Lexicalization of syntactic constructions in Thai

Amara Prasithrathsint, Ph.D
Chulalongkorn University, Thailand
Amara.Pr@chula.ac.th

ABSTRACT

In isolating languages, such as Thai, new words are mostly created through the process of compounding. Compound words are generally defined as words that are composed of old words the meanings of which interrelate in such a way that a new meaning comes out which is very different from the meanings of the words in isolation, e.g., *mɛ̂ɛ-náam* (mother-water) 'river,' *tûu-yen* (cupboard-cold) 'refrigerator.' However, there are certain structurally ambiguous constructions, which could be interpreted as compound nouns, noun phrases, or even sentences, e.g., *khon-khàp-rót* (person-drive-car) 'driver/a person who drives a car,' *sàt-kin-nɨ́a* (animal-eat-meat) 'meat-eating animal/ an animal that eats meat/ the animal eats meat,' *nǎŋsɨ̌ɨ-khǎay-dii* (book-sell-well) 'best seller/ books that sell well/ the book sells well,' *khrɨ̂aŋ-bin-tòk* (plane-fall) 'plane crash/ the plane crashed.' The question is how to draw a boundary between such compounds and syntactic constructions. So far there has been no study that focuses on this phenomenon in Thai. Therefore, this study aims to analyze such ambiguous constructions in order to find out how to decide whether they are lexical items, phrases, or sentences. Based on a three-million-word corpus of current Thai, the study reveals that the ambiguity results from certain typological characteristics of Thai, such as S-V-O, Head-Modifier, non-inflectional, no subject-verb agreement, no marking on finite verbs. Thus, any string of [N+V] can be a compound noun, a noun phrase (N+ modifying reduced relative clause), or a sentence (N or NP + finite verb). It is also found that only semantic criteria, which have been generally used, are not sufficient to separate lexicalized words from noun phrases and sentences. Syntactic criteria need to be taken into consideration, too. In sum, an [N+V] construction is considered to be a compound noun if it has a specialized or an idiomatic meaning, and that no word can be added to the construction without changing its meaning. The findings of this study support the theory of **lexicalization**; i.e., a new lexical item may derive from a formerly freely composed, grammatically regular, and semantically transparent phrase or sentence.