The Tai-Kadai Languages

Anthony V. N. Diller, Jerold A. Edmondson, Yongxian Luo

Resources for Thai Language Research

Publication details
Anthony Diller
Published online on: 11 Jun 2008

How to cite :- Anthony Diller. 11 Jun 2008, Resources for Thai Language Research from: The Tai-Kadai Languages Routledge
Accessed on: 21 Dec 2018

PLEASE SCROLL DOWN FOR DOCUMENT

Full terms and conditions of use: https://www.routledgehandbooks.com/legal-notices/terms

This Document PDF may be used for research, teaching and private study purposes. Any substantial or systematic reproductions, re-distribution, re-selling, loan or sub-licensing, systematic supply or distribution in any form to anyone is expressly forbidden.

The publisher does not give any warranty express or implied or make any representation that the contents will be complete or accurate or up to date. The publisher shall not be liable for an loss, actions, claims, proceedings, demand or costs or damages whatsoever or howsoever caused arising directly or indirectly in connection with or arising out of the use of this material.
PART 2

TAI LANGUAGES: OVERVIEWS AND RESOURCES
CHAPTER THREE

RESOURCES FOR THAI LANGUAGE RESEARCH

Anthony Diller

3.1. INTRODUCTION

3.1.1. Scope

Thai is spoken, at least as a second variety, by well over half of the total of 80 or 90 million speakers of Tai-Kadai languages. In some respects it reflects features of the greater grouping as a whole, but in other ways it is exceptional. Thai is also by far the most thoroughly described member of the group, with accounts going back several centuries. The purpose of this chapter is to call attention to some five hundred studies of Thai grammar and other aspects of the language. This is no means a comprehensive linguistic bibliography, which would need several times as many entries. Studies are selected here because they are representative of ongoing research and because they are useful in providing readers with further bibliography. Some attention is also given to how Thai linguistic research and its subfields have developed historically, including how studies cited relate to broader professional background issues, which also may shift diachronically.

Omitted below are many studies of merit, especially those in languages other than English. In particular, books, theses and journal articles written in Thai have generally not been included even though they contain innovative and revealing linguistic research on the language. Those seriously interested in Thai and in other Tai-Kadai languages will surely need to spend time in Thai university libraries and elsewhere where these illuminating materials are accessible.

The main focus here is on Thai; only a sample of work on Tai, Kam-Tai and Tai-Kadai is included. For further references on the wider family at its different levels, see other chapters in this volume and Huffman (1986a); anthologies edited by Gething, et al. (1976); by Khunying Suriya Ratanakul, et al. (1985, 1998); by Edmondson and Solnit (1988, 1997); and works of Morev (1991); Edmondson (2007); Luo (1997, 2007); Diller (2000); Ostapirat (2000, 2004); Thurgood (2007a). Only selected earlier works of Li, Haudricourt, Gedney, and other pioneers in the field are mentioned here; see Huffman (1986a) for fuller listings.

Thai authors are cited below by family (last) name, with some reluctance. Apologies are due to those who prefer given (first) name citation. As a statistical tendency, last-name citation has been the majority practice among Thai scholars writing for an international linguistics audience, so that is followed here. Such practice is also the norm in scientific and medical writing. Beware however that this contrasts with the humanities where Thai authorities are frequently cited and indexed by given name, a format also followed by many libraries, by Huffman (1986a), and in some other reference works.

Where relevant, several works are mentioned in more than one subsection below for convenience of readers with specific interests. Apologies to readers who find this repetition tedious.

3.1.2. Bibliographies and anthologies

The comprehensive indexical bibliography of Huffman (1986a) includes many works not mentioned here. Bibliographies of Kullavanijaya and Vongvipanond (1984) and of Burusphat (2002) are also of utility.

Many useful papers appear in anthologies and conference proceedings. Of great value and
convenience to Thai linguists are collections of papers presented to conferences of the Southeast Linguistics Society (SEALS) and to symposia in the Pan-Asiatic Linguistics series (PAL; see Luksaneeyanawin, et al., 1992; Premsrirat, et al., 1996) and similar symposia (Bamroongraks, et al., 1988). Other anthologies with a strong Thai grammar focus are edited by Bickner, et al. (1986); Abramson (1997); Tingsabadh and Abramson (2001); Harris, Burusphat and Harris (2007). For earlier anthologies, see Huffman (1986a). There is also treatment of Thai in sources where focus is more broadly Tai-Kadai or Southeast Asian (e.g. Ratanakul, Thomas and Premsrirat 1985).

Ongoing resources to keep in mind are the journals *Mon-Khmer Studies*, *Crossroads*, *Linguistics of the Tibeto-Burman Area*, *Journal of the Siam Society* and *Tai Culture*, as well as publications shown on the website of *Pacific Linguistics*. Journals produced in the Thai university context publish linguistic studies of quality in Thai and occasionally in English, among them: *Journal of Language and Linguistics, Science of Language Papers, Journal of Language and Culture*.

### 3.1.3. Grammars and overviews

A cogent reason to reign in representation of Thai in this volume is the recent publication by Iwasaki and Ingkaphirom (2005) of a splendid reference grammar of the language. Questions linguists may have about the language’s structures and functions will be answered and discussed in this comprehensive and well-indexed volume, with its many examples taken from naturally-occurring speech.

To place Thai grammar in its wider Tai context, I am aware of only one extensive comparative grammar of the Tai languages: the Russian work of Morev (1991). This is a work of insight, fine scholarship and includes an impressive bibliography of Russian sources. Perhaps subsequent research, including that reported in this volume, will stimulate updated comparative Tai or Tai-Kadai grammars.

The earlier reference grammar of Noss (1964), a standard for decades, retains its value with excellent examples and a good index, but today’s readers may need some patience in matters of terminology and orientation. Panupong (1970), developing an initiating/non-initiating distinction, presents an impressive study both of sentence-level syntax and of inter-sentence relations. Her study remains an important milestone for those pursuing structurally-defined relationships. In a rather different linguistic tradition, the Thai grammar (in Russian) of Morev, Plam and Fomicheva (1961) is an earlier landmark analysis also worthy of study. Among grammars written in Thai, my favorite remains Bandhumedha (1979), full of fresh ideas.

For the general reader more comfortable with categories of traditional grammar, a range of pedagogical grammars is available, among which Smyth (2002) is particularly informative, clear and dependable. For a more philological, historical and anthropological account of Thai, consult Anuman Rajadhon (1961, 1981).

For overviews, Haas (1969b) and Li (1974) present concise encyclopedic articles of admirable compactness and lucidity. For more specialized summaries, see Gedney (1967), reviewing Thai research up to that date. More recent overviews, covering the basics of phonology and syntax, include those of Hudak (1987), Bickner (1994), and Diller (2004). Premsrirat (2006) presents a useful sociolinguistic summary.

### 3.2. PHONETICS, PHONOLOGY AND ORTHOGRAPHY

#### 3.2.1. Distinctive units

The consensus of many authorities would admit the following distinctive (phonemic) units in Thai, although transcription systems vary and individual scholars might have alternative views. Prevoiced and lax stops: b, d; voiceless, unaspirated and tense stops: p, t, c, k (‘tense’ here perhaps implying also a simultaneous glottal and oral release; c being alveopalatal and affricated); voiceless aspirated stops: ph, th, ch, kh (ch affricated); nasals: m, n, ng [ŋ]; fricatives and aspirates: f, s, h; liquids: r, l (distinguished rather tenuously, with sociolinguistic
tendencies towards merger and overcorrection); semivowels w, y [j]. The majority, but not all, scholars in the field admit glottal stop to the consonant inventory but some opt for an analysis in which it becomes predictable; it is generally not shown in the transcription used here. Some consonants show a range of articulations, e.g. as sociolinguistically conditioned. Several initial clusters with second component -r, -l, or -w are permitted but are often simplified in less educated or less careful colloquial speech.

Only voiceless unaspirated stops -p, -t, -k, glottal stop, nasals and semivowels occur in final position. In some systems, final semivowel equivalents are indicated by vocalic diphthongs: [kay^11] ‘chicken’ may appear as [kaj^11] or [kai^11]. The final stops are unreleased, lacking the tense quality of initial p-, t-, k-. The widely-used Haas transcription represents stop finals as -b, -d, -g, but Abramson’s (1972) instrumental studies do not show voicing. Loanwords increasingly introduce finals like -s, -f, etc.

For vowels, nine come in short-long pairs: i i:, u u:, e e:, oe, oe: (mid central or slightly back unrounded [ə] or [ɨ]); o, oː; ae, aeː; a, aː; and o oː: (low back unrounded [ɔ]); three diphthongs ia, u’a and ua occur as phonologically long, with short variants found in a few exclamations or in other marginal vocabulary items. Long vowels can occur finally, as can short vowels plus glottal stop.

The five tones are usually described as mid (33), low (11), falling (52), high (45, 55 or 454) and rising (24). For most speakers, the high tone includes glottal constriction, more salient when vowel is long. (Tone is indicated here in superscripts of these paired numbers, in most cases with lexical rather than phonetic values, e.g. X^45 represents a contrastive high tone, regardless of its real pitch characteristics; unmarked syllables have so-called neutral tone.)

The preceding phonemic inventory is shown in a romanization differing only in minor respects from the semi-official system prescribed by the Thai Royal Institute. Other systems are encountered, but most approach a one-to-one correspondence with the semi-official system sketched above, including that of Haas (1964, 1969b). Workers in the field soon become used to variation in transcription systems, often merely a matter of symbols used but occasionally indicating more substantial claims as to phonetic detail. More on the phonetic nature of these units and how they are represented follows below.

For general discussion of the Thai inventory, see Tingsabadh and Abramson (1993a, b) and the recent grammars and encyclopedic reviews mentioned above.

Of historical interest is Bradley (1911), apparently the first instrumental analysis of the five Thai phonemic tones, although tones were noted much earlier: La Loubère (1691) counted six. Abramson (1962) marks the professional dawn of modern instrumental research in Thai acoustic phonetics and work of the highest quality by Abramson and colleagues has extended over nearly half a century.

3.2.2. Phonological approaches and overviews

For an overview of studies of Thai phonology up to the mid 1980s, see Rischel (1984). Over several decades, work of L-Thongkum, Lukaneeyanawin, Sutadarat, and other Thai colleagues has contributed significantly to progress in Thai phonetics and phonology and those researching these topics would surely need to attend to the full range of their work. Original phonological approaches are developed in these publications, e.g. Lukaneeyawin (1992) explains ‘three-dimensional phonology’. Note also Erikson’s and Gandour’s phonological contributions. Gandour’s work, together with colleagues, spans several decades and constitutes the principal body of neurolinguistic research analyzing how Thai language ability is affected under aphasic or other degenerative neurological conditions.

Redundancy rules constrain tone by patterns of vowel length and whether or not a syllable ends in a stop (referred to as ‘closed’ or ‘open’, or in more picturesque Thai-derived terms as ‘dead’ or ‘alive’). Closed syllables occur only with low, falling and high tones. There are also some sandhi-like rules, e.g. tones shift from the values above in fast speech, with some rising tones becoming high; the long vowels in first syllables of certain compounds are shortened. Example: [khaoŋ^53-lang^24] ‘behind’, spelling pronunciation; cp. [khaŋ^5-lang^24], normal speech. Such rules and debate over some points in the preceding summary are considered in
the classic study of Henderson (1949) and in Lodge (1986); also in work of many authorities noted below.

3.2.3. Consonants, vowels and tones

Work of Jimmy G. Harris (e.g. 1972, 1987, 2007) in articulatory phonetics analyzes specifics of the Thai sound system. Exact places and manners of articulation, e.g. of the units transcribed here as [d], [t], [th], [c] and [ch], are established through palatography. Harris provides useful comparisons with other languages and shows how identification of Thai phonetic phenomena has been oversimplified. In an earlier study, Brown (1962, 1965) also makes challenging observations regarding Thai consonant articulations, emphasizing complex articulations and their tonal effects.

Vowels and questions of vowel length are studied from various perspectives in works by Brown, L-Thongkum, Hartmann, Roengpitya, Svastikula and Tumtavitikul. Onsuwan (2000) inspects the stop/vowel interface. Abramson (2001) considers the stability of distinctive vowel length. For pharyngealization, see Henderson (1987); for nasalization, see Beddor, et al., (1999); for states of the glottis such as prephonation and unprephonation, see work of J.G. Harris. Diachronic studies of vowel development are mentioned in 4.5.

Tone and in particular how tone interacts with other phonological elements has stimulated much research in Thai phonetics and phonology. While most authorities concur that citation forms in Thai show five lexical tones, they are also aware that citation values shift in various contexts, leading to different phonological perspectives. In an early study, Leben (1971) argued for a segmental approach to Thai tone. Yip (1982) counters this with a laryngeal tier analysis. An autosegmental approach is outlined by Hoonchamlong (1990). A general overview of Thai tonal issues is given by Erickson (1976), emphasising physiology and by Intrasai (2001), emphasising acoustics; see also Gandour (1976); Robertson (1982). Tone and vowel length are considered by Gandour (1977, 1984), and in a wider comparative context by L-Thongkum and Teeranon and Intajamornrak (2007). Tumtavitikul (1993) studies how consonant onset affects tonal parameters; see also L-Thongkum (1992). Gandour, Potisuk and Harper (1996) discuss stress and vowel length. For tonogenesis, see section 3.4.6.


Musical recitation raises interesting research questions. Still of value is List’s (1961) study of linguistic tone and song melody. Tumtavitikul and Promkhuntong (2007) present results of an instrumental study of how Thai classical poetry is chanted. Closely related is how tone, vowel-length and other properties are manifested in speech-act particles, interjections and the like. Perhaps belonging here too is Cooke’s (1992) discussion regarding a possible sixth tone in Thai. Do these items show phonological properties somewhat different from other lexical material? Chuenkongchoo (1956) is among the earliest studies; Peyasantiwong (1979, 1981, 1986) further develops the analysis; also Lodge (1986). For conjunctions and linker syllables, consult Bee (1975). Chainmanee’s (1994) study of filled hesitation pauses breaks new ground in a related area.

3.2.4. Psycholinguistics and phonology; phonesthesia

Psycholinguistic investigations of tone include work of Abramson (1971) on whispered Thai and of Van Lanker and Fromkin (1978), who report different neural processing for contrastive speech tone and non-lexical musical tone. Psycholinguistic implications of the range publications of Gandour and colleagues deserve attention.

Perception studies using experimental protocols have often involved comparative as well as psycholinguistic issues. Wayland and Guion (2003) investigate tonal perception among native and non-native speakers. For consonant perception, including voicing issues, see the

General longitudinal and acquisitional studies include Tuaycharoen (1977, 1979) and Imsri and Idsardi (2003). Bilingual acquisition is the topic of Sarawit (1976).

Direct sound-meaning relationships, including sound symbolism, phonesthesa and quasi-morphological or morphophonemic functions of phonological material are the topic of studies by Henderson (1965), Kam (1980) and Chamberlain (1992). These processes seem to be of more importance for Thai than for English, etc., and merit further study.

3.2.5. Orthography and Romanization

For the Thai writing system, Haas (1956), the first complete description in English, remains a useful resource. Danvivathana (1987) presents the system in all of its detail along with historical treatment of how letter shapes and inventories have altered over the centuries. For those simply seeking a practical introduction to Thai orthography, Iwasaki and Ingkapirom (2005) and Smyth (2002) can be recommended, and many pedagogical manuals, CD-ROMs, etc., are available. The encyclopedia entries cited above also contain overviews. A compact summary of current symbols and basic principles, along with Lao comparisons, can also be found in Diller (1996a). This study includes a historical sound-change rationale for what may seem today like arbitrary and complex system of rules. For more on orthography and diachronic sound change, see section 3.4.6.

Given that the Thai writing system is not only practically efficient but also a longstanding component in the Thai sense of national identity, alternatives have never posed a real threat (Anuman Rajadhon 1961, 1981; Aeorsrivongse 1984; Diller 1993, 2002). True, Kings Rama IV and Rama VI each came up with a reformed writing system and a change to romanization was tentatively considered just after political changes of 1932, yet these were never popular options. A few putative simplifications in Thai spelling decreed during the regime of the 1940’s were quickly undone and forgotten (Hudak 1986).

King Rama VI (Vajiravudh, r. 1910-1925) proposed different transcription systems for native words and Indic loans in Thai, his own name, pronounced [wachi° rawut°], providing a good example of the latter type. Thai family names are often romanized this way: through transliterating etymological Sanskrit letter values rather than indicating modern sounds. Owners of Indic-component Thai names of this category have reported to me hesitancy to change romanization for a surname that was royally granted both in Thai form and also as romanized in the King’s etymological-Sanskritic transcription system. Inconsistent application of this two-fold system also accounts for frequently seen toponym pairs like Dhonburi, Thonburi; Ubol, Ubon, etc. (Ronakiat 2007). For more on romanization and transcriptions, including disparagement of the Haas system and apparently phonetics in general, see Prince Dhaninivat (1970).

3.3. SYNTAX AND SEMANTICS

3.3.1. Syntactic typology

Many, but far from all, authorities in the field would concur in a general way with the following first-approximation sketch. Most (including the writer) would also offer qualifications, counter-examples, definitional quandaries and further debate, as is clear from following entries. As to commonly-cited typological parameters, Thai characteristics include basic transitive syntactic order [S + V-trans + O], in more semantic frameworks represented as AVO, and most frequent intransitive order [S + V-intrans]. [V-intrans + S] occurs also occurs in presentational or existential contexts as mentioned below. Understood noun phrases are widely unstated and construed (i.e. are subject to zero anaphora or deletion, etc., depending on analytic framework) and topicalized orders are common. Taken together, these factors give
rise to a number of alternative pragmatically- or functionally-determined surface orders in actual discourse.

Nominal modification order in Thai strongly accords with the inherited Tai pattern [nominal head] + [modifier], with possessives and relative clauses also following head nouns. Interestingly, Indic compounds have introduced a contrasting [modifier] + [nominal head] ordering. Some commercial and institutional NP names use the Indic order: compare Chulalongkorn-mahawithayalay (Indic compound) and Mahawithayalay Thammasat (as though Tai/Thai noun phrase, although components are both Indic); Ao-Thai Gas [a:w11 thay33 kae:s45], ‘Gulf-of-Thailand Gas’, Indic compound ordering for the whole NP, but components are etymologically Thai and English; the ‘Gulf-of-Thailand’ subcomponent shows Thai/Tai head-modifier order. (This issue is now strictly academic, given Caltex’s acquisition.)

Prepositions precede their nominals. Many auxiliaries precede their main verb, but others follow. Positioning of lexical items with semantic functions of auxiliaries in English, or at least with translational similarities, is not clear-cut. Sometimes reinforcing correlative auxiliary components are found on both sides of a verb complex, e.g. the progressive aspect sequence [kam33lang33] X [yu: 11] ‘to be in process of Xing’ (Kullavanijaya and Bisang, 2004); the deontic sequence [sa:24ma:t52] X [day52] ‘to be able to X’.

Morphosyntactically, as a tonal language with many typically isolating features, Thai retains a core of basic monosyllabic Tai words. However, in the current lexicon this inherited base is statistically overwhelmed by vocabulary from other sources, often polysyllabic with tonally attenuated or perhaps ‘neutral’ reduced syllables. Some prefixal syllables show at least a weak semantic content: [ma-] codes a set of fruit-bearing flora such as [ma-muang52] ‘mango’ and [ma-phra:w45] ‘coconut’. This is transparently compound reduction: [ma:k11], now ‘areca nut’, originally had a wider ‘fruit’ meaning. Compounding of several types is active and common; some sandhi phenomena apply to compounding. Moderate use is made of full and partial repetition. Derivalional processes of several types are mentioned below but no use whatsoever is made of obligatory inflectional morphology to indicate tense, aspect, transitivity, specificity or number. These either are coded lexically, understood from context, or left vague.

Nominal word classes accepted by most scholars are common and proper nouns, including a copious supply of titles and epithets, shading into pronouns; also in the nominal class or else in classes of their own are classifiers, number words (i.e. lexical numerals), and deictics. As modifiers follow head nouns prolifically and endomorphically, rather large noun phrases can be built up. A few nouns are homonymous with verbs, mainly instrumental as in other languages: [thay24] ‘a plow; to plow’; [prae:ng33] ‘a brush; to brush’.

Open verbal classes are more controversial. For the protracted debate regarding adjectives and/or adjectival verbs, see section 3.3.6. Progressive grammaticalization is at the bottom of several definitional quandaries: this process not only moves full verbs into preverbal and postverbal auxiliary subclasses, but also into preposition-like coverbs marking semantic case for following nominals. These construction types impinge on wider serial verb patterns (3.6). Closed functional classes widely recognized include prepositions (but challenged by Warotamiskhkhad 1988; see also Indrambarya 1994), conjunctions, intensifiers and quantifiers, speech-act and polite particles and interjections. Whether or not one or more additional adverb classes might be needed depends on how generously one defines other classes or on which tests are applied.

Syntactic overviews are presented in the encyclopedic articles mentioned above, with issues developed at greater length in other sources cited. For those interested in how syntactic constructions function in actual speech, an excellent place to go first is Iwasaki and Ingkaphirom (2005). Mention should be made also of many fine theses and other studies in Thai not considered here that describe individual constructions with great insight.

3.3.2. History of Thai syntactic research

Historically, interest in Thai syntactic issues by outsiders is of long standing and in varied frameworks. Progress in understanding the nature of Thai can be traced from the brief but
valuable sketch and translated lexical lists of La Loubère (1691), a diplomat-trader, through to
the first explicit grammar in English by James Low (1828), an East India Company
trader-cartographer. He began by informing readers that Thai has ‘no grammar’ (meaning
inflectional morphology?) but then went on to treat parts of speech with understanding. His
book is a technical as well as linguistic milestone: it was printed using the first moveable-type
Thai printing-press font, developed by Ann Judson and her missionary colleagues (Winship
1986). Unfortunately the text was replete with myriad typographic errors. Noteworthy here is
Smyth (2001), who outlines the early study of Thai by outsiders and provides a useful
bibliography and discussion of the early grammars.

A substantial grammar was produced by another missionary, Bishop D. J. B. Pallegoix
(1850). He projected the full panoply of Latinate categories onto Thai, including even the
future perfect tense. Pallegoix’s Latin-Siamese-French-English lexicon, with Thai script and a
credible romanization, was an impressive tour de force. More importantly, as I suspect,
Pallegoix’s scholarly friendship and intellectual exchanges with the Buddhist monk Prince
Mongkut were influential in promoting the sentiments leading to a vision of Thai as the
standard normative language such as we know it today (Diller 2001a). After Mongkut’s
coronation in 1851 he was subsequently known as King Rama IV. He turned his attention to
language reform and to ‘correct’ Thai, apparently the first time a monarch had pursued
normative linguistic interests at that level of detail. He focused not only to lexical issues but
even on syntactic minutiae normally of concern only to linguists, such as prescription of
different classifier constructions (elephants and horses were to be counted directly, without an
idiosyncratic classifier [tua³³] ‘lit. ‘body’, since these animals were considered higher in
dignity than others). Khanittanan (1987a, b) considers linguistic features in royal writing as
they evolved, complemented by the more socio-historical study of Aesrivongse (1984, in
Thai), whose insights are behind Diller (1993, 2001a).

Pedagogical grammars and readers, too many to enumerate here, contain insightful
discussions of specific constructions and exemplify them in context as well. Especially
dependable earlier sources informed by linguistics include Haas (1945); Brown (1967);
Anthony, French and Warotamasikkhadit (1967); Jones, Mendiones and Reynolds (1969),
Yates and Tryon (1970); Kuo (1982). Landmark syntactic analyses written by those with
native-speaker intuitions and reasonably comprehensive in scope include the formidable,
influential and abstruse normative grammar of Upakit-Silapasan (1939, in Thai);
Warotamasakkhadit (1963), the first grammar written in a generative framework; Panupong’s
(1970) inter-sentence study mentioned above; Pankhuenkhat (1978) and Bandhumedha (1979,
in Thai), well-organized and full of insights.

3.3.3. Theoretical perspectives in syntax

Approaches to Thai syntax have included most of the familiar linguistic frameworks of the
mid- to late-20th-century: traditional, structuralist, tagmemic, generative, functional and more
discourse-oriented studies. Upakit-Silapasan (1939, in Thai), mentioned above, presents
essentially a traditional analysis employing familiar Western Classical grammatical categories,
but disguised by neo-Indic nomenclature and making use of several authentically Indic
notions, especially karaka (semantic deep case) theory. In methodology, Noss (1964) and
Panupong (1970) pursue the substitution-frame methodology standard for structural
linguistics of their period and are both convincing examples of that approach. As noted above,
generative work begins with Warotamasakkhadit (1963) and continues on, as that framework
advances, through representative post-Aspects productions such as Bandhumedha (1976) and
Surintramont (1979) on deletion, Warotamasakkhadit (1979) on topicalization, Rodman
(1977), on coordinating constraints and Wongbiasaj (1980), on movement transformations.
Bounding and subjacency are covered by Panpothong (2001). For generative semantics, Stine
(1968), on the instrumental case, is a good early example.

Deletion or nominal omission has frequently been a focus of theoretical treatment. As
noted above, in terms of the patterns S + Vintrans and S + Vtrans + O, noun phrases S and O
are often construed from context. Nominal omission applies to compounding processes as
well. In some conventional expressions, head nouns in compounds are optionally or regularly missing: [kho’24-tha:n33] ‘request alms’, but also ‘beggar’ with head noun [khon33] ‘person’ missing. Note also noodle, rice dishes and other culinary productions, e.g. [phat11-thay33] ‘Thai-style fried (noodles)’; [phat11-kaphraw33 mu:24] ‘(rice topped with) basil-fried pork’, possibly giving rise by analogical formation to [maek45-kaphraw33 mu:24] ‘McDonald’s-style bunned patty with) basil pork’. However, after true prepositions, which are few in the language, nominals resist deletion. Verb-derived coverbs sometimes also reject deletion of following nominal, sometimes not, providing a test for degree of grammaticalization. There are important constraints too, some sentence-internal and some sensitive to macro issues of discourse organization. These are considered with particular insight and cogency by Grima (1978, 1986). For taking missing nominals as empty categories as they were construed in the frameworks of the 1980’s, see Lehman and Pingkarawat (1985), Cole (1987), Kobsiriphat (1988) and Pingkarawat (1989). In the post-Government- and-Binding framework, see Hoonchamlong (1991) and Laksinaking (1991) on anaphora. Outsiders may be astonished that grammaticality judgments underpinning generative work of this type are not infrequently contested by native speakers themselves, a topic examined ethnolinguistically by Diller and Khanittanan (2002).

Numerous additional frameworks have been used. Under the institutional aegis of the University of Hawaii’s East-West Center, work making reference to case grammar and to the specific format ‘lexicase’ (see Starosta 2001), includes Kullavanijaya (1974), on word classes; Savetamalya (2001); Clark (1978), with comparisons to Vietnamese coverbs; and Clark and Prasithrathsint (1985), on synchronic lexical derivation. Deep case, that is semantic role relationships rather than those derived from surface syntax, informs the approach of Lekawatana (1970) and also of Vichulata (1978), who also develops a stratificational analysis. Tagmemic work includes Patamapongse (1971); Phinit-Aksion (1972); Punyodyana (1976). Deepadung (1989) exemplifies residential grammar. For Montague grammar, consult Godden (1981). Diller (1997) wonders whether ‘subject’ is needed in Thai syntax, referring to role and reference grammar. Schiller (1992) produces an autolexical analysis, a framework also used in Wong-Opasi’s (1994) treatment of compounding.

3.3.4. Compounding, nominalization and morphological processes

Compounding and questions of complex word formation are treated by Gehr (1951); Fasold (1968); Osipov (1969); Warotamasintop (1975); Vongvipanond (1992); Witayasakpan (1992); Kullavanijaya (1992); Manomaivibool (2000). L-Thongkum (1994) presents a comparative Tai-Kadai study. Vongvipanond (1992) treats doublets with components of related meaning, a type widespread in the Thai lexicon. See also Wong-Opasi (1994), noted above. The astonishingly large set of compounds relating to emotions and personal attributes with component [cay33] ‘heart, mind’ have attracted much syntactic, semantic and anthropological interest; see the comparative Southeast Asian survey of Matisoff (1986). For Thai data analysed in various ways, see Lee (1987); Diller and Juntanamalaga (1990); Moore (1992). Another common compound type has a component meaning ‘head’, which Juntanamalaga (1992) convincingly relates to kinesic tabu beliefs relating to heads and feet. Thai nominalization is similar to compounding in some respects. Prasithrathsint (1996), (1997), (2006), (2007) presents a compelling set of diachronic and comparative analyses covering this topic.

Some compounding heads referring to people such as [nak45], [cha:w33] and [phu:52] are in effect bound morphemes. [khon33], another ‘person’ word, functions as a common noun but also as a generic compounding head and as classifier. Other classifiers do so as well.

Quasi-morphological derivational processes involving vowel ablaut, tonal variation and even some consonantal interchange is sometimes referred to as elaboration and can be found throughout Southeast Asia, surveyed by Henderson (1967), Nacaskul (1976) and Williams (1991). For Thai, Haas (1964) provides many further examples. An early brief analysis of this material along with more straightforward reduplication is given by Haas (1942). Sookgasem (1997) provides a valuable expansion of reduplication types with discusion of theoretical
ramifications. Various patterns of semi-repetition (elaboration or echo-syllables) not only account for common lexical forms but also, for many speakers, show all the signs of active derivation processes. Some patterns of vowel and tone alternation are used to supply vibrancy and emotive wit to informal spoken language. Kullavanijaya (1997) shows how some of these processes are used in intensifying. The deictic system too shows some quasi-morphological paradigmaticity, discussed in Henderson (1967). Iwasaki (2004) shows how topic-marking particles have been derived from such forms. Diller and Juntanamalaga (1988) speculate as to how the current deictic system may be the residue of diachronic reorganization.

Traditional Thai grammatical study (e.g. Upakit-silapasan 1939; Anuman Rajadhon 1981) makes much of morphological processes affecting Indic loanwords borrowed into Thai. Some changes, such as deletion of many Indic final syllables, are simply a matter of assimilation to the Thai phonological system. Other changes, such as those in morphological recombination of Indic roots to create Thai neologisms, may show prefixation, assimilation, etc., that reflect processes in the Indic donor systems. Pali-Sanskrit terminology such as [sama:t 11] ‘samasa compound’, [son24thi45] ‘sandhi compound’, may be used and are even sometimes applied to compounds non-Indic in provenance. See Gedney (1947); Wan Waithayakorn (1970); Prasithrathsint (1994).

3.3.5. Nominal substitutes and classifiers

Pronouns and nominal substitutes immediately lead into sociolinguistics, as the forms in question are generally more sensitive to such constraints than they are to the more usual pronominal properties of number and even person. The definitional borderline between pronouns, conventionalized epithet-substitutes and ordinary common nouns has been drawn in various ways. Also, classifiers (below) are part of the story as they have anaphoric functions and show other pronominal traits as well. Early studies of note include Cooke (1965), which also compares Vietnamese and Burmese data using distinctive features; Campbell (1969), a useful comprehensive orientation; Palakornkul (1972, 1975) emphasizing sociolinguistic usage; Hatton (1973, 1978), discussing translation; Strecker (1980); Sugamoto (1989). Truwichien (1980) discusses the important topic of address avoidance with great insight, while her longer study (1985) impressively integrates relevant socio-cultural matters. Gething (1986) discusses similar issues. Hatton (1978) and Hoonchamlong (1992) elucidate differences between male- and female-speaker first-person forms and usage. This topic is further developed by Diller and Chirasombutti (2000), who suggest that Thai women are required by the prevailing linguistic system to ‘place themselves’ through self-reference selections in a more finely-determined social space than that required of male speakers. Chirasombutti (1995) provides further detail and a comparison with Japanese. Wijeyewardene’s (1968) contribution at the tabu end of this field is of great anthropological interest.

Hinds (1988) presents a lucid analysis of reflexives, including discourse-based examples. His argument sounds a note of caution for researchers quick to assume that anaphoric properties of reflexives in Thai such as [tua 33], [e:ng33] and [tua 34 e:ng33] equate formally to English translational counterparts. Namthammachat (1975) and most theses written in generative frameworks also treat these issues.

Classifiers can claim a respectable store of professional literature. Most classifiers are derived from homonymic common nouns like [tua 33] ‘body’; a few are from verbs, e.g. [phu:k11] ‘to tie’, as a verb; ‘bundle’ as a classifier for monastic manuscripts. They constitute nearly an open (sub)class, especially if combined with generics like [ya:ng 11] ‘kind, sort’ and measure words, both formal, like [ki33lo:33] ‘kilometer; kilogram’, and conventional, like [kae:w52] ‘(drinking) glass’, as used in a measuring context. These types share many classifier patterns. For many common nouns, classifiers are required in counting expressions, where regular order is [head noun] + [number word] + [classifier]: [no’:ng 45 so’:ng24 khon 33] [younger-sibling two classifier] ‘two younger siblings’. Classifiers are also used to suggest definiteness, especially with deictics in the pattern [head noun] + [classifier] + [deictic], but other modifiers can occur in the [deictic] position if they are used to specify, especially
contrastively. The counting and definiteness patterns can be combined. In these patterns and elsewhere in the language, head nouns are frequently ‘missing’: they are construed from discourse context, giving syntacticians room to theorize as to zero anaphora, deletion, empty categories, etc.

Idiosyncratic classifiers are most strongly required for counting items that are discrete, concrete, and part of Thai cultural life. Two or three dozen are used quite commonly in informal conversation, but many more are recognized: McFarland (1944) gives a list of 82; however, it takes the Thai Royal Institute (Ratchabanditsathan 1995) many more than that and a booklet of 128 pages to prescribe the complete system. As items become more abstract or obscure they tend either to be counted with a general classifier [an]$^{33}$, or counted directly with no overt classifier, or else with a ‘repeater’ construction where a single lexical form occurs in the pattern [head noun] + [numeral] + [classifier] both as head noun and as classifier. A few concrete items, such as several body parts, are also counted in this manner. The system is far from rigid, showing individual and sociolinguistic variation, bureaucratic prescription as above, late acquisition (some items typically learned at school), and, as noted above, diachronic instability both as to the forms used and as to what each classifies (e.g. [tua]$^{33}$, originally for animates with bodies, is now on the rise).

Haas (1942, 1978) was perhaps the pioneer in their analysis along modern linguistic lines and her study remains an admirably concise overview, keeping in mind the passing of time. As a great help both to language learners and to researchers, her dictionary (1964) specifies one or more classifiers for virtually every noun, including indication of those that use the ‘repeater’ construction. Hiranburana (1978) concentrates on a hierarchical semantic classification, with Placzek (1984, 1985, 1992) providing more detailed studies of this type, with interest in perceptual salience. Plam (1974) supplies a treatment both semantic and syntactic, as do Köhler (1979) and Hundius and Köhler (1983); see also Zhang (1992). Kookkiattikoon (2001) looks in depth at their syntax.


For comparative classifier studies, see section 3.4.4.

### 3.3.6. Verbs, transitivity, predicate constructions and grammaticalization

Thai verbs are a robust class but should a separate class of adjective be recognized? There are good reasons to take adjectives as a subclass of verbs, following Gedney (1947), Haas (1964), Panupong (1970), Hudak (1987) and others who subcategorize these items as a type of verb, perhaps ‘adjectival verb’. Such forms do not require (or even admit, in most cases) a copula such as [pen]$^{33}$ or [khu’:]$^{33}$ ‘to be’ to form complete predicates and they interact with auxiliaries, negatives and question/answer patterns mainly in the way full verbs do. But good arguments to the contrary have been advanced too, variously invoking comparative constructions, the ability to occur with certain prefixal formatives, semantic effects of repetition, etc. See Noss (1964); Prapa (1996), Smyth (2002), Iwasaki and Ingkaphirom (2005) and Thai traditional grammar. These authorities, it would seem, would need to acknowledge that complete Thai sentences could lack overt verbs. Nominalizing formatives [ka:n]$^{33}$ for verbs and [khwa:m]$^{33}$ have been often been used to distinguish these categories but a number of verbs with cognitive-emotive meanings like ‘understand’, ‘detest’, etc. accept both prefixals equally well. To call adjectival verbs ‘stative verbs’ has been one attack, but one might well wish to call verbs like [yu:]$^{11}$ ‘to be located’ stative verbs as well and these fail most of the putative adjective tests.

As though to make tightly defining verbs yet more onerous, Thai main verbs are difficult to distinguish absolutely from forms one might wish to take as auxiliaries: markers of timing,
As noted, Thai verbs show no formal marking as to transitivity, nor do they overtly indicate finite-nonfinite status. The ability of nominals to be deleted or not to occur overtly presents a moderate challenge to transitivity analysis, but most authorities concur that many lexical verbs are basically in a transitive class, a few like [hay52] ‘to give’ in a ditransitive one, many more are intransitive and still others alternate in potential transitivity status, perhaps along with other minor semantic effects. Landmark studies of verb classes include Noss (1964), Panupong (1970) and Sindhavananda (1970). More recent discussions are by Sriphen (1982) and Thepkanjana (1992), who develops the useful notion of a transitivity continuum, taking account of verb semantics. Savetamalya (1992), emphasising patient subjects, discusses transitivity using a case-grammar approach. Negation of predicates and elsewhere is considered by Kanasut-Roengpitya (1974) and Lagsanaging (1992).

Semantics and pragmatics interact in transitivity issues. Panupong (1978) and Diller (1997) wonder about how best to analyze single verb forms exhibiting alternating transitivity. Part of the question must include pragmatics: in a rather common construction type, topicalized undergoer or patient/theme object occurs without overt agentic subject: O-undergoer (S-agent) V-trans. Does this merge in a gradient way with S-theme V-intrans? Perhaps related to this problem is a small class of high-frequency verbs such as [mi:33] ‘to have; there is/are’ and [koe:t11] ‘to be born; to happen’. Here a single form appears to have both a transitive use and also to occur intransitively in a pattern of V + S when showing existential or presentative meanings. Sookgasem (1992) clarifies the issues, also considered in the lexigsaw system by Indrambarya (1996).

This leads to the question of marked passive or pseudo-passive constructions, well-researched topic of long-standing interest (Bergen 1875). Prasithratthatsint (1988) documents substantial diachronic change in a range of passive-like constructions. A common issue in the current language involves a verb [thu:k11] ‘to come in contact with, touch’. This is widely recognized with a shifted, generalized meaning more like ‘to undergo, suffer’ and with the function of an adversative passive: [mac:w33 thu:k11 ma:24 kat11] [cat undergo dog bite] ‘the cat was bitten by the dog’. In this construction, the form [thu:k11] retains nearly all of the formal syntactic collocational properties of a main verb, if not quite retaining normal verb semantics. This has lead some authorities to resist labelling the construction as ‘passive’ per se, as argued forcefully and entertainingly by Noss (1972b), holding that using ‘passive’ here would be too much of an imposition of Western grammatical conceptions. (Is it churlish to observe that Noss registers no similar problem with ‘subject’ and other Western category labels, which his grammar liberally utilizes?)

In any event, as Khanittanan (1979), Prasithratthatsint (1988) and others have maintained, the construction has been widely used to translate English and other Western-language passives. Perhaps partly because of resulting translation genres, usage of the [thu:k11] construction is assuredly becoming used in less adversative contexts, especially in Thai technical discourse (‘the metal was dissolved in acid’) and among middle-class speakers (‘I was invited by him to the party’). This recalls Prasithratthatsint’s (1988) demonstration that earlier Thai ‘passive’ constructions have been moderately changeable. A syntactically similar, but less common, form [do:n33] ‘to be hit by’ is more stable in its adversative semantics. The (pseudo-)passive problem is treated structurally and functionally in several doctoral theses and in briefer works, among them: Filbeck (1973a); Lekawatana (1975); Thanyarat (1983); Wongbisaj (1979b) in a generative framework and Savetamalya (2001) in a lexigsaw one. Morev (1996) investigates these matters from the perspective of diathesis, which includes consideration of causatives and other transitivity-shifting issues. Gero (1977) and Gsell (1979) are concerned with a similar range of issues.

Serial verb constructions, directionals and causatives have been the focus of much syntactic research. The preceding discussion indicates that the pseudo-passive markers mentioned above have many properties of verbs; as such they may be implicated in the wider phenomenon of serial verb constructions. Definitions have varied but many take this type of construction to consist of two or more verbs or verbal predicates strung together without overt aspect and modality. In serial coverb constructions they intrude on prepositions.
marking of coordination or subordination. In many instances, or in all if so defined, at least one
nominal argument is shared, such as subject/agent or object/patient. Usually at least one
such nominal is understood, i.e. is an empty category or zero anaphor. Crucial here also is the
observation that along with many other serializing languages Thai lacks morphological
marking distinguishing finite from non-finite. Also, to assume that the first verb in sequence
were in all cases the dominant one leads to quandaries. Constructions of this type are a
Southeast and East Asian areal feature, as documented by Clark (1978, 1992) and Clark and
Prasithrathsint (1985) in a lexicase framework. Bisang (1996), with grammaticalization and
‘great attractors’ in mind, analyzes Thai examples in this wider context too, as does Post
(2007). Analyses reveal both broad similarities and specific differences with comparable data
in nearby languages.

Needleman (1973a) is among the first studies in which a formal post-Aspects generative
framework confronts the challenge of the seemingly flat multi-verb structures of Thai verb
serialization, followed by Filbeck (1975), Vis (1978) and Sereechareonsatit (1984).
Thepekkanjana (1986) should be credited, it seems to me, with a notable advance by turning
attention to specific semantic subclasses of verbs as a constraint in how longer complexes are
contextually built up and strung together, a direction developed by Chuwicha (1993) and in
other work. Somewhat similar in orientation, but using the lexicase apparatus with robust use
of synchronic derivation, Clark and Prasithrathsint (1985) offer an analysis of verb forms
showing context-sensitive differences in sense or structural properties. Working in the same
basic framework, Wilawan (1992, 1993) goes on more radically to supply an argument
rejecting the ‘serial verb’ characterization entirely and taking the relevant constructions to be
coordinate or subordinate clauses within a system of sentence adjuncts. For more
consideration of these matters, further references and a scheme based on
symmetrical/asymmetrical properties of these constructions, see Diller (2006a), which also
mentions properties of shared arguments, negation, modality and timing.

The verbs [pay 33] ‘to go’ and [ma: 33] ‘to come’ and similar verbs of motion enter into
serial constructions of great frequency, mixing directional, temporal, aspectual and evaluative
nuances in intriguing ways that have stimulated much study. These forms are considered in
many of the sources cited above, but studies focus on them specifically as well. Gandour
(1978a) associates ‘come’ and ‘go’ with deictic properties, with Treerat (1990) going on to
come (!) to grips with how syntax, predicate semantics and discourse context interact.
Temporal-aspectual functions of [pay 33] offer a particular challenge, with interpretations of
time sometimes suggesting past completive, sometimes future continuative. [khaw 52] ‘to
enter’, hence inchoative, and some other directionals contribute to the topic as well. Bickner
(1985) and Rangkupan (2001) analyze usage of directionals as it indicates psychological
perspective in narrative and in other contexts.

The verb [hay52] ‘to give’ in serial constructions promotes a different polysemous complex
involving notions like benefactive, causative and complementizer of controlled action. Areal
considerations are again relevant; see comparative analyses by Hermann (1979); Pooisrakit
differences in intended versus inadvertent causality in constructions with [hay52], [tham 33] ‘to
do’ and the compound [tham3hay52]. Other studies are by Kumlert (1976), Khamsaen (1978),
Indrambarya (1992), with Warotamasikkhadit (1994) offering questions as to prevailing
terminology. Grimba (1978), while not limited to [hay52] constructions, presents an especially
cogent discussion of the associated types of zero anaphora. See also chapter 17.

Temporal-aspectual information may be marked by the use of directionals to indicate
temporal-aspectual nuances is mentioned above. Other marking of this type, along with
modality, can usually be traced back to main-verb sources too, although such verbs may or
may not still be in active usage. A problem encountered by those attempting to establish a
fixed linear order for lexical forms in auxiliary constructions, e.g. Anthony (1964), Dellinger
(1975), is that many of the dozen or more commonly used forms can function in different
surface positions with slightly different meanings and collocational dynamics; see critique of
Warotamasikkhadit (1979). Thus the form [a:t11] can function epistemically as ‘apt to, likely
to’ but also deontically as ‘able, capable of’. The epistemic position is more peripheral, preceding positions of irrealis marker [caː11] and negative [may25], with the deontic alternate closer to the verbal core and following positions of such formatives. The favored structuralist solution has been to expand the lexicon: Noss (1964) distinguishes two [aː11] homonyms in different form classes.

Kimsuvan (1992) considers alternations with [yuːː11] ‘be located; be happening; at’. This form along with others has been investigated in detail by Kullavanijaya and Bisang (2004) in a selection-theory approach. Similar syntactic and semantic alternations apply to several other forms. [day25] ‘can’ (and many further possible glosses) is notorious, with functions ranging widely over the epistemic, deontic and temporal-aspectual semantic territory. Enfield (this volume) has given much attention to the counterpart of this form in Lao and much of what he describes in this volume for the Lao would apply to Thai as well. Other approaches involve linkage of semantics to various treatments of phrase structure, an especially well-designed analysis being that of Sookgasem (1990) in Head-driven Phrase Structure Grammar. See also studies of Kanchanawan (1978), with interest in machine translation; Ekniyom (1979), using internal reconstruction; and theses of Kullavanijaya (1968); Scovel (1970); Boonyapatipark (1983). Howard (2000) and Koenig and Muansuwun (2000) inspect perfectivity in detail, e.g. use of the form [laeːw45], originally a verb ‘to finish’.

Much of the above discussion relates in one way or another to grammaticalization paths, especially to situations where diachronically prior constructions and senses coexist in the current language along with their evolved alternates, with little or no phonological indication of difference. Enfield (2006) presents a cogent discussion of grammaticalization issues relating to Lao, in most cases with direct application to Thai as well. Not only verbs and derived auxiliaries are at issue, but also other paths like [verb – coverb – preposition], cp. [caːk11] ‘to leave (now very restricted); from’ or [noun – preposition], cp. [lang24] ‘back (body part noun); in back of (as though preposition or adverb)’, with yet another function being classifier for houses. For more on what is essentially grammaticalization, if not explicitly so identified, see work of Clark (from 1978 onwards), Jagacinski (1991), Juntanamalaga and Diller (1992), Bisang (1996), Diller (2001b). Iwasaki (2004) considers the grammaticalization of topic-marking form [nia52]. See Part Four of this volume.

3.3.7. Other parts of speech and constructions

Questions are taken up in work by Kullavanijaya (1980) and Santaputra (1984). A common yes-no question type is coded by a final particle written as though [may24] but usually pronounced [may45]; the ultimate diachronic source is undoubtedly the preverbal negator [may42]. This and other final particles with various speech-act, politeness and discourse functions are analyzed in a number of studies, Cooke (1989) providing a comprehensive orientation, with other studies by Bhamaraput (1972), Peyasantiwong (1981), Kendall, Yoon and Hye-Suk (1986), Horie (Ingkaphirom) and Iwasaki (1996). Phonological reduction of these forms is studied by Peyasantiwong (1979). Bandhumedha (1979) presents a convenient grouping of families of particles and perceptive analyses of functions. For specific treatments, see Cooke (1979) for the [si45] set and Neill’s (1989) narrative-based analysis of [na45] and [chay52 may45]. Diller and Juntanamalaga (1992) take up the [oːy24] set. Strings of particles can occur together. Whether such forms are subject to ordering rules is considered by Warotamasikkhadit (1975) and Prasitrathisinth (1974).

Conjunctions are considered by Thomas (1979) and by Clark (1994), who presents a comparative study and argument for a topicalizing function. Jagacinski (1991) takes up complementizer [waːː2] and other complement types. For relative clauses and factitive noun clauses marked by formative [thiːː2], see Ekniyom (1971); Suktrakul (1975); Sornhiran (1978); Kuno and Wongkhomthong (1981b); Savetamalya (1996) and, for a convincing diachronic analysis, Kullavanijaya, this volume. Comrie (1996) treats clauses of this type comparatively with particular attention to Japanese. Morev (1994) describes possessive constructions. For reciprocals, Bee (1972) comes to grips with the poly-functional [kan33], variously a reciprocal, gathered-plural marker and male first-person pronominal form.
Constructions and their functions of sentences with copular forms or equatives \[\text{pen}^{33}\] and \[\text{khu}:^{33}\] are taken up by Kuno and Wongkhomthong (1980, 1981a). As part of the wider picture, Warotamasikkhadit (1969, 1976a) treats other idiosyncrasies of the verb \[\text{pen}^{33}\] along with verbless sentences.

A postverbal nominal in a construction that might seem SVO superficially on inspection might not be a typical O-type patient semantically: it might turn out to have a locative, manner or instrumental function. Instrumental constructions are described by Warotamasikkhadit (1986). The postverbal position can also accommodate affected body-part terms, e.g. in pain expressions. These introduce the issue of how pain is represented and how this affects construction choice, as discussed by Diller (1980) and by Iwasaki (2002); also in a more comparative anthropological mode by Fabrega and Tyma (1976).

Quantification is taken up by Stine (1981). Whatever the theoretical approach, there can be little doubt that Thai quantificational phrases are prone to occur at the end of predicates, even if this means splitting an earlier noun phrase. Wongbiasaj (1979a) considers this type under quantifier floating. In a related area, Haas (1946) and Kullavanijaya (1997) give examples of a range of intensifying techniques, several of which involve final position as well.

3.3.8. Discourse, conversation and sociolinguistically-based studies

Not a few of the studies cited above consider units greater than single sentences, among them Panupong (1970), Grima (1986) and Hinds (1988b). Iwasaki and Ingapirom (2005), and in their other work, frequently rely on conversation-based examples and cogently take such wider contexts into account. Thomas (1988) argues that Thai grammar needs to recognize the language as 'paragraph-efficient' rather than as based on rules confined to single clauses, with Vongvipanond (1988) arguing for two types of linkage devices: 'macro-cohesive' and 'micro-cohesive'.


1986 was an especially fruitful year for dissertations devoted to Thai discourse by native speakers. Burusphat (1986) presents a comprehensive discourse analysis based on narrative folklore, with texts presented in a convenient appendix, fully transcribed, glossed and translated. Chanawangsa (1986) studies cohesion from several perspectives, while Chodchoey (1986) uses transcribed materials to uncover strategies in oral discourse.

Conversational Analysis (CA) refers to micro-analysis of discourse along sociological lines, including special attention to pauses, repetitions, self-corrections, kinesics, etc., usually ignored in other frameworks. Moerman (1988) presents stunning examples of the insights that this style of analysis can reveal. An extra dimension of complexity in his study is the combination of Northern Thai, Lue and Central Thai in the text material, which is presented in meticulous oral transcriptions. Bilmes (1992) makes similar points. Turn-taking and speaker overlap is of interest in these studies, with Hinds (1988) providing a provocative parallel between Thai driving behavior and informal conversational interaction. While filled pauses are also of interest in CA, Chaimanee (1996) instead devotes attention to them in a comparative study of native and non-native speakers’ hesitations.

Khanittanan (1987a, 1987b, 1988b) uncovers diachronic trends in discourse genre and makes a compelling case for the evolution of a more ‘autonomous’ and abstract type of Thai, partly as the result of different communicative needs arising from modernizing social trends. This genre is less tied to traditional constraints of immediate interpersonal communication, apt to use abstract nominalizations and to make nominal material explicit. In a similar vein, differences between spoken and written discourse are clarified by Chodchoey (1988). Person (1996) turns attention to specialized genre: oral sermons of a Buddhist monk.

For Thai-Japanese comparisons, see Ruetaivan’s (1999) study on how motion events are represented in narrative discourse.

3.3.9. Semantics and lexical fields

Several of the syntactic studies already cited show a strong semantic focus. The distinctive feature framework is used by Cooke (1965) in his comparative study of pronominal reference. Hiranburana (1978) develops a related hierarchical-taxonomy approach to classifying Thai classifiers, with several works of Placzek (e.g. 1992) producing a more finely-tuned analysis with particular semantic attention to shape. Gething (1986) discusses the extent to which distinctive features relate to cultural matters. As mentioned above, Thepkanjana’s work (e.g. 1986) is notable for uncovering how semantics of verb classes can constrain the makeup of serial verb constructions.

Compounding has attracted semantic attention. The studies cited with regard to compounding [cay33] ‘heart, mind’ to yield emotional terminology treat semantics as well as syntax. Juntanamalaga’s (1992) study classifies the cluster of meanings associated with [hua24] ‘head’ as it occurs in compounds. Vongvipanond (1992a) presents a substantial semantic and syntactic analysis of compounds where component parts are synonyms.

Additional studies link meaning with syntactic constructions in specific semantic fields. As mentioned, Thai pain terms are described in a comparative context by Fabrega and Tyma (1976) and by Diller (1980), who points to interesting syntactic features of these constructions. Their syntactic analysis is substantially advanced by Iwasaki (2002). A similar topic, traditional Thai disease terminology, is developed in a more ethnographic study by Bamber (1987). Still relating to physiology but in a cheerier subfield, Reed (1976) produces a semantic analysis of Thai gastronomic terms.

Representative thesis-length treatments of semantic topics include Gething (1972) and Terayanont (1988). Varied approaches are attested. Gething (1968, 1972) develops a structural redundancy methodology which he applies to analysis of nominals. His later work (1975, 1979) treats the semantics of locatives and other expressions with comparative reference to Lao. Diller (1994), in a comparative volume, attempts to address issues raised in the ‘semantic primitives’ enterprise. Folk taxonomy is the focus of studies by Simmonds (1978) and by Stott (1978), describing vernacular forest nomenclature. Historical semantics is taken up by Khanittanan and Placzek (1982), who trace how inherited Tai ‘psychic’ vocabulary, such as [khwan24] ‘(roughly) psyche, spirit’, has undergone semantic readjustment with the introduction of Indic loans such as [win33ya:n33] ‘(roughly) soul’.

Another historical study of anthropological interest is that of Gething (1977).

Toponyms as a reflection of culture, social organization and environment are treated in impressive studies of Hartmann (2007) and Prasitrathsint (2007b). These link names of villages and other geographical units in Thailand to the wider Tai context, along with consideration of the Southeast Asian ecosystem and socio-political factors.

Kinship: of anthropological merit and impact are studies of Thai kinship terms and their extended usages. Gething (1986a) discusses distinctive features organizing Thai kinship and occupational terms. In a more extended study, Prasitrathsint (2001) presents a componential analysis of the complex system with useful comparative references. For diachronic/comparative development, one could consult also Black Tai (Fippinger, 1971) and check Strecker’s (1980, 1984) Proto-Tai reconstructions. Formative studies are by Benedict (1943, 1945), with Chinese comparisons. Also of interest to anthropologists is the fact that Thai speakers appear to enjoy ‘disrupting’ their lexically normative kinship system in
manners exotic (to outsider anthropologists, at least): parental terms [pho:;42] ‘father’ and [mae:42] ‘mother’ are used by parents (among others) in referring to or addressing their children: thus, one calls one’s daughter ‘mother’; but the daughter may call her mother [mae:42] too. Also, these terms may precede children’s given names or nicknames in the manner of a title. There is yet more: sibling terms are the widely used in address and reference among spouses; see Haas (1969).

3.4. SOCIOLINGUISTIC, COMPARATIVE AND HISTORICAL STUDIES

3.4.1. Sociolinguistic perspectives on Thailand

Studies cited above dealing with pronouns, kinship terms, speech-act particles and discourse are especially likely to be of interest to sociolinguists. In this subsection we note more general descriptions of the rather complex setting in Thailand important in understanding communicative functions of Central Thai in their social setting. For those requiring an overview of Thailand’s sociolinguistic situation, Premsrirat (2006) is recommended. The issue of *International Journal of the Sociology of Language* 186 (2007), devoted to Thai sociolinguistics, includes range of current work, some mentioned below. For those interested in the development of the subfield of Thai sociolinguistics, comparisons could be made with the sociolinguistic review of Debyasuvarn (1973) and also with an issue of the journal *Language Sciences* 10 (1988), devoted to papers on Thai sociolinguistics of that era. Roop’s (1969) earlier study of language diversity remains useful for perspective.

Beebe (1975, 1976, 1981) presents an impressive set of Labovian studies correlating articulatory alternatives with variables like occupation, age and gender. There is also treatment of these issues a language-learning context by Brown (1967), who usefully distinguishes colloquial ‘Bangkok Thai’ from the standard language; the former variety shows cluster simplification, substitution of /l/ for /r/, etc.

Multilingualism and language hierarchy in Thailand is the theme of the impressive and informative survey of Smalley (1994), expanding his (1988) introduction. For Smalley, Thailand’s local dialects and minority languages are part of a hierarchical structure with standard Central Thai at the top. Speakers tend to become bilingual or multilingual in varieties positioned over them in the hierarchical model. Following from Smalley’s observations, investigators find that local dialects of the Tai family in Thailand are by now most frequently components in speakers’ bidialectal (or multilingual) competence: outside of the Central-speaking area virtually all speakers with basic public schooling have at least a functionally passive competence in Central Thai, although active competence and most daily-life communication may be in a local variety. These local varieties are not infrequently mixed with or influenced by the standard language. Chamberlain (1972) and Khanittanan (1973) describe tonal influence of the standard language on Northeastern or Lao varieties. Diller (1979) investigates how tones and segments in Southern and Central Thai are systematically combined in sociolinguistically salient ways. In particular, hybridization is described whereby Southern Thai tones are articulated for what is otherwise Central Thai. This can apply even to the standardized written form. For the north, the interplay of Kam-mueang (Northern Thai), Lue and Central Thai in their sociological context is amply elucidated and documented by Moeriman (1965, 1988) in a closely-focused conversational-analysis framework. Similarly, but in a different framework, Premsrirat’s (2007) study of endangered languages in Thailand calls attention to this urgent line of research, recalling Bradley’s (1992) discussion of the disappearance of Ungong. Morita (2003, 2007) studies the important question of assimilation and language shift among Sino-Thai speakers; her bibliographies are a useful resource for those interested in this issue. From another quarter, Chunsuvimol (1980) considers communicative networks of Thai workers in Singapore.

Dialectology studies of local and regional varieties in Thailand are too numerous to be listed here: many are catalogued and described by Tingsabadh (1984), with leading studies discussed in Smalley (1994). Many fine theses completed in Thai universities describe
individual local varieties. Brown (1962) and Hartmann (1980) treat issues of subgrouping. These works and sources mentioned therein can be consulted for Kam-mueang (Northern Thai), Southern Thai and so-called Isan Thai (essentially Lao varieties spoken in Thailand). Chantavibulya (Panupong) (1959) appears to be the first detailed linguistic description of a local dialect (Songkhla, Southern Thai). For comparative reference, a remarkable nearby Southern tonal system is acoustically and physiologically analysed by Rose (1997).

Central Thailand is far from uniform in local speech. For the central-west area, Suphanburi Thai is described by Tingsabadh (1988, 1992), who makes the intriguing and methodologically challenging observation that speakers of this variety really cannot cite their local forms in isolation. Court (1972) describes the unusual tonal system of Traat.

Early landmark comparisons with focus on tone compare varieties within Thailand, and sometimes beyond: Haas (1958); Egerod (1961); Jones (1965). Brown (1962, 1965) is the first book-length work to cover the entire country in detail. It establishes a convenient visual display for comparing tonal systems and remains a valuable dialectology resource. Many later works describe new varieties or provide further analyses, e.g. Strecker (1979); see Tingsabadh (1984).

Sociolinguistic and political commentary on Thai as a standard national language is provided by Aesrivongse (1984, in Thai), whose viewpoint is summarized and extended by Diller (1988, 1991).

3.4.2. Registers, polite speech and special styles


Native speakers of Thai find it easy to characterize samples from written and spoken registers in evaluative terms like [supha:p 52] ‘polite’ or its opposite. In effect this presupposes a diglossic arrangement which includes the distinguishing of lexical pairs that can be described as ‘higher’ and ‘lower’ in terms of speech-level, although the relationship among resulting registers is far from a simple binary one. Diller (1985, 1993) suggests that both lexical and syntactic evaluations contribute to this continuum, but it should not be concluded that written Thai always shows ‘high’ selections and spoken Thai ‘low’ ones. Sometimes written Thai opts for ‘lower’ expressions as in journalism: Khanittanan (1994, 2007) shows how styles are manipulated for effects in different segments of news presentation; see also Srinarawat’s (2007) informative work on political slang, another use of ‘lower’ expressions for effect.

Politeness markers and similar indicators of interpersonal dynamics are ubiquitous and especially salient in colloquial Thai. These categories are coded by various means: most obviously by address and reference forms and by final particles such as polite final particles [kha45], [kha52] (female speaker) and [khrap45] (male speaker) and the less polite [ha35], [wa45], etc., (Bandhumeda, 1979; Peyasaniwong 1981; Cooke 1989), but also by various lexical options along the high-low continuum mentioned above. This continuum generally reflects ascribed deference as well as formality. But the system is complex. Lexical selections in formal Thai cannot be taken as necessarily indicative of marking interpersonal deference, e.g. in formal sources high-status and low-status males alike may be referred to as [na:y33 X]: ‘Mr. X’, [na:y33 thak45sin24] ‘Mr Thaksin’ (a former Prime Minister); [na:y33 dae:ng33] ‘Mr Daeng’ (an unemployed unskilled manual worker). Compare more colloquial Thai: [na:y33 X] seems less deferential than [khun33 X] ‘Mr. X’ (lit. ‘honorable X’). Note that [khun33 X] is uncommon in formal Thai, e.g. in official reports, etc. Even syntactic choices are implicated in this complex. Khanittanan (1988a) elucidates strategies and issues with great insight, as do Kummer (1992) and Bilmes (2002), while Deephuengton (1992) turns to disagreement strategies. Work of Moerman, Hinds and Iwasaki keep such factors in mind as well.
Thai is remarkable for special registers relating to institutions of royalty and monkhood, although the lexical makeup of these registers shares many selections with the ‘high’ or ‘polite’ [supha:p] level mentioned above. Particularly given the impact of broadcast media, most Thai speakers are at least passively familiar with the hundreds of lexical substitutions these registers prescribe, if not always able to control them actively. Haas (1951) and Jones (1971) lay out the complexity of the system, especially as regards degrees of deference ascribed to ranks of royal persons. Gedney (1961) provides a convenient outline of these registers, while Kanasut-Roengpitya (1973) documents them in greater detail. Diller (2006b) also gives a sketch of ‘royal Thai’ calling attention to some of its derivational processes and suggesting that functionally it shows some parallels with special avoidance genres in other languages.

Poetic Thai and other literary genres have their own special vocabulary and syntax and show points of contact with the type of language used with royalty and the monkhood. Also, the formal metrical requirements of traditional poetry provide a number of clues both as to abstract phonology and as to how Thai may have developed diachronically. Linguists undertaking work of this type include Warotamaskkhadit (1968a); Bickner (1981, 1992); Chittasophon (1984); Gedney (1980, 1989b); Hartmann (1989); Hudak (1990 and this volume).

3.4.3. Loans, contact and bilingualism

General contact issues are introduced in traditional philological studies, see Anuman Rajadhon (1961, 1981). A linguistic orientation is taken by Warie (1973, 1979). Sometimes overlooked is the importance of foreign language education as an aspect of contact, considered by Chirasombutti (2007).

In the preceding section, lexical items of the higher diglossic registers were noted. These in general are etymologically not inherited Tai vocabulary but are loans introduced over many centuries. Perhaps the most concentrated and conspicuous loan element in Thai is vocabulary from Indic languages Pali and Sanskrit. Gedney (1947), in a work still of great value, presents a formal analysis and catalogue of many hundreds of words of this sort, mainly nouns and verbs. Historically, Indic vocabulary in Thai can be classified into two types: (i) traditional loans, accepted into Thai from the thirteenth to nineteenth centuries; and (ii) coined neologisms created in the twentieth centuries through official institutional action. There are hundreds of both types in current use and distinguishing them is often not a simple matter. Type (i) were typically introduced in pre-modern times through Khmer or Mon intermediaries, subject to assimilation processes in those languages first, or else from Pali Buddhist sources, i.e., terms originally used in religious contexts came be used more generally. Most of these loans are transparent in provenance, but not all. Harris (2007b) convincingly shows how Mon or Khmer mediated Sanskrit marica ‘hot (capiscum) pepper’, borrowed into Thai with sound changes, resulting now in [phrik].

Type (ii) loans were created by literary and technical authorities in an environment of modernization requiring lexical enrichment, but coupled with nationalistic oresthetic sensibilities disparaging direct loans from English or other Western languages. The solution was to code the Western-based concepts with neologisms constructed from Indic (mainly Sanskrit) morphemes. The neo-Indic forms created also conformed analogically to various assimilated prototypes already established through type (i) forms; for example, there was no attempt to pronounce neo-Sanskrit material the way it would be pronounced in India. Sometimes shifts have occurred in the designated semantics of type (ii) items, especially those of the 1920s or 1930s. Wan Waithayakorn (1970), one of the key figures in the coinage process, and Prasithrathsint (1994) discuss the specifics of how these (ii) items were introduced. See also Court (1984).

Austroasiatic, Austronesian and Sino-Tibetan language families have contributed many vocabulary items to Thai at different diachronic periods. Relatively recent loans from Khmer, Malay and southern Chinese varieties like Taeciw are often transparent, while earlier strata feel to Thai speakers like authentic Thai vocabulary and require linguistic scholarship to
deduce provenance. Careful analysis and correct differentiation of diachronic strata from various sources of vocabulary now used in Thai can be considered still in the infancy stage. Surely however such studies are a precondition for convincing discussions of remote or ultimate Tai and Tai-Kadai relationships, be they genetic or some form of contact, if that distinction can indeed be maintained.

Contact with Khmer and other Austroasiatic languages is a topic pursued by Khanittanan (2001) and Diller (2002), who both argue for an early period of substantial Thai-Khmer bilingualism. Evidence for this claim lies in the sweep and magnitude of Thai vocabulary that is Khmer-derived (Varasarin 1984). While some occurs at the higher speech-levels, much is basic and currently in every-day usage: [camu:k11] ‘nose’; [tapho:k42] ‘hip’; [khu’:33] ‘to be (equivalent)’; [doe:n33] ‘to walk’; [koe:t11] ‘be born’; [set11] ‘to finish’. More tellingly: a number of auxiliaries, conjunctions, prepositions and other grammatical formatives are clearly of Khmer or Mon-Khmer provenance, as is the disyllabic lexical pattern, seen in ‘nose’ above, now so thoroughly incorporated into Thai that it seems native.

Turning to Austronesian, Suthiwan (1992) focuses on Malay loans into Central Thai. She is able to show several different diachronic strata on the basis of tone assignment and stop devoicing. One Ayudhya-era level is associated with a Thai court adaptation of the Javanese and Malay Panji tales and many of the loans are literary in character, but other strata have toponyms, names of fruits, etc., that are in common usage. Suthiwan also considers Malay loans into Southern Thai, as does Court (1975), who includes insightful diachronic deductions and discussion. More contentious are the earlier levels still. Schlegel (1902) advanced the hypothesis, revived and enhanced (?) by Benedict (1942, 1975, etc.), that Proto-Tai vocabulary shows enough plausible Austronesian cognates to support a standard genetic (hence Austro-Tai) relationship. This topic is pursued in detail elsewhere, whether pro, con or prevaricating: (e.g., Gedney 1976; Diffloth 1977; Reid 1984; Hartmann 1986b; Matisoff (1990); Thurgood (1994); Diller 2000; Ostapirat 2004).

Another important contact domain concerns Chinese relationships, again a matter of contention. Work of Egerod (1957, 1959b), Manomaivibool (1975, 1976, 2000), Li (1976) and Luo (2000; also this volume and sources mentioned therein) suggest that there are very early, if not genetically inherited, strata of Chinese in the Tai (and Tai-Kadai) languages. Manomaivibool (1976) carefully differentiates early strata, e.g. [plu’ak11] ‘peel; bark’ versus [phiw24] ‘skin’, which she takes to be from the same Chinese lexical source 膚, but reflecting different stages. Luo shows the magnitude, basic nature, semantic spread and emphasizes regular correspondence patterns of many items involved. Thurgood (2007a), on the other hand, calls attention to an irregularly-corresponding lexical group taken by Gedney (1979) as evidence for a new set of Proto-Tai initials. Thurgood suggests rather that the items in question represent Chinese loans of various strata, hence the irregularity. This problem is perhaps indicative of more general methodological tensions attending Chinese-Tai (or Tai-Kadai) diachronic research. Whatever the nature of the ultimate Sino-Tibetan relationship, vocabulary of early eras (e.g. terms for numbers, body parts, basic verbs, etc., on at least the Proto-Tai level) is not to be confused with much more recent southern Chinese (mainly Swatow/Taeciw) loans into Central Thai (Egerod 1959a). Unlike the earlier vocabulary, later items are felt by speakers of Thai to be ‘Chinese’ and sometimes even have phonological properties that effectively mark them as such (e.g., unaspirated initial stops in items with high or rising tones, precluded as a regular possibility for inherited Tai vocabulary). Such vocabulary refers especially to foods, cooking processes, business terms and other transparently Chinese cultural concepts, but also to some pronouns now commonly heard in Thai speech, at least on Bangkok streets. Of relevance here are Morita’s (2007) work on Chinese assimilation and Srinarawat’s (1988) on language use of Chinese-background speakers in Bangkok. These are interesting to contrast with Huang’s (2007) study of Zhuang-Chinese assimilation.

Portuguese, Persian, Arabic and other languages recalling Indian Ocean pre-modern commerce have brought into Thai designations for grapes, cabbage, roses, soap and other cultural items; see Harris (2007a, 2007b).
English has had a moderate to heavy impact on Thai, explored sociolinguistically by Khanittanan (1979), Nacaskul (1979) and by Chutisilp (1984). Work on the development of the so-called adversative passive marked by [thu:k] ‘to undergo’ often notes the use (or misuse) of the construction in translating English passives (e.g. Prasithrathsint 1988; Diller 1993). Lexically, of historical interest is the fact that many English loans like [sathe:33chan52] ‘station’, commonly used in the nineteenth century, were replaced by Sanskrit neologisms in the early twentieth century, such as [satha:24ni:33]. In this case, the Indic neologism even recalls the form of the English prototype as well as its etymological connections (note Proto-Indo-European sta: ‘stand’, appearing in both the Sanskrit and English forms). In a sense then hundreds of such Sanskrit neologisms are an attempt to represent English concepts semantically but through portraying them in neo-Indic phonological guise. This is more in keeping with early twentieth-century nationalistic and literary sensibilities than would have been the simple inclusion of barely assimilated English forms. Nonetheless, sensibilities shift and recent decades have not avoided a substantial influx of direct English borrowing (Senawong 1992). A torrent of such loans now imubes ‘pop’ culture, teen slang, sports reporting and technology. Commercial establishments like tailor shops frequently display transliterated English names. Human nicknames like ‘Nut’, ‘Golf’ and ‘Bird’ are analyzed by Nacaskul (1987), who clarifies the phonology of such loans; even dogs get English names. The humanities and law, on the other hand, tend to eschew direct English borrowing and such loans are regularly absent from proclamations, constitutions, etc. For the English-to-neo-Indic process, see Wan Waithayakorn (1970); Prasithrathsint (1994). Gandour (1979) and Bickner (1986b) discuss the complicated issue of how tone is assigned to English loans. Thai-English bilingualism is studied by Cefola (1981) and by Davis and Schoknecht (1994). Psycholinguistic studies of Burnham, et al. (1992) are of relevance here as well. In reverse, Cohen (1987) provides an entertaining look at expatriate foreigners’ acquisition (or not) of Thai in Bangkok.

3.4.4. Comparative studies

Several studies compare standard Central Thai with other Tai varieties of Thailand, some with focus on attitudinal factors: Northern (Lanna) Thai (Pankhuenkhat 1976; Nokaeo 1989); Yong (Davies 1979); Northeastern Thai (Palikupt 1983).

Comparisons with Khmer (Cambodian) include focus on predicates by Martini (1956, 1957); descriptions of remarkable syntactic parallelism by Nacaskul (1971) and by Huffman (1973); study of complement constructions by Poo-israkij (1995); and diachronic lexical analysis Varasarin (1984). Khanittanan (2001) and Diller (2003) develop diachronic proposals related to these comparative studies.

Comparative classifier studies are instructive in several ways. Jones (1970) presents a landmark survey of Southeast systems. It establishes areal patterning of classifier syntax, especially among Tai languages: classifiers follow their head noun as in Thai to the west and south, but precede it, as in Tai varieties like Nung or Zhuang to the east and north. (As for an isogloss, a leading Black Tai linguist, the late Cam Trong, told me that both patterns are used in his language for different purposes. In fact, most speakers of Central Thai admit some flexibility in special contexts, such as the regular [phi: 42-no’:ng45 so’:ng24 khon33] ‘two siblings’ as compared to the more restricted [so’:ng24 khon33 phi: 24-no’:ng45] ‘both of them, the siblings’. As for particular lexical items, work of Jacob (1965) indicates that Khmer has supplied Thai with several classifiers and vice-versa. Conklin (1981) is a comparative study encompassing Tai (including Thai) and Austronesian classifiers. In a cross-language survey of classifiers and language standardization, Barz and Diller (1985) examine evaluative and normative feelings about classifiers, found to differ across Indo-Aryan and Southeast Asian languages. For Thai especially, sociolinguistic constraints are fine-tuned and crucial in many Thai classifier selections, as established by Juntanamalaga (1988). Thai is placed in a comparative Tai context by Burusphat (2007), who presents a comprehensive treatment of animate classifiers; see also Morev (2000).


Thai-English comparisons often have an applied-linguistics focus. Pioneering studies include Kruatrachue (1960) on phonology and Chaiyaratana (1961) on syntax. A sample of others includes studies on pronominalization, considered by Chomaitong (1976); on definiteness, by Lamchote (1971); on aspect, by Noochoochai (1978); and on intonation, by Kanchanathat (1977). For applied studies featuring error analysis, translation problems or difficulties Thai speakers have with specific features of English, representative studies are by Ariyapitipun (1988), Meemeskul-Martin (1984), Palmer (1969), Richards (1968), Schmidt (1988), Suwattee (1981), and by Van Syoc (1964).

3.4.5. Historical and philological studies

Diachronic sources relating to Tai family and to Tai-Kadai are mentioned in section 1.1. This and the next subsection mainly mention works with a focus specifically on (standard and/or Central) Thai, or else the language of Sukhothai; however, controversial hypotheses tracing how Central Thai has developed over the past seven centuries depend crucially on comparative-historical Tai background to provide plausible starting points for presumed Thai-internal diachronic changes. These hypotheses also refer to orthographic history. Even though some may devalue study of written sources as being institutionally ‘philology’ rather than linguistics, I see no good reasons to disregard orthographic evidence when trying to unravel Thai diachronic sound change.

See Anuman Rajadhon (1961, 1981), Hartmann (1986a), Court (1996) and Diller (1996b 2001a) for historical hypotheses as to how Indic-based orthographies and literary culture, including Thai writing, have developed and spread in the Southeast Asian context. The inventor(s) of Thai writing certainly had Khmer orthography in mind, and perhaps Mon, but significant innovations were introduced as well. These included tone-marking, horizontal rather than vertical representation of clusters, phasing out of redundant vowel-initial graphemes and the creation of new segmental symbols as needed. The latter were usually accomplished through modifications made to existing Khmer-type letters representing similar sounds; thus a new [f] symbol was an enhancement of a given [ph] letter. The impression is strong that the original impetus for Sukhothai script involved great care and attention to Tai sound-system detail along with a semiotic attempt to represent perceived phonetic closeness through graphical similarity. Result: diachronic linguists should take this orthography seriously.

Although it is clear that Thai orthography in general can be traced back to South Indic scripts, intermediate points are still professionally debated. The tradition that Thai writing originated in the reign of King Ramkhaeng of Sukhothai (r. approx. 1279-1298) has been challenged but to date no material evidence of a different origin has been adduced. Also, the historical relations of Thai and Lao scripts are somewhat contentious, although existing material evidence strongly indicates that a Sukhothai-type script had spread east to Lao-speaking areas by the early sixteenth century and constituted the prototype for standard Lao writing. Mon writing, on the other hand, was the basis for Lanna (Northern Thai) script, currently under resuscitation, also known to the east as tham (i.e. ‘dharma script’, as it was used for Buddhist texts). Discussion and further sources are included in Danvivathana (1987).
and in Diller (1996a). Hudak (1986) describes official simplifications in Thai spelling during the 1940’s, later rescinded.

For those with interest in how lexical, syntactic and semantic aspects of the Thai language has evolved since the Sukhothai era of about seven centuries ago, a good place to start is with the commentaries and texts of Na Nagara and Griswold (1992). They present and discuss key literary sources of the Sukhothai period, although their perspectives are mainly philological and historical. A difficulty encountered in some philological work is the tendency to assume that Sukhothai texts are a direct reflection of earlier stages of Central Thai, or even coincide phonetically with modern Thai, whereas Sukhothai Thai probably stood in a less direct line with it. Note that Brown, 1962, 1965, considers Southern Thai varieties to be the more direct descendants of Sukhothai Thai. To their credit, Na Nagara and Griswold (1992) frequently cite cognate material from non-Central dialects. Weroha (1992) too provides comparisons with local varieties. Mikami (1984) summarizes other features of Thai of the early period, while Bamroograks (1987) makes significant progress in understanding Sukhothai discourse patterns. Prasithrathsint (2007), in a convincing comparative and diachronic study of nominalization, shows that at least the type with prefixal \([khwa:m\ A2]\) ‘matter (of...);’ was present in the language of Sukhothai, with nominalization on the increase in succeeding centuries.

For the complex diachronic development of the pronominal system, Strecker’s (1984) treatment of Proto-Tai pronouns is a good place to start, with dialect studies such as Filbeck (1973b) useful to keep in mind. Studies like these indicate that the Thai system is the result of substantial diachronic shifting and innovation. Iamchinda (1992, in Thai) is a seven-hundred year survey covering the period over which written data on pronouns are available.

Jones (1971) makes accessible King Chulalongkorn’s important essay on Thai titles and ranks, a study nicely complemented by Tingsabadh and Prasithrathsint (1986, in Thai), who analyze the use of address terms over some two centuries.

Etymology, areal contact and comparative-historical semantic shifting are insightfully treated by Matisoff (1986, 1992), who analyzes kinship terms and analogues of Thai ‘heart/mind’ compounds in \([cay\ 33]\). These are considered in the wider Southeast Asian context, however issues are raised that need to be kept in mind when focusing on semantic shifts more specifically in Thai. An especially noteworthy study of the latter type is that of Khanittanan and Placzek (1982), who show how the inherited item \([khwan\ 24]\) ‘spirit, soul’ underwent semantic reorganization when an Indic term \([win\ 33ya:n\ 33]\) with similar meaning was introduced in Buddhist contexts. For more on etymology, see Burnay and Coedès (1920); Li (1956, 1971, 1977); Anuman Rajadhon (1961, 1981); and Thai-language sources such as Na Nakhon (1973).

Turning to diachronic phonology and phonetics regarding segmentals, we can surmise that the vowel system of Central Thai is an area where Khmer comparisons are merited, but inscriptive Khmer is not without its own interpretive challenges (Jacob 1965, etc.). Tai-internal accounts of earlier vowel inventories and of how they have evolved into the current Central Thai system have led to divergent proposals, especially regarding vowel-length (Sarawit 1973; Hartmann 1976b; Li 1977; Brown 1979; Strecker 1983; Luksaneeyanawin 1992). In any event, Sukhothai orthographic vowel distinctions show affinities both with Angkorian inscriptive Khmer and also with modern Thai, even though the varieties may not be in direct linear relationship. Dhananjayananda (1997) calls attention to one difference: what is now the [e]/[e:] distinction is not marked until the seventeenth century. The same general picture perhaps applies to long-short distinctions in low front and back vowels [ae] and [o’] as well. But does this mean that the distinctions were pronounced earlier but not marked in text until later, or that new contrastive articulations originated subsequently as well? Diachronic aspects of [ay] and [aw] sequences are elucidated in work of Bickner (1992). Finally, as L-Thongkum, Teeranon and Intajamornrak (2007) have convincingly established, diachronic consideration of vowel length change and of tonogenesis, considered below, must be considered as intertwined issues.
Work on Thai syntactic shifts has been less controversial but has a potential contribution to make to typological studies. Analyses of Sukhothai discourse by Bamroograks (1987) and Prasithrathsint (2007) have been mentioned. Grammaticalization, mentioned in 3.6 above, has been an important process in Thai diachronic syntax but I am not aware of a comprehensive text-based study organizing just how all relevant changes have occurred historically. Khanittanan (1987a, 1987b, 1988b) and Prasithrathsint (1988, 1996) have produced leading work in establishing how syntactic patterns of written Thai have shifted in the past two centuries or so, whether qualitatively or quantitatively. Prasithrathsint has traced in great detail the increase of passive-like constructions and nominalizations in written sources. Khanittanan, using a succession of royal prose compositions, identifies a number of features that characterized evolving prose writing of the mid-nineteenth century: zero anaphora of understood subjects; topic-initial sentences; paratactic constructions rather than marking with overt conjunctions; lists with quantifiers in final position. These generally reflect features today associated with spoken Thai (Messenger 1980; Chodchoey 1988; Hinds 1988a).

Over time, a new style of formal written Thai has taken shape, characterized by denser nominalization, clausal embedding and other types of subordination and more overt specification of nominals, rather than leaving zero anaphors to be construed. This has coincided with increased normative interest in specifying what is ‘correct’ Thai (Diller 1993, 2001a). Lexical selections also play a role here, with higher diglossic choices and technical Indic vocabulary characteristic (Wan Waithayakorn 1970; Prasithrathsint 1994). Resulting in what Khanittanan (1988b) refers to as a more ‘autonomous’ style, discourse of this type is comparatively depersonalized and even aloof from Thai interpersonal social dynamics. ‘Autonomous’ Thai does not encode a range of interpersonal communicative factors the way the lower colloquial style typically does with particles, finely calibrated address-reference selections and other choices. This style of written Thai is currently maintained, but as mentioned in subsection 3.4.2 above, written Thai genres also admit a more colloquial type of Thai as well.

Contrasts between these written styles are sometimes manipulated for special effects, such as to produce eye-catching journalistic headlines in a lower more oral style, while providing ‘serious’ content reporting in the higher more autonomous style (Khanittanan 2007). Also, some Thai authors are adroit in exploiting the difference through stark contrasts in descriptive passages versus dialogue; Chat Kopchitti would be an example. On the other hand, many educated Thai speakers (academics in particular) can, when occasion requires it, produce a spontaneous type of oral discourse showing many features of the higher written style. A methodological consequence of this syntactic vibrancy seems to be that native Thai speakers may disagree with each other as to grammaticality judgments (Diller and Khanittanan 2002). One can hope that future scholarship will contribute to more understanding of these interacting oral and written genres and of their role in the constitution of Thai linguistic competence.

3.4.6. Tonogenesis and its quandaries

In diachronic phonology, one area where an understanding of Tai-wide issues impinges crucially on the more parochial history of Central Thai concerns tonal development. Also, since some of the crucial evidence generally accepted for sound changes is from written texts, matters of orthographic interpretation and development become involved as well.

The Thai writing system has been subject to diachronic phonology in some ways and resistant to it in others. Methodologically, it has been taken as important tool in uncovering and analyzing sound changes relating to tone, thus meriting linguistic concern. Most authorities posit an original (or Proto-Tai) system of three tones for open syllables and limited tonal options for stopped syllables; in general, they took one of the open tones (Li 1977). A sweeping devoicing process is thought to have subsequently occurred affecting voiced obstruents, along with loss of aspiration in formerly aspirated sonorants. For Thai (and Lao), the devoiced stops further acquired aspiration. Lexical items with the newly-arising initials were kept from merger by consonant-induced allophonic tonal distinctions, which then...
became phonemic. That is, the new unaspirated fully-voiced sonorants (e.g. /m/ < *hm) and aspirated voiceless obstruents (e.g. /ph/ < *b) did not produce homophonic lexical mergers with vocabulary items with the original unaspirated sonorants (e.g. /m/ < *m) and aspirated stops (e.g. /ph/ < *ph). This complex shift is thought occurred after the codification of (Sukhothai) Thai writing, probably in the fourteenth and fifteenth centuries.

Some evidence for this change consists in spelling variation during the period and in environmental loans introduced before the change took place, such as Malay *durian, sago* becoming Thai [thu45 rian33], [sa:24 khu:33] (Suthiwan 1992). Further evidence is the fact that Thai (following Sukhothai) orthography regularly represents the shifted sonorants with a digraph headed by (now) silent initial h- symbol, e.g. as hm, hn, hl; hence [luang24] ‘large, grand’, written as though kluang to reflect the aspirated sonorant initial of the thirteenth century. If the initial sound had been a simple voiced sonorant or something else at the time, it is difficult to explain why that particular digraph was selected for the relevant vocabulary. To clinch the argument, there are also some early Thai loans into nearby Austroasiatic languages that preserve pre-shift initials like /hm/ (Gedney 1965).

In this scenario, the effect was to bifurcate the original tonal system. For Central Thai, devoicing was accompanied not only by aspiration of relevant stops but also by other sound changes in which phonation type of initial consonant affected tone. New tonal mergers took place as well. The result is the three classes of letters referred to as ‘high’, ‘mid’, and ‘low’, with different tonal effects. Central Thai happens to have remerged in such a way as to distinguish five phonemic tones. Since the Thai writing system is widely considered to reflect the situation before these major shifts, interpreting its present complexity is in effect a matter of taking into account the results of the complex diachronic sound changes. It also means that the Thai and Lao writing systems are abstract in the sense that they can be interpreted in various ways to accommodate tonal systems of local dialects in Thailand and Laos, a fact with consequences for literacy policy.

A consequence of the changes sketched above is that in the first instance (the Sukhothai period) Thai tone markers [may45 e:k11] and [mai45 tho:33] would have had direct phonetic interpretations, whereas after the changes, their interpretation would become phonologically abstract leading to the variable values the markers now exhibit with different groups of consonants. The remaining two tone markers were added after the shifts mentioned above were complete, as can in fact be detected from their restricted distribution, although the physical shape of markers has been subject to shifting (the former cross symbol used for [mai45 tho:33] is now used for [mai44 cat1 awaa33]). Speculations as to the origin of the tone markers and how their interpretation must have shifted are offered in Diller (1996b), who considers Sanskrit chanting marks. Further perspectives range from the earlier work of Brown (1962), Jones (1965, 1966) and recent observations of L-Thongkum, Teeranon and Intajamornrak (2007).

A complication in the preceding model has been known for years, e.g. as discussed by Gedney (1947, sections 175-177), Brown (1965, section 2.3) and others. Not directly accounted for in the tonogenesis-devoicing process sketched above are Indic-provenance items like [bun33] ‘merit’, ultimately from Pali puñña. These items would seem to imply a voicing process for stops p > b and t > d (but not c > j or k > g). Note that this would work in opposition to the devoicing changes described above. If voicing for ‘merit’, etc., had originally occurred beforehand in an intermediate language like Old Khmer, why then was ‘merit’ not caught up in post-Sukhothai tonogenetic devoicing? On the other hand, if it was still pronounced with initial p- in Sukhothai times, why is it [bun33] today, while inherited Tai vocabulary like [pa : 11] ‘jungle’ does not show the same change? Solutions to this have differed. Maybe puñña had been previously assimilated in Old Khmer with a preglottalized initial such as [p bun] and it was this that was adopted by Sukhothai (bilingual?) speakers, treating it as having the same initial [p] as in Tai words like [pa:n6] ‘village’. This escapes the problem for Thai by casting it back onto Old Khmer (why did the Khmers treat Indic p- this way?). The case seems not to be closed.
Many have contributed to the development of the tonogenetic line of research, from earlier scholars such as Bastian (1868), Bradley (1909a, 1909b) and Burnay (1927) on to professional publications more linguistic in orientation of Haudricourt, Gedney, Jones, Li and others. Brown (1962, 1995, 1975), Erikson (1975), Rischel (1984), L-Thongkum (1997), Thurgood (2007b), and L-Thongkum, Teeranon and Intajamornrak (2007) and sources therein present detailed articulatory hypotheses concerning tonogenesis. Anivan (1988) considers phonetic shifts in Central Thai tones in more recent times after establishment of the basic tonogenetic patterns.

Gedney (1985) discusses the vexed question of timing for the main tonogenetic changes. Loan-word evidence from Malay (Suthiwan 1992) was mentioned above for taking the Ayudhya-era fourteenth and fifteenth centuries as the period of changes for Thai. (Malay contact is more convincing for this line of argument than similar indications in Pali and Sanskrit loans, as at earlier stages the latter have come into Thai second-hand through Khmer or Mon, raising extra diachronic possibilities.) If this timing hypothesis is indeed the case, then a troublesome methodological issue arises. The majority of Tai varieties would need to have shared the same basic devoicing and tonogenetic sound changes: a great phonological upheaval over a wide and disparate area in a short time. Only a few rather isolated varieties of the Sino-Vietnamese frontier area are reported as having escaped systematic devoicing (Haudricourt 1949; Ross 1996; L-Thongkum 1997). Zhuang dialects far removed in time, place and social dynamics from Central Thai would have had to undergo similar shifts. Following a normal application of the Comparative Method, without considering orthographic or loan evidence, one would need to project the changes much further back in time than the fourteenth century (see Jones 1965, 1966; Gedney 1985). Here indeed is an instance where the proto-Tai reconstructions that are accepted are of paramount importance in establishing how (Central) Thai has changed since the Sukhothai period. Is the Sukhothai starting-point positioned before or after the sweeping changes?

To make matters more complex, some, but by no means all, nearby Austroasiatic languages with classical writing systems or relevant evidence of other sorts appear to exhibit similar devoicing shifts, some involving vowel-related and/or phonation-related register phenomena. This includes Khmer, certainly critical in Thai’s linguistic history. Further consideration of these sound-change issues is clearly merited. General consideration of the substantial Austroasiatic, and specifically Khmer, impact on Thai has been mentioned: see Nacaskul (1971), Huffman (1973), Varasarin (1984), L-Thongkum (1992), Khanittanan (2001) and Diller (2002). See also chapter 7.

REFERENCES


RESOURCES FOR THAI LANGUAGE RESEARCH


62  THE TAI-KADAI LANGUAGES


— (1989a) Selected Papers on Comparative Tai Studies (Michigan Papers on South and Southeast Asia 29), Center for South and Southeast Asian Studies, University of Michigan.


Gullstrøm, Carl Friedrich (1832) Remarks on the Siamese Language by the Rev. Mr Gullstrøm, communicated through the late Robert Fullerton, London: Cose; also in Siam Repository (1869) I: 51-58.


— (2007c) ‘Notes on Problems in the Phonetic Descriptions of Dental and/or Alveolar Stop Strictures’, in Jimmy G. Harris et al. (eds), 105-111.


Indigenous Languages of Thailand Research Project (1977) *Bibliography of Tai Language Studies*, Bangkok: Central Institute of English Language, Office of State Universities.


Jacob, Judith M. (1965) ‘Notes on the Numerals and Numeral Coefficients in Old, Middle and Modern Khmer’, *Lingua* 15: 143-162.


— (1980) ‘A Question or not a Question: This is a Question’, paper presented at the 13th International Conference on Sino-Tibetan Languages and Linguistics, Charlottersville: University of Virginia.


Kuo, William (1982) Teaching Grammar of Thai, University of California at Berkeley, Center for South and Southeast Asian Studies.


Morev, Lev N. (1964) Osnovy sintaksia taiskogo iazyka [Fundamentals of Thai Syntax], Moscow, Nauka.


Na Nakhon, Pluang (1973) *Panha Phasa Thai* [Thai Language Problems], Bangkok: Rum San.


— (1975) ‘How Useful are Citation Form in Synchronic Thai Phonology?’, in Jimmy G. Harris and James R. Chamberlain (eds), 274-284.


— (1978) *An Introduction to Thai Syntax*, Bangkok: Mahidol University Southeast Asia Center.


REFERENCES FOR THAI LANGUAGE RESEARCH


Thongkum, Luangthongkum: see L-Thongkum.
Upakit-Silapas, Phraya (Nim Kanchanachiwa) (1939) Lak phasa Thai [Principles of the Thai Language], republished 1979, Bangkok: Thai Wattanaphanit.


