
Feedback on the EAGLES proposal from Asian languages

Chu-Ren Huang (Chinese)

Elizabeth Zeiton (Formosan)

Paul J.K. Li (Formosan)

Dipti Misra Sharma (Hindi)

Tokunaga Takenobu (Japanese)

Hae-Yun Lee (Korean)

Asanee Kawtrakul (Thai)

Thatsanee Charoenporn (Thai)

Virach Sornlertlamvanich (Thai)

Methodology

- ▶ Distribute the EAGLES documents
 - ▶ EAGLES Document EAG-CLWG-MORPHSYN/R May 1996
 - ▶ EAGLES Document EAG-CLWG-SYNLEX August 1996
- ▶ Test applicability on each language (NOUN and VERB)
 - ▶ By comparing existing linguistic resources (lexicon, tagged corpus etc.)
- ▶ Comparison and summarization

NOUN (EAGLES)

- ▶ Type (L1): comp/prop
- ▶ Gender (L1): m/f/n
- ▶ Number (L1): sg/pl
- ▶ Case: (L2a) nom/gen/dat/acc (L2b) voc/ind
- ▶ Countability: cou/mass

NOUN – Type –

- ▶ Common noun/proper noun are applicable
- ▶ Additional class
 - ▶ dependent (Korean) ?
 - ▶ compound noun (Japanese)
 - ▶ classifier (Japanese, Korean, Thai)
 - ▶ Languages not to distinguish sg/pl tend to use classifier
 - ▶ Chinese: *liang3 zhi1 gou3*
two CLS dog
 - ▶ Korean: *kai twu mali*
dog two CLS
 - ▶ Japanese: *ni hiki no inu*
two CLS p dog
 - ▶ Thai: ?

NOUN – Gender, Number, Case, Countability –

- ▶ These attributes are not applicable to all languages except Hindi.
- ▶ Hindi needs attribute Animacy (animate/inanimate).
- ▶ Case is marked by particles (postpositions) in Japanese and Korean.

Japanese:	Mary- ga	John- ni	kisu-sita
Korean:	Mary- ga	John- eykey	Khisu-hayssata
	Mary-NOM	John-ACC	kissed
- ▶ No explicit case marking in Chinese. Corpus markup may prefer thematic/argument role instead of case.
- ▶ Case is marked by position and meanings in Thai.

VERB (EAGLES)

- ▶ Type: main/aux/modal
- ▶ Finiteness: fin/nofin
- ▶ Verb form/Mood: ind/sub/impr/cond/inf/part/ger/sup
- ▶ Tense: pres/impf/fut/past
- ▶ Person: 1/2/3
- ▶ Number: sg/pl
- ▶ Gender: m/f

VERB – Type –

- ▶ main/aux/copulative are applicable.
- ▶ Additional class
 - ▶ support verb (Korean, Japanese, Hindi)
 - ▶ derives denominal verbs
 - ▶ Korean: kongpu-**hata** (study do)
 - ▶ Japanese: benkyo-**suru** (study do)
 - ▶ compound verb (Hindi) : should be separate entry?
 - ▶ Distinction between state/action is important in Thai.
- ▶ Modal is usually expressed by auxiliary verbs (???)

VERB – Finiteness –

- ▶ Not applicable

VERB – Verb form –

- ▶ Verbs conjugate in Korean and Japanese. But it's form is determined by the succeeding word.
- ▶ Verbs do not conjugate in Chinese and Thai (isolating languages).

VERB – Mood –

- ▶ Mood is expressed by auxiliaries and particles (Japanese, Thai) and suffixes (Korean).
- ▶ The classification of mood is more semantic oriented (Japanese ???), some of them are expressed by modal verbs in English.

VERB – Tense –

- ▶ Tense is expressed by other than verb conjugation such as auxiliaries, suffixes and adverbs in many languages.

VERB – Person, Number, Gender –

- ▶ Not applicable except Hindi.

VERB – Aspect –

- ▶ Aspect is expressed by auxiliaries and suffixes.

VERB – Voice –

- ▶ Attribute values active/passive are applicable.

VERB – Reflexivity –

- ▶ Not applicable to all languages.

VERB – Main verb function –

- ▶ Attribute values transitive/intransitive are applicable to all languages except Thai.
- ▶ In Chinese, state/active marking is more useful.

General comments

- ▶ There seems to be clear contrast between inflectional language and others.
- ▶ Since many Asian languages are not inflectional, information marked by inflection is expressed by auxiliaries, particles and affixes.
- ▶ Information to be described in a lexicon and how it is marked in surface representation should be clearly distinguished. In the EAGLES proposal, the classification of information seems to be affected too much by surface representation (inflection). For example, we would have consensus on attribute values of case, but how it is realized depends on languages.

Proposals

- ▶ New class "classifier" should be created under the noun type. This could be Level 1.
- ▶ Affixes (currently classified in Residual) should be promoted to the first class category. In particular, affixes play very important role in Formosan language.
- ▶ Classification of adposition should be more precise for agglutinative languages (Korean and Japanese). This could be Level 2b.
- ▶ Classification of honorific system would be necessary at Level 2b (Korean and Japanese).

Example (Japanese)

- ▶ **Situation:** There were two sisters among girls. A boy passed by them, and the elder sister told the younger sister the utterance:
- ▶ “He is a high school boy.”
- ▶
 - “(Kare-wa)= ϕ kôôtôgakkô-no gakusei-san (desu)= ϕ yo”
he-ACC highschool-GEN student-HON is MOOD
- ▶ Ending postposition “yo” vs “ne”
 - ▶ Represeting a mood to inform something to the hearer.
 - ▶ “yo”: The speaker assumes that the hearer does not know it.
 - ▶ “ne”: The speaker assumes that the hearer knows it already.
(confirmation)