional zones, reflecting prototypicality consistent with Rosch and Mervis’s (1975) prototype theory. This theory posits that categories are structured around central members, with less typical ones still included despite sharing fewer defining features. In Thai, verbs like *tàt* ‘cut’ (for blade-based separations) form lexical cores, while transitional verbs like *fan* ‘hack’ and *tə̀w?* ‘puncture’ blur boundaries. This supports a gradient, usage-based model of lexical categorization, as illustrated in Figure 1.

**Figure 1**

*Semantic overlap among *tàt* ‘cut,’ *tʰúp* ‘smash,’ and *tʰim* ‘pierce,’ with transitional verbs indicating blurred category boundaries (Pichetpan, 2021).*

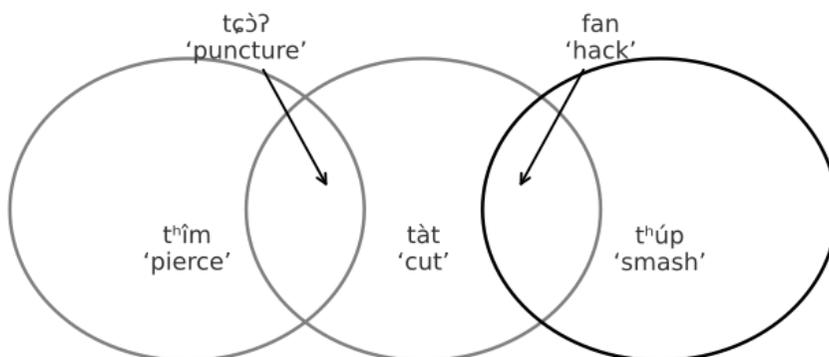


Table 3 below complements this by comparing prescriptive categories with observed usage patterns (Pichetpan, 2021), showing how contextual factors, such as tool ambiguity, force, and separation type, guide verb interpretation. Rather than reflecting fixed boundaries, categorization in Thai adapts to perceptual and functional cues, highlighting the role of pragmatic reasoning.

**Table 3**

*Prescriptive Classifications and Empirical Overlap Zones Among Thai Caused-Separation Verbs: *tàt* ‘cut,’ *t<sup>h</sup>úp* ‘smash,’ and *t<sup>h</sup>ím* ‘pierce.’*

Verb	Prescriptive Class	Empirical Overlap	Observed Contextual Triggers
<i>tàt</i> ‘cut’	<p><i>Material Destruction (Cutting)</i></p> <ul style="list-style-type: none"> <li>• Premsrirat (1987, p. 157): use of sharp tools; excludes food.</li> <li>• Phanthumetha (2016): under “separating some parts.”</li> </ul>	<p>Overlaps with <i>smash</i> or <i>pierce</i> in cases of tool ambiguity or partial separation</p>	<ul style="list-style-type: none"> <li>• Cutting scenes with striking motions with assumed blades (e.g., karate-chopping a carrot) may be variably described as <i>fan</i> ‘hack’ or <i>t<sup>h</sup>úp</i> ‘smash,’ depending on tool shape (blade vs. blunt edge) and impact intensity.</li> <li>• Peripheral cutting scenes like puncturing with pointed objects (e.g., twigs), deviate from <i>tàt</i> ‘cut’ by yielding only partial separation. The verb <i>tǎw?</i> ‘puncture’ blends the separative feature of <i>tàt</i> ‘cut’ with the vertical thrust of <i>t<sup>h</sup>ím</i> ‘pierce,’ thus blurring category boundaries.</li> </ul>
<i>t<sup>h</sup>úp</i> ‘smash’	<p><i>Breaking (Smashing)</i></p> <ul style="list-style-type: none"> <li>• Phanthumetha (2016): striking with weighted round/rod tools, possibly under “breaking,” (p. 927).</li> </ul>	<p>Overlaps with cutting in strong downward actions using ambiguous tool edges.</p>	<p>Scenes involving fast, downward impact with an indeterminate tool may shift interpretation from smashing to blade-referential cutting.</p>

**Table 3**

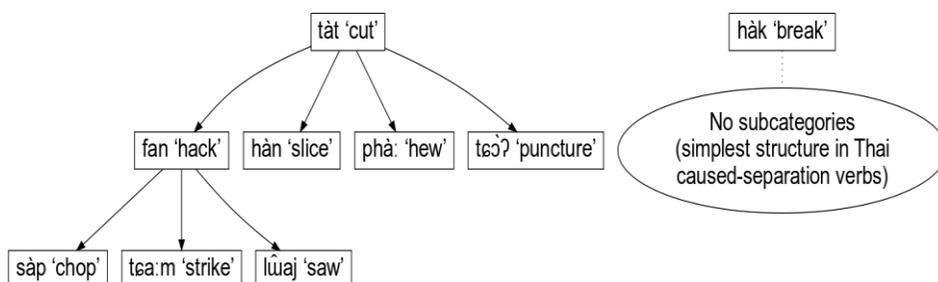
*Prescriptive Classifications and Empirical Overlap Zones Among Thai Caused-Separation Verbs: tăt ‘cut,’ t<sup>h</sup>úp ‘smash,’ and t<sup>h</sup>ím ‘pierce.’ (cont.)*

<b>t<sup>h</sup>ím</b> ‘pierce’	<i>Material Destruction (Piercing)</i> • Phanthumetha (2016): emphasizes destruction and state change; under “separating some parts.”	Overlaps with cutting in localized punctures resembling blade use.	Piercing actions may be construed as cutting when the resulting separation resembles a sharp incision, linking them to the cutting domain despite involving pointed tools and incomplete division.
------------------------------------	--	--	--

In addition to semantic overlaps, Pichetpan (2021) examines variation in lexical granularity. Some verbs form detailed subcategories while others span broader and flatter semantic fields. For instance, *tăt* ‘cut’ has rich subcategories depending on instrument and target, whereas *hàk* ‘break/snap’ is semantically broader. Figure 2 contrasts these structures.

**Figure 2**

*Comparison of hierarchical subcategories of tăt ‘cut’ vs. the flatter structure of hàk ‘break.’*



These findings question the viability of a uniform componential taxonomy across all verbs within the domain. Variability suggests that verb

categories emerge from usage patterns, familiarity, and communicative salience (cf. Kopecka, 2012); consequently, some develop more nuanced semantic networks while others remain broader in scope.

Furthermore, adaptability also extends to syntax. Thepchuaysuk (2016) identifies eight syntactic alternations, later collapsed into four types based on transitivity and result encoding (pp. 114, 280). His findings show that verb behavior depends not only on semantic content but also on argument structure and syntactic flexibility. A verb's semantic classification shifts based on whether it appears in transitive or inchoative constructions, challenging the idea that syntax merely reflects meaning.

Cultural and cognitive dimensions affect verb categorization as well. Prescriptive models, such as Premsrirat (1987), incorporate culturally salient domains like food preparation into taxonomic groupings. This approach reflects cultural relevance but presumes category stability, prioritizing formality over nuance. In contrast, empirical approaches foreground socio-cognitive variation. Pichetpan (2021, p. 375) emphasizes speaker familiarity, communicative goals, and contextual salience in verb categorization. Similarly, Thepchuaysuk (2016, p. 24) stresses individual cognitive interpretation in labeling separation events, challenging rigid classification.

Together, these contrasts underscore the tension between structured taxonomies and adaptive usage patterns. Prescriptive models offer a stable classification, potentially useful in lexicography and pedagogy, but often lack sensitivity to speaker variation and pragmatic nuance. Empirical models, while more reflective of real-world usage, complicate systematic categorization by introducing dimensions like syntactic alternation, speaker intention, and contextual framing.

In conclusion, the analysis of Thai caused-separation verbs reveals that categorization extends beyond semantic features, involving syntactic structure, context, and cognition. Integrating prescriptive and empirical approaches can strengthen Thai verb categorization. A probabilistic, hierarchical

hybrid framework could balance structured classification with usage fluidity, improving applications in language teaching, lexicography, and natural language processing

### **2.2.3 Determinants of Verb Choice: Semantic, Syntactic, Contextual, and Cognitive Dimensions**

Thai caused-separation verb choice reflects the interplay of semantics, grammar, and context, examined here across three dimensions: (1) selection parameters, (2) contextual and cognitive influences, and (3) usage stability versus flexibility.

#### **1) Lexical Semantics and Constructional Constraints**

Prescriptive accounts classify verbs hierarchically using componential features such as action type, force, and object properties. For instance, Premsrirat (1987, p. 157) differentiates *tàt* ('cut with a sharp instrument.') from *sakàt* ('chisel stone or hardwood.') based on tool-action pairings, while Phanthumetha (2016) focuses on resulting physical states.

Empirical studies extend this by showing that verb choice is also shaped by syntactic factors. Thepchuaysuk and Thepkanjana (2017, pp. 310-311) group verbs by transitivity and resultativity, highlighting how argument structure governs selection. Unlike prescriptive models (see Phanthumetha, 2016; Premsrirat, 1987), which blur distinctions between action- and result-oriented verbs, empirical data show that action verbs appear in transitives to emphasize instrument use, while result verbs surface in inchoative or passive forms (Thepchuaysuk, 2016).

In sum, although prescriptive models offer structured categories, empirical research reveals verb selection as a dynamic interface between semantics and syntax.

## 2) Contextual Modulation and Speaker Intent

Prescriptive models center on semantic features in verb categorization, often sidelining context and speaker variation. Although variability is acknowledged (e.g., Premsrirat, 1987, p. 146), such models lack formal treatment of contextual effects.

Empirical studies, by contrast, foreground how verb choice reflects dynamic, context-driven reasoning. Pichetpan (2021, p. 112), using video elicitation (Bohnenmeyer et al., 2001), finds that Thai speakers adjust verb use based on event framing, tool perception, and communicative goals. For example, a knife-hand gesture applied to a rigid carrot elicits *t<sup>h</sup>úp* ‘smash’ rather than *tát* ‘cut,’ likely due to how the hand is construed (p. 181). Similarly, *hàk* ‘break’ shifts meaning with object type (Thepkanjana and Uehara, 2007). Such patterns support a probabilistic selection model governed by perceptual cues, instrument-object affordances, and pragmatic framing.

Additional pragmatic considerations, including speaker familiarity and intent, also influence verb choice, aligning with cognitive linguistic accounts (Pichetpan, 2021, pp. 134, 184; Taylor, 2003). Event prototypicality further affects usage, with high-frequency verbs suiting prototypical scenarios, while unconventional events prompting lexical variability (Pichetpan, 2021, p. 43; Thepchuaysuk & Thepkanjana, 2017, p. 300).

Table 4 summarizes how contextual, pragmatic, and prototypical factors, such as framing, tool salience, object properties, syntax, and speaker intent, modulate verb choice.

**Table 4**

*Contextual, Pragmatic, and Prototypical Determinants of Verb Choice in Thai Caused-Separation Events*

Factor Type	Influence Factor	Potential Effect on Verb Choice
Contextual	Event Framing	Guides emphasis on instrumental action ( <i>transitives</i> ) vs. result ( <i>intransitives</i> ), aligning with verb type distinctions (Thepchuaysuk, 2016), and Thai’s <i>multi-verb sequencing</i> of subevents (e.g., <i>instrument use</i> → <i>separation</i> → <i>result</i> ) (Pichetpan, 2021).
	Instrument Saliency	Favors instrument-specific over general verbs, reflecting Thai categorization by tool type (Pichetpan, 2021).
	Object Properties	Predicts verbs based on object rigidity, texture, or dimensionality, reflecting Thai’s object classification (rigid vs. flexible) (Pichetpan, 2021).
	Syntactic Frame	Regulates alternations through <i>transitivity</i> (e.g., <i>causative/inchoative alternation</i> for verbs like <i>hàk</i> ‘break’ or <i>tɛ̌ȟ:k</i> ‘tear’; Thepchuaysuk, 2016) and permits <i>serial verbs</i> under <i>strict combinatorial rules</i> (e.g., <i>hâj</i> ‘CAUS:GIVE’ needs non-agentive, non-manner V2; Pichetpan, 2021; Thepchuaysuk, 2016).
Pragmatic	Speaker Intent	Shifts between <i>agentive manipulation</i> (e.g., <i>instrument-focused verbs</i> like <i>tɛ̌ȟ:áj</i> ‘use’) and <i>resulting state</i> (e.g., <i>causative resultatives</i> use <i>hâj</i> ‘CAUS.PURP:GIVE’ to signal achieved separation) (Pichetpan, 2021).

**Table 4**

*Contextual, Pragmatic, and Prototypical Determinants of Verb Choice in Thai Caused-Separation Events (cont.)*

Factor Type	Influence Factor	Potential Effect on Verb Choice
Prototypicality	Cognitive Centrality	Promotes stable verbs in familiar contexts (e.g., <i>cutting with a blade, snapping rigid objects</i> ) due to cognitive salience, whereas atypical events (e.g., <i>chisel use</i> ) show greater variation, reflecting weaker representation and flexible categorization (Pichetpan, 2021).

These findings suggest that verb selection is not semantically fixed but arises from the interaction of context, cognition, and speaker goals.

### 3) Usage Patterns: Verb Stability vs. Contextual Flexibility

Prescriptive studies emphasize lexical stability, grouping verbs based on assumed regularity. For example, Premsrirat (1987) places Thai cutting verbs like *tɕ<sup>h</sup>uān* ‘slice’ and *tɕuāk* ‘hack’ under the component ‘rapid,’ suggesting fixed use in swift motion contexts (p. 161).

Empirical research, however, challenges this fixity, showing Thai speakers adapt verb usage according to context, syntax, and pragmatics. Pichetpan (2021) documents uses of *tɕɔʔ* ‘puncture’ for both ‘poking cloth’ and ‘chopping a carrot with a chisel,’ reflecting flexibility driven by instrumental salience. Thepchuaysuk (2016), and Thepchuaysuk and Thepkanjana (2017) affirm this variability, showing that construction types may influence verb use, as instrument- and result-focused frames tend to promote consistency, but also allow for alternation through passive or causative/inchoative forms. Pichetpan (2021) adds that speaker intent and event salience further modulate usage (pp. 130-135).

Notably, even prescriptive models admit some flexibility. Thepkanjana and Uehara (2009) note that Thai causative predicates distinguish between process and result, unlike English; for instance, *tàt* ‘cut’ typically implies full separation but can be canceled in serial constructions (p. 605). Despite being framed prescriptively, this pattern aligns with empirical observations.

In sum, while prescriptive frameworks stress coherent verb classification and selection, empirical evidence reveals Thai caused-separation verbs as both patterned and adaptable. A hybrid model that fuses structural expectations with context-driven flexibility can offer a more balanced view of Thai verb use.

### 3. Conclusion and Future Directions

#### 3.1 Synthesis and Implications

This subsection synthesizes prescriptive and empirical perspectives on Thai caused-separation verbs, arguing that rigid taxonomic models are insufficient. A probabilistic, usage-based framework that integrates structural, semantic, and pragmatic dimensions more effectively captures Thai’s intricate patterns of verb selection, serial constructions, and context-driven variation.

As shown in Section 2.2.1, prescriptive models privileges fixed and limited classifications, often underrepresenting Thai’s syntactic variability and constructional variety. Empirical studies, by contrast, illuminate syntactic adaptability, including serial verb constructions and alternations across transitive, inchoative, and resultative forms. These findings underscore the importance of accounting for syntactic flexibility in modeling verb behavior. Incorporating such variation enhances both theoretical precision and pedagogical relevance.

The findings in Section 2.2.2 show that prescriptive semantic frameworks based on verb-instrument pairings and resultant states tend to impose categorical boundaries. However, empirical evidence reveals that real-world usage involves overlapping and gradient patterns shaped by speaker

construal and event-specific parameters. A probabilistic-hierarchical model that integrates structured generalizations with semantic fluidity could offer a more accurate and flexible account, especially for transitional verbs that span multiple event types.

As noted in Section 2.2.3, prescriptive approaches often base Thai verb selection on semantic criteria and tool-event pairings. Yet empirical studies suggest that speakers also consider contextual salience, speaker intent, and event prototypicality—adjusting verb usage to align with communicative goals and situational nuance. This supports models that accommodate variation and context sensitivity.

These insights hold applied relevance: in language teaching, usage-based models can expose learners to realistic patterns of verb alternation and variation; in lexicography, dictionary entries should reflect actual usage frequency and category overlap; and in Natural Language Processing (NLP), probabilistic models better handle ambiguity and context-specific interpretation.

Ultimately, combining the structural clarity of prescriptive models with empirical attention to context and cognition yields a more robust and practical account of Thai caused-separation verbs.

### 3.2 Research Directions

To advance Thai caused-separation verb classification and selection, future research should move beyond fixed semantic-syntactic classifications to explore how meaning is constructed in context. This includes examining discourse structure, interactional dynamics, and register variation, with attention to features like formality, information structure, and speaker intent.

**Corpus-based probabilistic modeling** provides a solid methodological base. Large-scale corpora can be used to track frequency distributions, syntactic patterns, and collocations (e.g., agent-theme-instrument configurations). Such patterns support probabilistic frameworks that bridge structured taxonomies to real-world variation shaped by gradient behavior, alternation,

and cross-category usage. Computational techniques such as Hidden Markov Models (HMMs) and Conditional Random Fields (CRFs) can model verb alternation and predict context-sensitive lexical choices (Lafferty et al., 2001; Rabiner, 1989).

**Field-based and discourse-oriented methods** are equally valuable. Techniques such as structured elicitation, narrative tasks, and spontaneous conversation yield naturalistic data essential for refining usage-based models.

**Socio-cognitive approaches** should also be pursued, as factors like speaker familiarity, event salience, and contextual predictability influence real-time verb selection, which can be examined through psycholinguistic tools such as reaction-time experiments and self-paced reading.

Taken together, integrating perspectives on usage, cognitive patterns, and probabilistic trends can help us better understand Thai caused-separation verbs and underscore their significance across both theoretical linguistics and practical language applications.

### References

- Berlin, B., & Kay, P. (1969). *Basic color terms: Their universality and evolution*. University of California Press.
- Bohnemeyer, J., Bowerman, M., & Brown, P. (2001). Cut and break clips. In S. C. Levinson & N. J. Enfield (Eds.), *Field Manual 2001, Language and Cognition Group* (pp. 90–96). Max Planck Institute for Psycholinguistics.
- Du, J., Li, F. T., Ge, Y., & Zhang, J. (2023). Separation events in Mandarin, Russian and Korean: A crosslinguistic event-categorization study. *Review of Cognitive Linguistics*, 21(2), 377-410.
- Ekman, P. (1972). *Emotion in the human face: Guidelines for research and an integration of findings*. Pergamon Press.

- Goodenough, W. H. (1965). Yankee kinship terminology: A problem in componential analysis. *American Anthropologist*, 67(5), 259-287.
- Hale, K., & Keyser, S. J. (1987). *A view from the middle* (Lexicon Project Working Papers No. 10). MIT, Center for Cognitive Science.
- Hsiao, F.-Y. (2015). Encoding and categorizing cutting and breaking events in Mandarin Chinese. *Journal of East Asian Linguistics*, 24(1), 55-100.
- Kopecka, A. (2012). Semantic variation and universals in the categorization of cutting and breaking events: A study of Polish. *Cognitive Linguistics*, 23(2), 281-318.
- Lafferty, J., McCallum, A., & Pereira, F. (2001). Conditional random fields: Probabilistic models for segmenting and labeling sequence data. In C. E. Brodley & A. Pohoreckyj Danyluk (Eds.), *Proceedings of the 18<sup>th</sup> International Conference on Machine Learning* (pp. 282-289). Morgan Kaufmann.
- Langacker, R. W. (2008). *Cognitive grammar: A basic introduction*. Oxford University Press.
- Majid, A., Bowerman, M., van Staden, M. V., & Boster, J. S. (2008). The cross-linguistic categorization of everyday events: A study of cutting and breaking. *Cognition*, 109(2), 235-250.
- Majid, A., Bowerman, M., van Staden, M., & Boster, J. S. (2007a). The semantic categories of cutting and breaking events: A crosslinguistic perspective. *Cognitive Linguistics*, 18(2), 133-152. <https://doi.org/10.1515/COG.2007.005>
- Majid, A., Gullberg, M., van Staden, M., & Bowerman, M. (2007b). How similar are semantic categories in closely related languages? A comparison of cutting and breaking in four Germanic languages. *Cognitive Linguistics*, 18(2), 179-194. <https://doi.org/10.1515/COG.2007.007>
- Murphy, G. L. (2002). *The big book of concepts*. MIT Press.
- Papafragou, A., & Selimis, S. (2010). Event categorisation and language: A cross-linguistic study of motion. *Language and Cognitive Processes*, 25(2), 224-260.

- Phanthumetha, N. (2016). *คลังคำ: คู่มือการใช้ภาษาไทยแนวใหม่ [Thai word bank: A new guide to using the Thai language]*. Amarin Printing and Publishing.
- Pichetpan, N. (2021). *Verbs of caused-separation in Thai and Khmer: Lexical semantics and language convergence in Mainland Southeast Asia* [Doctoral dissertation, University of Sydney]. Sydney eScholarship Repository. <https://ses.library.usyd.edu.au/handle/2123/25661>
- Premssirat, S. (1987). A study of Thai and Khmu cutting words. In D. Bradley (Ed.), *Papers in South-East Asian Linguistics No. 10: Khmu, a minority language of Thailand* (pp. 145-187). Pacific Linguistics, Series C.
- Pye, C., Loeb, D. F., & Piao, Y.-Y. (1995). The acquisition of breaking and cutting. In E. Clark (Ed.), *Proceedings of the 27th Annual Child Language Research Forum* (pp. 27-36). CSLI Publications.
- Rabiner, L. R. (1989). A tutorial on hidden Markov models and selected applications in speech recognition. *Proceedings of the IEEE*, 77(2), 257-286.
- Rosch, E., & Mervis, C. B. (1975). Family resemblances: Studies in the internal structure of categories. *Cognitive Psychology*, 7(4), 573-605.
- Royal Institute of Thailand. (1950). *พจนานุกรม ฉบับราชบัณฑิตยสถาน พ.ศ. 2493 [Royal Institute Dictionary, B.E. 2493]*. Royal Institute of Thailand.
- Royal Institute of Thailand. (1982). *พจนานุกรม ฉบับราชบัณฑิตยสถาน พ.ศ. 2525 [Royal Institute Dictionary, B.E. 2525]*. Royal Institute of Thailand.
- Royal Institute of Thailand. (2011). *พจนานุกรม ฉบับราชบัณฑิตยสถาน พ.ศ. 2554 [Royal Institute Dictionary, B.E. 2554]*. Royal Institute of Thailand.
- Saji, N., Hong, C., & Wang, C. (2024). Learning semantic categories of L2 verbs: The case of cutting and breaking verbs. *PLOS ONE*, 19(1), e0296628. <https://doi.org/10.1371/journal.pone.0296628>

- Sudmuk, C. (2005). *The syntax and semantics of serial verb constructions in Thai* [Doctoral dissertation, The University of Texas at Austin]. ProQuest Dissertations & Theses Global.
- Talmy, L. (1985). Lexicalization patterns: Semantic structure in lexical forms. In T. Shopen (Ed.), *Language Typology and Syntactic Description* (Vol. 3, pp. 57-149). Cambridge University Press.
- Taylor, J. R. (2003). *Linguistic categorization* (3<sup>rd</sup> ed.). Oxford University Press. <https://doi.org/10.1093/oso/9780199266647.001.0001>
- Thepchuaysuk, K. (2016). *คุณสมบัติทางวากยสัมพันธ์และอรรถศาสตร์ของคำกริยาแสดงเหตุการณ์แยกส่วนในภาษาไทย [Syntactic and semantic properties of verbs of separation in Thai]* [Doctoral dissertation, Chulalongkorn University]. Chulalongkorn University Intellectual Repository. <https://cuir.car.chula.ac.th/bitstream/123456789/55288/1/5380104422.pdf>
- Thepchuaysuk, K., & Thepkanjana, K. (2017). คุณสมบัติทางวากยสัมพันธ์และอรรถศาสตร์ของคำกริยาแสดงเหตุการณ์การแยกส่วนภาษาไทย [Syntactic and semantic properties of verbs of separation in Thai]. *Journal of Studies in the Field of Humanities*, 24(2), 278-317.
- Thepkanjana, K. & Uehara, S. (2009). Resultative constructions with “implied-result” and “entailed-result” verbs in Thai and English: A contrastive study. *Linguistics*, 47(3), 589-618. <https://doi.org/10.1515/LING.2009.020>
- Thepkanjana, K., & Uehara, S. (2007). Semantic extension of the verb of breaking in Thai and Japanese. *Manusya: Journal of Humanities*, 10(3), 95–114.
- Wallace, A. F. C., & Atkins, J. (1960). The meaning of kinship terms. *American Anthropologist*, 62(1), 58-80. <https://doi.org/10.1525/aa.1960.62.1.02a00040>