

**Punctuation
and other text-category indicators
in written Thai text:
Issues and implications for Thai L2
Reading instruction**

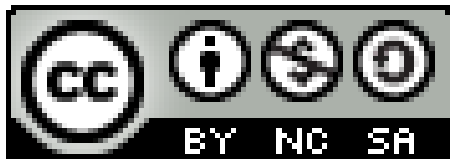
**Yuphaphann Hoonchamlong
University of Hawaii'i at Manoa**

yuphapha@hawaii.edu

SEALS 23

Chulalongkorn University

May 30th , 2013





Outline

- Characteristics of Thai orthography and writing convention
- Punctuation & “text category indicators” in Thai
- Discussion & demonstration of
 - challenges for L2 learners learning to read Thai
 - Proposed reading strategies for L2 learners learning to read Thai




About Standard Thai

- Tai language family
- Tonal language- 5 tones in Standard Thai
- SVO and topic-comment sentence structure
- “isolating/analytic” language (no inflection/conjugation)
- C(C)V(C) syllable structure
- Derivation (compounding and reduplication) as dominant and productive morphological process
- A “pro-drop” language, dominance of zero anaphora in inter- and intra-sentential reference
- “discourse-oriented” language



characteristics of Thai orthography

- ❖ script developed from "Grantha", an ancient Brahmi script of South India, since 1283 A.D.
- ❖ alphabetic (script represents all the phonemes of speech)
 - ❖ fairly systematic letter–sound mapping
- ❖ the vowel symbol(s) are in fixed position(s) around the initial consonant 
- ❖ tone markers to represent tones, the position of the tone marker is on top of the initial consonant.

Thai Vowel positions

In relation to initial consonants

1. “before” /ee/ เ อ (F) เบน
2. “after” /aa/ อ า (F) บาน
3. “above” /a/ อ̣ ำ F นัก
4. “below” /u/ อ̣ ุ F ทน
5. Combination form

/aw/ เ อ ำ เขา

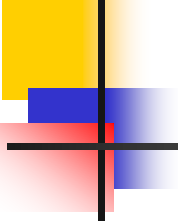
/ia/ เ อ ีย F เวียน





Thai orthographic convention

- **Direction:** write from left to right
- **Punctuation**
 - no space among words/morphemes
 - rarely use punctuation marks to mark clause and sentence boundaries
 - space may be used to separate words in lists, before and after numbers, clauses or sentences but can be random
- **No upper-case/lower-case distinction**



Example of Thai writing convention (using English to demonstrate)

howtofryanegg

overmediumheat warmasmallamountofoilorbutter inasmallskillet
crackanegginto thepanandcookuntilthewhiteappearssolid
about 3 to 4 minutes

How to Fry an Egg

Over medium heat, warm a small amount of oil or butter in a small skillet. Crack an egg into the pan. Cook until the white appears solid, about 3 to 4 minutes.



Punctuation (1)

- A set of non-alphanumeric characters that are used to provide *information about structural relations among elements of a text*, including commas, semicolons, colons, periods, parentheses, quotation marks and so forth. (Nunberg 1990: p.17)
- Other graphical features/devices of the text:
 - font & face alternations, capitalization, indentation
 - “null element”: spacing

Text-category indicators



Punctuation (2)

- Text = a passage of written language
- Text categories (units of written text)
 - Paragraph
 - Text sentence
 - (dash-interpolation)
 - (parenthetical)
 - Etc.



Functions of text category indicators

1. **Delimiters**: mark one or both ends of an element
2. **Separators**: are inserted between elements of the same category type
3. **Distinguishers**: typographically set off an element of a particular type from its surroundings



Punctuation and text information (1)

- Indicates grammatical features of text, e.g.
 - **apostrophe** (') marks possessive: John's book
 - **period** (.) marks end of a written "text sentence"
 - **question mark** (?) indicates a direct question when placed at the end of a sentence
 - **comma** (,) indicates a separation of ideas or elements within the structure of a sentence
- **Space**: indicates word boundary
- **Capitalization**: indicates the beginning of a sentence. (and names)



Punctuation and text information (2)

- Indicates phonological features of text
 - **comma** (,) indicates pausing & intonation patterns

A language that uses word spaces and punctuation marks (such as English) presents the reader with a text **pre-analyzed into words, clauses and sentences**

History of Thai punctuation

■ Old Thai punctuation marks

◎	ตาไก่ ฟองมัน	U+0E4F	marks the beginning of a short statement. marks the beginning of a stanza. (poetry)
◉	ฝนทองฟองมัน		marks the beginning of a new section of text. (prose)
๐๓	โคมุต	U+0E5B	marks ending of a story.
๐๓๐๓			
๐๑	อังกั้นเดี่ยว ขั้นเดี่ยว		- marks ending of a sentence. - marks ending of a stanza (poetry)
๐๑๑	อังกั้นคู่	U+0E5A	marks ending of a section or a story. (prose) marks ending of a stanza (poetry)
๐๑๐	อังกั้นวิสรรชนี		marks ending of a sentence or a section (prose) marks ending of a stanza (poetry)



A	฿ 0E0A	บ 0E1A	ส 0E2A	◌̣ 0E3A	๓ 0E4A	๗ 0E5A
B	฿ 0E0B	ป 0E1B	ท 0E2B		๓ 0E4B	๗ 0E5B
C	ฒ 0E0C	ผ 0E1C	ฬ 0E2C		๓ 0E4C	
D	ญ 0E0D	ฝ 0E1D	อ 0E2D		๓ 0E4D	
E	ฎ 0E0E	พ 0E1E	ฮ 0E2E		๓ 0E4E	
F	ฏ 0E0F	ฟ 0E1F	๗ 0E2F	฿ 0E3F	๓ 0E4F	

Unicode Standard 6.0 THAI



Modern punctuation in Thai (1)

- 1835 D.B. Bradley brought Siamese printing press to Siam from Singapore
- 1844 “The Bangkok Recorder”- first Thai Language newspaper; “paragraph indentation” was used.
- Western educated Thai elites started using “European” punctuation in their Thai writing
- 1900 Ministry of Education published a “how to” textbook on modern punctuation



Modern punctuation in Thai (2)

- 1917 King Rama VI advocated the use of the new punctuation as detailed in 1900 book, proposed spacing between words
- 1981 The Royal Institute set up a committee to review punctuation in Thai
- 1987 RI published a punctuation guideline



BUT.....

We still see Thai text written like this....

howtofryanegg

overmediumheat warmasmallamountofoilorbutter inasmallskillet
crackanegginto thepanandcookuntilthewhiteappearssolid
about 3 to 4 minutes

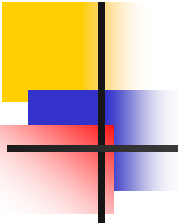


Why?

- Thai makes use of “lexical” text category indicators
 - “connectives” (relators, conjunctions)
 - “punctuation words”

- Single connectives:

และ (and), แต่/แต่ว่า (but)

- 
- Set connectives
 - Giving examples

เช่น *aaa bbbb ccc* (และ) *dddd* (เป็นต้น)

ได้แก่ *aaa bbbb ccc* (และ) *dddd* (เป็นต้น)

ไม่ว่าจะเป็น *aaa bbbb ccc* (หรือ) *dddd*

for example *aaa bbbb ccc* and *dddd* etc.



- Time relation:

เมื่อ ---E1----- (แล้ว) E2 ก็+Verb----

When ---E1-----, ---E2-----

- Concession (contrast): variant forms

ถึงแม้ว่า---E1----- (ก็ตาม) (แต่) E2 (ก็)+verb

แม้ว่า ---E1----- (ก็ตาม) (แต่) E2 (ก็)+verb

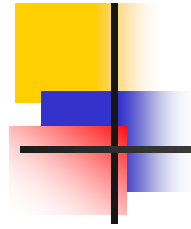
ถึง ---E1----- (ก็ตาม) (แต่) E2 (ก็)+verb

Even though ---E1---, --E2-----

- Time period

ตั้งแต่ ----(Time 1)----- (จน) ถึง ----(Time 2)-----

since ----(time 1)----- until ----(time2)-----



Reading in L2



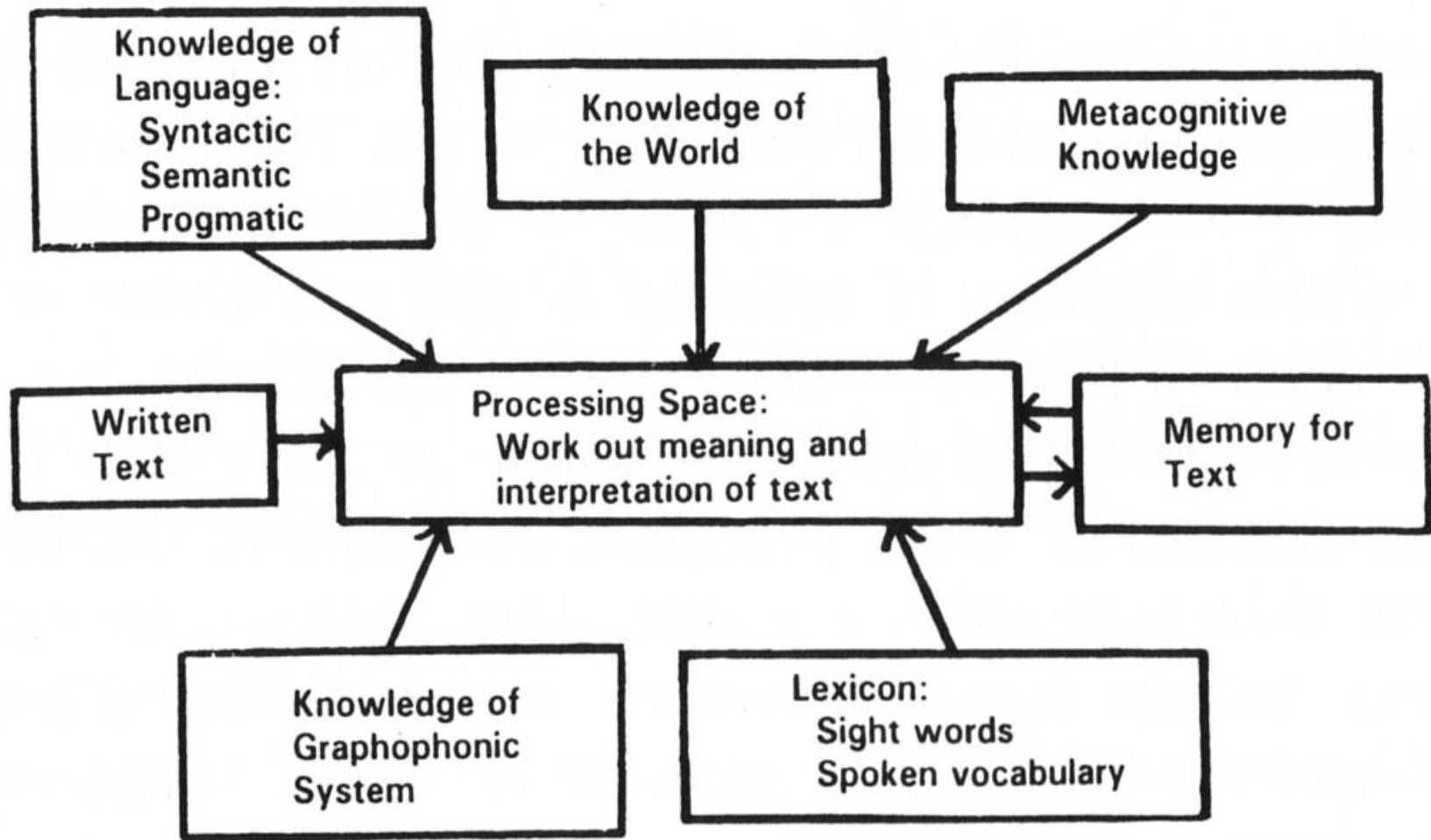
Issues in reading/literacy

- Reading entails the extraction and integration of linguistic information printed in text (Koda 2008)
 - how meaning is linguistically coded
 - how linguistic information is graphically represented (phonological/morphological)
- Reading involves a linkage/mapping of
 - (spoken) language and
 - writing system



Cognitive processes in reading

- decipher the symbols(graphemes),
word division/identification
(lower level)
- comprehend the meaning of the
graphemes and the message conveyed
(higher level)
 - word meaning
 - text meaning (clause, sentence,
discourse)



interactive model of reading proposed by Rumelhart (1977)
as adapted by Ehri (1998: 6)



Challenges for L2 learners in learning to read Thai

Lower level:

- ❖ learn to "decode" or "decipher" unfamiliar scripts
- ❖ learn to identify syllable boundary

Higher level

- ❖ learn to identify word boundary and word meaning
- ❖ learn to identify clause/sentence boundary
- ❖ Learn to identify "zero" pronouns and other ellipsis



Proposed strategies for TFL learners

- Start with “paragraphs”, the unit of Thai text with clearly defined boundary.
- Pre-intensive reading process
 - Identify potential clause and sentence boundaries (and their relations), using connective words/sets as the main clue (along with spaces)
 - Identify main verbs (or verb clusters)
 - Identify boundaries of unknown words from strings of unknown characters by looking for potential doublets and compounds
 - Identify potential names (as opposed to words)
 - Observe sentence and paragraph structure



Examples



คุณเจริญและคุณมาลี สายทอง สามีภรรยา มีลูกด้วยกันสองคน คนโตเป็นผู้ชาย ชื่อจรรุญ ชื่อเล่นว่าจัน คนเล็กเป็นผู้หญิงชื่อมาลัย มีชื่อเล่นว่าแจง จันอายุ ๑๑ ขวบ แจงอายุ ๔ ขวบครึ่ง เสาร์วันหนึ่งคุณเจริญและคุณมาลีฝากลูกไว้กับคุณยาย แล้วออกไปซื้อของที่ตลาดนัดสวนจตุจักร ที่นั่นมีของขายหลายชนิด มีทั้งอาหารและของใช้ต่างๆ

- 1 คุณ-เจริญ-และ-คุณ-มาลี สายทอง สามี-ภรรยา-มี-ลูก-ด้วย|กัน-สอง-คน//คน-โต-เป็น-ผู้ชาย
- 2 ชื่อ-จรรุญ / ชื่อเล่น-ว่าจัน// คน-เล็ก-เป็น-ผู้หญิง/ชื่อ-มาลัย/ มี-ชื่อเล่น-ว่า-แจง// จัน-อายุ 11 ขวบ/
- 3 แจง-อายุ 4 ขวบ-ครึ่ง// เสาร์-วัน-หนึ่ง-คุณ-เจริญ-และ-คุณ-มาลี-ฝาก-ลูก-ไว้-กับ-คุณ-ยาย แล้วออก-ไป
- 4 ชื่อ|ของ-ที่-ตลาด|นัด-สวน-จตุจักร// ที่|นั่น-มี-ของ-ขาย-หลาย-ชนิด/ มี-ทั้ง-อาหาร-และ-ของ|ใช้-ต่าง|ๆ/
- 5 ราคา-ก็-ไม่|ค่อย-แพง

- 1 'Khun' Charoen and Khun Malee Saithong husband-wife have two children altogether. // The elder one is male.
- 2 (he) is named Charoon/, nick-named Chan// the younger one is female/, named Malai / (she) has nickname Chaeng// Chan is(age) 11 years old/
- 3 Chaeng is(age) 4 and a half years old. // One Saturday Khun Charoen and Khun Malee left the children with Grandma then (they) went out to
- 4 shop at Jatujak swapmeet. // Over there there are many things for sale/, including various food and utensils/
- 5 The price is not very expensive as well.

นักท่องเที่ยวชาวต่างประเทศกล่าวว่า กรุงเทพฯเป็นเมืองที่น่าเที่ยวชมเมืองหนึ่ง
 เพราะกรุงเทพฯมีสถานที่ที่น่าเที่ยว น่าสนใจหลายแห่งเช่น วัดพระแก้ว วัดโพธิ์
 วัดเบญจมบพิตร วัดอรุณฯ พระบรมมหาราชวังและตลาดนัดสวนจตุจักร เป็นต้น

- 1 นัก-ท่อง/เที่ยว-ชาว-ต่าง-ประเทศ-**กล่าว-ว่า** “กรุงเทพฯ-**เป็น**-เมือง-**น่า|เที่ยว-น่า|ชม**-เมือง-หนึ่ง
- 2 **เพราะ**-กรุงเทพฯ-**มี**-สถานที่-**น่า|เที่ยว** น่า|สนใจ-หลาย-แห่ง-**เช่น** วัด-พระแก้ว วัด-โพธิ์
- 3 วัด-เบญจมบพิตร วัด-อรุณฯ พระ|บรม|มหา|ราช|วัง-**และ**-ตลาด|นัด-สวน|จตุจักร **เป็นต้น**”

- 1 Foreign tourists **said that** “Bangkok **is** one of the cities good for sightseeing
- 2 **because** Bangkok **has** many interesting places to sightsee **such as** Temple (Wat) aaa, Wat bbb,
- 3 Wat ccc, Wat ddd, eee **and** Jatujak swapmeet, **for example.**”

นัก- ชาว- น่า+verb x-ใจ
 เที่ยว ท่อง+เที่ยว สถานที่ สน+ใจ น่า-สน+ใจ
 น่า-เที่ยว+น่า-ชม





Proposed Strategies for reading Thai (1)

- Identify phrase, clause, sentence boundary
 - Need to know “connectives” (relators, conjunctions), their variant forms and their positions in a clause and learn them as “sight words”
 - Use “connectives” as “lexical” text category indicators (“punctuation words”)

Strategies for reading Thai (2)

- Identify the position and function of “zero” pronouns and other elliptical units

เที่ยว-เมืองไทย ไม่-ไป-ไม่-รู้

Travel Thailand not go not know

เที่ยว เมืองไทย ~~ถ้า~~ ~~เรา~~ ไม่ไป ~~เที่ยวเมืองไทย~~ ~~เรา~~ ~~ก็~~ ไม่รู้

Travel Thailand. If one doesn't go (travel Thailand), one won't know.

- Need to know both syntactic structure and expected position of “connectives”

Strategies for reading Thai – word level

(1)

- Identify syllable boundary
 - Use cues from vowel forms to determine position of initial (I)/final (F) consonants

I ๗ (F)	บ้าน
๒ I ๗	๒๒๗
๒ I F	๒๓๓
๒ I ย F	๒๓๓๗

- Use cue from tone marks to determine I position (tone mark on I)

[Discussed in J. Marvin Brown's AUA Language Center Thai Course Reading and Writing: Mostly Reading]

Thai Vowel positions

In relation to initial consonants

1. “before” /ee/ เ อ I (F) เอน
2. “after” /aa/ อ า (F) บาน
3. “above” /a/ อ̂ I F น้ก
4. “below” /u/ อ̣ I F ทน
5. Combination form

/aw/ เ อ อ เอ

/ia/ เ อ ย F เียน





Thai script examples

เขาเป็นนักเรียนทุน

เขา เป็น นัก เรียน ทุน



ยายมานาน

ยาย มา นาน

น้องชอบขนมอบกรอบ

น้อง ชอบ ขน^๓ม^๓อบ^๓กรอบ

Strategies for reading Thai- word level (2)

- Identify (make an educated guess) of (potential) word boundary
 - Need to know about word morphology of Thai compounds and reduplications  
 - lexical items in Thai seems to be largely bisyllabic compounds
 - Use maximum collocation approach

a. ฟิต-เพ็ล็กซ์-เลือก-ทำ-ธุรกิจ-เกี่ยว-กับ-**ข้อ-ต่อ**-และ-สาย-อ่อน-ทุก-ชนิด

b. ที่-นาย-ปิ่น-กล่าว-มา- เป็น-เพียง-ข้อ-อ้าง-หรือ-**ข้อ-ต่อ-ผู้**-หนึ่ง-เท่านั้น

c. ออส-เตร-เลีย-อาจ-หยิบ-ยก-เรื่อง-นี้-ขึ้น-มา-เป็น-**ข้อ-ต่อ-รอง**-ปิด-ตลาด-ใน-บาง-สิน-ค้า-ของ-ไทย

ข้อ-ต่อ (joint)

ข้อ-ต่อ-ผู้ (argument)

ข้อ-ต่อ-รอง (bargaining point)



Thai compounds (1)

4 main types of compounds

1) Regular compounds (คำประสม)

เครื่อง + บิน เครื่องบิน

machine fly airplane

2) Semantic “doublets”*(คำซ้อน): compounds with components that share semantic similarity (synonyms, near-synonyms, same semantic field), typically components are mono-lexemic/mono-syllabic

บ้าน + เรือน บ้านเรือน

house house

* **Vongvivapanond, Peansiri E. 1992. "Lexicological Significance of Semantic Doublets in Thai."** in Papers on Tai Languages, Linguistics, and Literatures in Honor of William J. Gedney on his 77th Birthday. Compton, C. J. and J. F. Hartmann, eds. DeKalb: Center for Southeast Asian Studies. Monograph Series on Southeast Asia, Occasional Paper No.16.

Thai compounds (2)

3) Phonetic “doublets”: a non-meaningful syllable added to a mono-syllabic/lexemic word with similar phonetic features (as a filler to make a 2 syllable word)

มาก + มาย mâak-**maay** a lot

กว้าง + ขวาง kwâaŋ-**khwǎaŋ** wide

4) Elaborate Expressions: multi-syllable (typically 4) expressions based on the above compound types by inserting syllable(s) for the sake of rhyme (such as English artsy-craftsy) or syllable(s) that have some semantic relations.

เจ็บ+ป่วย pain+sick cèp+pùay

เจ็บ ไข้ ใต้ ป่วย cèp khây **dây** pùay

pain fever get sick
pain fever get sick



References

- Alderson, C. (2000). *Assessing Reading*. Cambridge: Cambridge University Press.
- Cook, V., & Bassetti, B. (2005). An introduction to researching second language writing systems. In V. Cook & B. Bassetti (Eds.), *Second language writing systems*. (pp. 1-67). Buffalo, New York: Multilingual Matters Ltd.
 - Foster, P., Tonkyn, A., & Wigglesworth, G. (2000). Measuring spoken language: A unit for all reasons. *Applied Linguistics* 21, 354-375.
 - Halliday, M. A. K., & Hasan, R. (1976). *Cohesion in English*. London: Longman
 - Koda, K. and A. Zehler. (eds.) (2008). *Learning to Read Across Languages: Cross-Linguistic Relationships in First- and Second-Language Literacy Development*. New York: Routledge.
 - Nunberg, G. (1990). *The Linguistics of Punctuation*. Stanford: Center for the Study of Language and Information.



In Thai

- Royal Institute. (2008). *Guidelines for usage of punctuations and other symbols; guidelines for using space; guidelines for writing abbreviations*. 7th printing. Bangkok: Aroon Printing.
- Boon Intharamphan. (2002). The Use of punctuation marks in Thai according to King Vajiravudh's proposal. In *Journal of the Royal Institute of Thailand*. Vol 27. No. 2. Apr-Jun 2002, p. 586-591.
- Nidhi Eoseewong. (2002). Punctuations. In *Words as weapon: on Language, culture and power*. Pp. 31-38. Bangkok: Matichon.



sa-wat-dii
สวัสดี

Q&A



Thai Language Program

<http://thai.hawaii.edu>



[http:// www.hawaii.edu/thai](http://www.hawaii.edu/thai)