Aspects of Tagalog Compounding

Videa P. De Guzman

The University of Calgary

1. Introduction

In the face of a variety of sometimes conflicting proposals for a theory of morphology (notably, Aronoff 1994; Anderson 1992; Lieber 1983, 1992; Carstairs-McCarty 1992; Botha 1984; Selkirk 1982), the nature of Tagalog compounding is explored in this study. It attempts to determine the syntactic, morphological, semantic and pragmatic aspects of this process and determines how best to account for compound words in Tagalog, a major language of the Philippines. The approach used is a lexicalist one in the sense of locating morphological and semantic rules of compounding in the lexicon component of the grammar, where both derivational and inflectional rules are also contained (See De Guzman 1991). The view taken here is that the lexicon contains the set of lexical items that speakers know. New items borrowed or created may be added any time.

It will show that Tagalog compound words have different structures. Like most compound words in other languages, they are typically made up of two words (roots or stems) of the same or different lexical categories, some combinations being more productive than others, and some combinations being rare or even non-occurrent. Similar to French where some compounds show de joining the two words, another frequent form of compounds in Tagalog consists of two words with a linker -ng between them. It will be interesting to find out what triggers these types of prejudices in category combinations as well as their differences in form and meaning.

Despite the claimed lack or absence of syntactic categories in Tagalog, but for words and particles (Gil 1993; Himmelmann 1991; Bloomfield 1917), I will proceed using the three major or open class lexical categories N(oun), V(erb) and A(djective) argued on the basis of semantic and morphological grounds (see De Guzman 1996) not on their syntactic distribution. There appears to be no other systematic and fruitful way of describing the structure of compound words and the relations that exist between their constituents, and even of other complex words for that matter, without referring to such categories. It seems to be nonsensical to say that any two words can combine to form a new compound word, when there are differing relations that exist between the constituents of varying combinations. Moreover, how are the more productive word combinations to be distinguished from those that are less productive. In addition, the meanings of the individual members that carryover to the compound words cannot be generalized accordingly without considering their word classes, not to mention the strict order in which the words combine.

2. In Defense of Lexical Categories

Tagalog is claimed to have no syntactic categories (Gil 1993; Himmelmann 1991; Bloomfield 1917) because any word, which have been traditionally identified as a Noun, a Verb, or an Adjective, may generally occur with each of the identified particles. Thus, the only two classes to be distinguished are words and particles. This claim,
however, does not clearly point out that the same word when used after the major syntactic particles ang, ng (pronounced nang) or sa always takes on the meaning of a nominal. To illustrate: an action word such as takbo ‘run’ when used as a command to someone or as a predicate means ‘Run!’ which in no uncertain terms is telling someone to perform an action. On the other hand, when it appears after the particle ang, it now refers to ‘someone’s or something’s running’. For example:

(i) masamá? ang takbó ng kotse/aso
    bad run car/dog
    ‘The way the car/dog runs is bad.’ (Lit., ‘The running of the car/dog is bad.’)

A typical complex form of the same word, tumakbó ‘run; ran’, with the active voice verb affix -um-, may also occur after the above-mentioned particles, but in such constructions it refers to ‘the person/thing that runs/ran’. For a descriptive word such as payát ‘thin’ which typically occurs in either the predicative or the modifier position means ‘the thin one’ when it appears after the particles above. For example:

(ii) itapon mo ang payát
    throw.away you thin
    ‘Throw away the thin one.’

If the meaning of a word, independent of such particles, is affected by its occurrence after these particles, it is not at all convincing that the word has not changed its syntactic function. To ignore these observed differences in meaning and function does not help in making an adequate account of word categories.

Considering a non-distributional or non-syntactic way of classifying words, De Guzman (1996) shows the viability of maintaining at least the three major classes of words, namely, N(oun), V(erb) and A(djective). Support for the distinctions comes from their semantic and morphological properties. Following Croft’s (1991:53ff) cross-linguistic characterization of the major syntactic categories, the semantic classes Object, Action and Property correlate with the categories N, V and A, respectively. Morphologically, these three categories may occur in either root or complex forms. They undergo morphological processes such as affiliation, reduplication, stress shift or a combination of these. Ns are usually simple roots. In complex forms, they have the affixes (a) -an referring to a place for N, e.g. bigás ‘rice’ >—> bigás-an ‘rice bin’; (b) ka- -an referring to some abstract notion pertaining to N, e.g. buhay ‘life’ >—> ka-buhay-an ‘livelihood’.

There is a whole host of complex Ns that are derived from Vs as well as As (See Schachter and Otanes 1972 for a comprehensive listing of derivations). In contrast, Vs, both roots and stems, typically occur with the voice affixes -um-, -an, -i- and -an, indicating which cooccurring complement of the verb is marked with the ang particle in a verbal clause structure. They also form paradigms expressing three aspects. For example, for takbó, we have the active, non-finite form tumakbó. Its finite forms expressing aspect are: tumakbó ‘completed’, tumatakbo ‘incompleted’, tatakbo ‘contemplated’. And this is true for the other verb forms with the other voice affixes. The V form with the voice affix -an, in contrast to the Ns with the same affix form, has the syntactic consequence of having a cooccurring locative nominal as the ang phrase. To illustrate:

(iii) pupunta/hán ng doktór ang pasyente
    will.go.locative doctor patient
    ‘The doctor will go (see) the patient.’ (Lit., ‘The patient will be gone to by the doctor.’)
It may also be pointed out that some of these forms may also be derived into Ns, sometimes with a stress shift, and be associated with a location or place where something is V-ed. For example:

(iv) saan ang puntáhan ng mga tao kung Sabado
where going.place plural person particle Saturday

o Linggo?
or Sunday

‘What place do the people frequent on Saturday or Sunday?’ (Lit., ‘where is the going-place of the people on Saturday or Sunday?’)

There are other morphological processes that certain classes of Vs may undergo to express reiteration or repetition of what the V denotes, which processes do not apply to Ns and As. Other forms reflect plurality of the agent performing the action or they may be associated with the meaning ‘intensive’. What is interesting is that these processes also apply to denominal and deadjectival Vs.

Descriptive or property words categorized as A may also be either roots or complex words. The most common affix is ma- attached generally to abstract Ns to mean having or characterized by what is expressed by N. For example: payá ‘thin; slim’, tuyó? ‘dry’; ma-gandá ‘pretty’, ma-tapang ‘brave, courageous’. It is significant that certain unaffixed As have V or N correspondences that are identical in form, except for the difference in their stress pattern. For example: Vs with penultimate stress basag ‘break’, ayos ‘arrange’, gamit ‘use’ correspond, respectively, to As with ultimate stress baság ‘broken’, ayós ‘arranged’, gamít ‘used’; Ns with penultimate stress buhay ‘life’, gutóm ‘hunger’, galít ‘anger’ correspond to As with ultimate stress buháy ‘alive’, gutóm ‘hungry’, galít ‘angry’, respectively. Another important distinguishing feature of As is the ability to express differing degrees or intensities of the property denoted by the root or stem through affixation, reduplication or the occurrence of specific particles. For example: payá ‘thin’, mapayá-t-payá ‘moderately thin’, kasyag-payá/t-payá ‘very thin’, kasingpayá ‘as thin as’, pinakapayá ‘thinnest’. For a more comprehensive account of the semantic and morphological distinctions among N, V, A and their derivations, the reader is directed to De Guzman 1996.

Having established these three major lexical categories and having referred to the linker =ng above, whose meaning is equivalent to that of the relative particle na or the possessive particle ng, we cannot overlook another particle that may also occur in compound words. It is the conjunctive particle at ‘and’ often appearing as a cliticized =t. We can now refer to these lexical categories to describe the nature of compounding in Tagalog.

3. The Nature of Tagalog Compounding

A compound word is defined here as a word, which bears a lexical category, that consists typically of two or more words. A word is a lexical entry that is either a root or a stem (a root plus affix(es)). The meaning of a compound word usually derives from the meanings of its constituents, drawn from the semantic and/or pragmatic relations that exist between them. Certain compounds created unsystematically may bear meanings beyond the meanings of their constituents. Tagalog exhibits compound expressions of varying combinations using the major categories N, A and V. It will be noted that there are also those where a linker (abbreviated in the examples below as LKR) or a conjunctive particle between the two constituents has to be present.
3.1 Types of compound words

Four types of compounds are identified in the following subsections, namely, endocentric, exocentric, synthetic and copulative.

3.1.1 Endocentric (or headed) compounds

Between the two words that usually form a compound word, one is said to function as the head and the other the non-head. Tagalog is typically left-headed (Cf. Williams 1981 and Lieber 1983, 1992); the category of the new compound word is typically the same as that of its head. Depending on the category of the head, the function of the non-head in relation to its head can be determined, and given this relation, the meaning of the whole compound can typically be drawn. With a N as head, the non-head is usually a modifier or qualifier which restricts the reference of the head N. It can also function as the head’s complement. Given the three lexical categories that may logically combine, the possible combinations with a N resulting in N compounds are: N + N, N + A, and N + V. Of these, N + V is non-occurrent; N + N is the most productive of all possible combinations (as observed in other languages as well) and N + A is moderately productive, contrary to what is observed in English whose analogue is A + N. The examples below (as well as in the succeeding subclasses) are given with some indication of their productivity as shown by the number of items listed under each pattern.

(1) N + N

(a) tubig + ulan
water + rain
‘rain water’

(b) tanod + bayan
guard + town
‘policeman’

(c) bunga=ng + kahoy
fruit=LKR + tree/wood
‘fruit (from trees)’

(d) bata=ng + lansangan
youngster=LKR + street
‘street kid’

(e) suka=ng + Iloko
vinegar=LKR + Iloko
‘Iloko vinegar’

From the combination of two Ns, we can draw different kinds of meanings reflected in the compounds. One of them is a N that is modified or specified by the following N. In 1(a), for instance, tubig-ulan ‘rain water’ indicates what kind of water it is, i.e. tubig ng ulan ‘water of/from the rain’. Examples l(b) and l(c) express some kind of

---

1 This identity is attributed to what Lieber (1983) labels feature percolation principle. Not only the category feature but also the argument features of the head, if any, percolate up to the complex word.
possessive attribution equivalent to the phrasal expressions *tanod ng bayan* ‘guard of the town’ and *bunga ng (puno=ng) kahoy* ‘fruit of trees’, respectively. Another relation expressed in both l(d) and l(e) is that of location or source. In l(d), *bata=ng lansangan* ‘street kid’ defines the child/youngster as coming from or frequenting/living on the streets. It could be understood, too, as ‘child belonging to the street’, hence, depicting a possessive relation. Expressed syntactically, this compound may be delivered as either *bata ng lansangan* ‘child of the streets’ or *bata sa lansangan* ‘child on the streets’. The last compound, l(e) *suka=ng Iloko* ‘Iloko vinegar’ gives the source or place of origin of the head N. Along with this meaning is the implicit connotation that it is the best kind of vinegar. Thus, it also means *suka ng Iloko* ‘Iloko’s vinegar’ or *suka sa Iloko* ‘vinegar from Iloko’.

At this juncture, it may be instructive to point out that when the first element of a compound as an independent word ends with a vowel or, more accurately, ends in final /h/, a glottal stop, /ʔ/, or the dental nasal /n/, it is linked to the following constituent with the connector =ng, which is equivalent in function to the relator form na. The phonological outcome is the deletion of these final consonants and the attachment of the linker.

Another source of N + N compounds shows one member as a complex form. For example, *pa-aral-án[N =ng]LKR + bayan[N] ‘public school’; ka-bungguʔ-án[N =ng]LKR + balikat[N] ‘a close acquaintance’ (Lit., ‘a person that one rubs shoulders with’). What these complex forms reveal is the capacity of one category to be derived into the same or another category, in these cases through affixation. The root or base in the examples are the Vs *aral* ‘to study’ in the first and *bungguʔ* ‘to hit; to slam’ in the second. The derived N words get related semantically to some other words of the same category and form compounds that are suitable for use in given situations.

(2) $N + A | N$

(a) buhay + mayaman

life + wealthy

‘life of leisure; rich life’

(b) bató=ng + buháy

stone=LKR + alive

‘smooth, hard stone; livingstone’

(c) dugó=ng + mahál

blood=LKR + dear/expensive

‘royal blood’

The second subclass consisting of N + A is clearly of the descriptive type. Having an A that specifies the kind of N in the compound, 2(a) *buhay-mayaman* ‘wealthy/rich life’ expresses ‘a kind of life (style)’. Thus, this pattern can be further seen in: *buhay-mahirap* ‘life of poverty; poor life’; *buhay-masaganaʔ* ‘life of prosperity’, etc. The same is found in 2(b) *bató=ng buháy* ‘a very smooth kind of hard stone; livingstone’ where the stone is

Note that *kahoy* in the compound *bungang-kahoy* is an abbreviation of the compound *puno=ng kahoy* ‘tree’.

This is similar to English compounds such as Idaho potatoes, Alberta beef, Washington apples, etc.
described as being alive equivalent to the phrase *buháy na bató*. In 2(c), *dugó*=ng *mahdál* ‘royal blood’ also describes what kind of blood, where the A *mahdál* conveys not just ordinary blood, but being ‘dear, expensive’, it must be of a superior type or class, thus of royal stature. Compared with its syntactic analogue, the expression is usually of the form *marangál na pamílya* or *pamílya*=ng *marangál* ‘honorable/noble/respectable family’. It may be instructive to remember that whereas the order of the head and its modifier is usually freer or reversible in syntax, it is always fixed in compound words.

An interesting form consisting of N + A, whose meaning can be deduced from the meaning of its members, does not bear the category of its head. Semantically, the compounds are descriptive expressions, hence, they belong to the category A.

(d)  
isip + matandá?  
mind/thought  old/adult  
‘mature’ (Lit., ‘mind of an adult’)

(e)  
isip + bata?  
mind/thought  young/child  
‘immature’ (Lit., ‘mind of a child’)

Such examples confirm the existence of compounds whose category does not derive from the supposed left-hand head. In other words, the syntactic feature of the head fails to percolate to the compound word. It appears that it is the category of the non-head which is acquired by the compound.

The next category of compounds based on the possible combinations of the three major categories is that headed by A. The combination A + A is extremely rare. The only observable pattern is A + N. The other very rare combination is A + V; rare because similar to the pattern N + V, V is neither a typical modifier or specifier nor a complement. With A as the left-hand member of the compound, it appears to be the modifier of the following N. Unless this non-head N is taken as the A’s complement, the resulting compound cannot be a legitimate compound bearing the category A. Based on their meaning, however, there is no question about their being categorized as A. Consider the following examples with A + N combination:

(3)  
A + N]A

(a)  
kapós + palad  
short.of palm.of.hand/fate  
‘unfortunate’

(b)  
tigíb + dusa  
full/laden grief/suffering  
‘grievous’

---

4 This pattern will be discussed later under the type called copulative compounds.

5 It has been suggested in Lieber (1983) via the argument-linking principle that N + V and A + V are non-occurrent because the non-head element V cannot be the head’s specifier. There is a good chance, however, that V may be lexically derived into another category say N or A. The rare compound A + V noted here will be treated under the heading exocentric compounds.
With the compounds 3(a)–(c), their syntactic correspondence with approximately identical meanings are: kapós ng/sa palad ‘lacking in fate/fortune’, tigíb ng/sa dusa ‘laden with grief/suffering’, mababa-ng(na) loób ‘will that is low; humble or meek’.

Compare the above examples with the following items which are taken from Schachter and Otanes (1972:111) categorized as compound As (without any category label for each constituent) and cited in Lieber (1992:46) in which these are analyzed as consisting of A + N A:

(d) amóy + isdá?
   smell/odor fish
   ‘fishy smelling’ (‘fish odor’)

(e) lasa=ng + isdá?
   taste=LKR fish
   ‘fishy tasting’ (‘fish taste’)

It may be argued that the first words in these two examples are both Ns and not As, as indicated by the gloss I provided in parentheses. These meanings indicate that the non-head N on the right specifies the head, for 3(d) ‘having the odor of fish’ and for 3(e) ‘having the taste of fish’. In both, the kind of ‘odor’ and ‘taste’ are specified as being that of ‘fish’. It may be said that these Ns take the following N as their complement. In syntactic terms, the compounds in 3(d) and 3(e) run parallel to amóy ng isdá? ‘smell of fish’ and lasa ng isdá? ‘taste of fish’, respectively, keeping the relation of possession intact. These forms can productively be expanded by substituting different Ns in the second constituent whose semantic features are congruent with those of the given heads, e.g. amóy surot/suka?/pawis ‘bedbug/vinegar/sweat odor’; lasa-ng kapé/pátis/tuyó? ‘coffee/fish sauce/dried-fish taste’.

Some other compound words which are unarguably A + N come from expressions denoting time of day. For example:

(f) madalí=ng + araw
   quick=LKR sun/day
   ‘dawn; early morning’ (Lit., ‘quick day’)

(g) hatí=ng + gabí
   half=LKR evening/night
   ‘midnight’

(h) dapít + hapon
   over.there/late afternoon
   ‘twilight; late afternoon’

Both amóy ‘odor’ and lasa ‘taste’ are also categorized as Ns in Diksyunaryo-Tesauro Pilipino-Ingles by J. V. Panganiban (1972).
In 3(f)–(h) the right component N is modified by the first element A and refers to a particular point in time and with this meaning, the compound can only be categorized as N. For 3(f), madalí? (na lang) at araw na ‘soon (already) and it’ll be day’ is a good paraphrase that refers to the point in time labeled ‘dawn’. On the other hand, 3(g) hating-gabi ‘midnight’ is marked by ‘being half of the evening/night’, equivalent to the phrase kalahati ng gabí ‘one-half of the night’. In 3(h) the first constituent A is virtually obsolete, except in this context. Contrary to the prediction of Lieber’s Licensing Conditions for Tagalog and her claim that “Tagalog in fact has no compounds which are either syntactically or semantically right-headed” (Lieber 1992:43, 47), this last set of examples, in addition to those cited previously, shows that the right element can also function as head.

When the head is a V, the cooccurring non-head is usually its complement. This is to be expected because each lexical entry V in the lexicon, if transitive, would be specified as requiring a cooccurring complement N. Thus, the pattern V + N is quite a productive form of compound, unlike its analogue N + V in English. But V + A is non-existent, the reason having been mentioned above that A is not a typical complement of V. The pattern V + V will be discussed along with A + A, both patterns considered as falling under the copulative type of compound. Below are examples of the common V compounds:

(4) \[ V + N \]

(a) alsal + balutan
    lift/raise            baggage
    ‘to evacuate or move out suddenly’

(b) magdalá=ng + hiyá?
    carry=LKR           shame
    ‘to be embarrassed’ (Lit, ‘to carry/bear shame‘)

(c) magdalá=ng + tao
    carry=LKR           person/human being
    ‘to be pregnant’ (Lit., ‘to carry/bear a human being‘)

(d) magbigáy + galang
    give              respect
    ‘to show respect’

(e) magbigáy + pugay
    give              praise
    ‘to offer praise’

(f) magbango(n)=ng + puri
    stand.up/lift=LKR    honor
    ‘to redeem one’s honor’

It will be noted from the above examples that the V constituent may be a simple root or a complex form — a stem marked by a voice affix. The complement N usually functions as the object of the V and the corresponding phrasal structure would have the typical ng particle before the object complement in the same order when the verb is
marked with the active voice affix m- attached to a pag-stem of the V. Thus, *magbigáy ng salu-salo* ‘to give a feast/party’, *magdalá ng pagkain* ‘to take some food’, etc. Note, however, that these syntactic phrases do not have corresponding compound words. It could be because these are such commonplace expressions, there is no need to formulate new (compound) words in their place.

The various combinations of compounds attested by the examples above show that they differ in structure from the corresponding syntactic phrases that express generally the same meanings. When the meaning of the compound is no longer transparent and hence, unpredictable from the meanings of its constituents, we are now dealing with the type of compounds labeled *exocentric*.

### 3.1.2 Exocentric (headless) compounds

The exocentric type of compound has forms identical to the endocentric ones. Syntactically, most combinations appear to have a head and a modifier or a complement, but the functional relation between the constituents do not necessarily carry over to the semantic interpretation of the whole compound. For example, a V + N form such as *hanap* ‘search’ + *buhay* ‘life’ does not in any way mean ‘to look for life’ but rather it is a nominal meaning ‘occupation; job’. In this instance, we see that the category of the supposed left head does not percolate to the compound. It may be suggested that this N compound takes after the category of the following complement since we have seen in previous examples that the category of a complement may be the same as the category of the compound. However, a closer analysis of the individual meanings of the two constituents together brings out the semantic content of the whole compound. Literally, the compound means ‘to search or look for life’. The purpose of this search is actually ‘to sustain life’ and in order to accomplish it one has to work and earn a livelihood. Thus, one’s *hanap-buhay* is that person’s answer to his/her search for life’s sustenance. This indicates that with this type of compounds the semantic content transcends what the individual meanings of the constituents denote. Rather, in most cases, the meaning can usually be gleaned from what the first word asserts and its pragmatic relation to the meaning or to a particular characteristic embodied in the meaning of the second member. Typically, the modifier or complement element contributes its connotative more than its denotative meaning, and the category the compound takes depends on its composite meaning. Other exocentric compounds, on the extreme end of the scale, have meanings that are completely unpredictable, better described as lexicalized. Such properties explain why exocentric compounds are referred to in general as semantically as well as syntactically headless. Let us look at some examples following the identical structures given under the endocentric type. Note, however, that the category of some of the resulting compounds in each pattern is based on their meanings.

\[(5) \quad \text{N} + \text{N|N}\]

- \[(a) \quad \text{dalaga=ng} \quad + \quad \text{bukid}
\]
  - maiden=LKR
  - farm

  ‘red snapper’ (Lit., ‘farm maiden’)

---

7 The category label of the compounds have some morphological implications as well which have not been shown here due to certain limitations.

8 The literal meaning is the other meaning of this ambiguous compound, i.e., ‘a maiden who grew up in the farm’.
In the first case, 5(a) is a lexical item which should be distinguished from the identical endocentric compound that means ‘a maiden of/from the farm’. It may be suggested though that the pretty color of this fish makes it just as attractive as many beautiful farm maidens. The following two examples, 5(b) and 5(c), show the relation of complementation and possession, yet the meaning of the second element cannot be a direct guide in determining the relevant meaning of each compound. 5(b) requires an insight into the nature of ‘rig-drivers’ who, being on the road practically all day long, are in contact with a variety of passengers ranging from simple common folks with lots of stories of all kinds to tell-tales or gossipers. In view of this, whatever he hears and later relates to someone else is second-hand information or plain hearsay, and therefore unreliable and is not usually taken seriously. In 5(c) the implicit understanding that ‘child’ here refers to an ‘unborn baby’ provides the appropriate interpretation of ‘house’. Once again, this is another instance of the unsystematic nature of exocentric compounds when it comes to interpreting them. The speaker somehow knows which semantic and pragmatic features are paramount in treating the relation between the elements.

(6) \[ N + N \] 

(a) asal + hayop 
behavior animal 
‘rude; ill-mannered’

(b) balát + sibuyas 
skin onion 
‘sensitive’

(c) bibíg + anghél 
mouth angel 
‘prophetic’

(d) isip + lamók 
mind/thought mosquito 
‘stupid’

(e) boses + ipis 
voice cockroach 
‘soft-voiced’

(f) matá=ng + lawín 
eye=eLKR eagle 
‘sharp-eyed’
In the first two examples, the discernible relation between the constituents is that of modification or specification. 6(a) speaks of ‘a kind of behavior’ and 6(b) ‘a kind of skin’. Even in this relation, the more common usage of these compounds is with the meaning carried by them as exocentric compounds. Pragmatically, the characteristic of the second N that is being addressed by the first N is implicit. For example, in ‘animal’, it is its being ‘wild’ that is prominent rather than its usefulness, its size, strength or capabilities; in ‘onion’, its being smooth, thin or delicate not its taste, shape, or effect on one’s eyes when being sliced. As such though, 6(a) is used to describe a person who is very rude, whose behavior is no less than despicable, and 6(b) to one who is ‘sensitive; whose feelings get hurt easily’ in contrast to its opposite, someone who is balát-kalabáw ‘insensitive; Lit., skin of a carabao (thus, thick-skinned)’. On the other hand, even with the obvious possessive relation obtaining in 6(c) corresponding syntactically to bibíg ng anghel ‘mouth of an angel’ and 6(d) to isip ng lamók ‘mind of a mosquito’, these meanings have to be reinterpreted further. From the speaker’s pragmatic knowledge, reference to ‘angels’ or ‘mosquitoes’ relative to the heads ‘mouth’ and ‘mind’, respectively, the meaning drawn is that in 6(c) whoever displays ‘prophetic’ prowess is described as being bibíg-anghel. It is believed that angels only speak the truth or what is surely to come. In 6(d), the mosquito’s being tiny and insignificant is considered the primary, relevant feature that contributes to the meaning ‘mind/thought’ which must not count much. The fact that it can be pesky or that its bite can be the cause of some serious illness does not even figure in the semantic interpretation. Thus, the meaning associated with the compound isip-lamók is that of the A ‘stupid; having very little gray matter, if at all’. Similarly, the compounds in 6(e) and 6(f) require knowledge of the precise features of the N possessor that characterize its relation with the first N. Thus, in the former, one knows to pick out the quiet, undetectable movement of cockroaches whereas in the latter, it is the eagle’s keen vision, not its powerful wings.

(7) \[ A + N | N \]

(a) \[
\text{bago}=ng + \text{taón}
\]
\[ \text{new=LKR} + \text{year} \]
‘New Year’

(b) \[
\text{hulí}=ng + \text{habilin}
\]
\[ \text{last=LKR} + \text{request} \]
‘last will’

(c) \[
\text{patáy} + \text{gutom}
\]
\[ \text{dead} + \text{hunger} \]
‘a vagabond’

(d) \[
\text{kusa}=ng + \text{loób}
\]
\[ \text{voluntary=LKR} + \text{inside} \]
‘initiative’

In all of the above examples, the left-hand element A modifies the following N. Given their exocentric reading, the first two have each a specific referent, with 7(a) referring to ‘New Year’s day’ and 7(b), to the ‘last will’ of a person before or near death. As endocentric compounds and syntactic phrases (with the same strict order required), 7(a) refers to any or an entire ‘new year’ and 7(b), any ‘final request’ made among others not necessarily by a dying person. In 7(c) the person addressed as patáy-gutom
could not be gleaned directly from the equivalent phrase *patay (na) ng/sa gutom* ‘(almost) dead because of hunger’ or its endocentric reading as A ‘dying of hunger’, which can, of course, be attributed to anyone who may have simply missed a meal or two. The situation of a ‘vagabond’ is, however, always characterized as ‘dead hungry’. In 7(d) a person’s ability to act without being told is somewhat related to the first constituent, but instead of being muddled by the following N, it becomes more emphatically connected, i.e. the volition comes from within. No syntactic phrase corresponds to this compound.

(8) \( A + V \)A

(a) \( bago=ng + ahon \)
\( \text{new=LRK get.on.shore/land} \)
‘ignorant’ (Lit., ‘newly landed’)

This pattern is quite a rare combination and occurs only as A. The literal meaning given above is seen to have been demoted to a negative status. Replacing the V with the V *dating* ‘to arrive’ will not be perceived as a similar compound as 8(a); it has the interpretation ‘newly arrived’ said of someone who has just arrived. Yet, the borrowed Spanish words *alsá* ‘lift; rise’ or *saltá* ‘to disembark; to land’ are acceptable substitutes with exactly the same meaning.

The following pattern is just as productive as its endocentric counterpart. The meaning observed, however, which is associated with each compound is uniquely nominal.

(9) \( V + N \)N

(a) \( basag + ulo \)
\( \text{to.break head} \)
‘a fight/brawl’

(b) \( hanap + buhay \)
\( \text{to.search.for life} \)
‘occupation; job’

(c) \( buká=ng + bibíg \)
\( \text{to.open=LRK mouth} \)
‘sayings; common expression’

(d) \( hampás + lupa \)
\( \text{to.strike ground} \)
‘a vagabond; a tramp’

(e) \( pasá(n)=ng + krus \)
\( \text{to.carry.on.one’s.back=LRK cross} \)
‘a heavy burden, usually referring to some deep personal concern’

Some of the meanings of the examples above may appear to be not too far removed from the combined meanings of the members of the compound if and when their connotations are accessed or appealed to in each case. These examples do not have corresponding parallels as endocentric compounds.
V + N\textit{a}

(a) agaw + buhay
    snatch life
    ‘dying’ (Lit., ‘to snatch life’)

(b) sampáy + bakod
    hang fence
    ‘amateurish; inexperienced’ (Lit., ‘to hang (laundry) on a fence (to dry)’)

V + N\textit{v}

(a) maningala(?)=ng + pugad
    to.look.up=1KR nest
    ‘to begin courting a maiden’ (Lit., ‘to look upward for a nest’)

Considering (9) and (10), it remains to be a puzzle that Tagalog compounds consisting of \textit{V} + \textit{N}, quite productively, function more as \textit{Ns} (also observed in Mandarin Chinese, Anderson 1985:47–48) or even as \textit{As} rather than as \textit{Vs}, again contrary to the prediction of the Percolation Principle (Lieber 1983:252–253). This situation, especially frequent in exocentric compounds, is not considered a violation of the principle of headedness, but an exception. Williams (1981:250) simply charges that to headlessness. However, if we take the semantics of each constituent and their relation against the background of how they are reinterpreted by the speakers’ familiarity, belief or attitude towards specific features of one or both members and even the use of the expressions in specific situations, they can provide a clearer explanation for their formation and semantic interpretation.

3.1.3 Synthetic compounds

One source of a good number of compounds is from what is called \textit{synthetic} compounds. These are forms in which one of the constituents is a deverbal \textit{N}. Unlike in English synthetic compounds where the deverbal element is usually the second element, e.g. truck\textit{driver[N, hand-written[A, God-fearing[A, etc., the deverbal in Tagalog synthetic compounds may occur on either the left or right side of the compound. One of the common deverbalizing affixes in Tagalog is \textit{pang-} which derives \textit{V} bases into instrumental \textit{Ns}, thus, rendering the derived \textit{N} with the meaning ‘used for \textit{V}-ing’. 9 For example:

(12) \textit{pang-} \textit{V}\textit{N} + \textit{N}\textit{N}

(a) \textit{pang-}\textit{patay (=pamatay)} + lamok
    used.for.killing mosquito
    ‘something for killing/getting rid of mosquitoes’

(b) \textit{pang-}\textit{patid (=pamatid)} + uhaw
    used.for.cutting thirst
    ‘thirst quencher’

9 The velar nasal in the prefix \textit{pang-} undergoes assimilation according to the following consonant.
(c) **pang**-pasak (=**pam**asak) + butas
used for patching or covering hole
‘a person who, without prior notice, is used to substitute for another; a gap-filler’

It will be noted that the compounds in the above forms may either be endocentric, as in 12(a) and 12(b) or exocentric, 12(c), but all are of the category N. The right-hand N which functions as the complement of the left-hand deverbal N is also viewed as the complement of the base V. As mentioned previously, there is no question that the syntactic and semantic features required by each V, as specified in each lexical entry, play a role in both sentence and compound word formations.

Another productive affix that derives Ns from Vs is the locative suffix **-an** which results in the meaning ‘place for V-ing’. To illustrate:

(13) \[ N + V\text{-}an \mid N \]

(a) hapág + kain-án
table eating place
‘dining table’

(b) silíd + aral-án
room place for studying
‘study-room’

(c) bahay + áliw-án
house amusement place
‘a red-house’ (Lit., ‘house to amuse one’s self in’)

As in the preceding examples, the first two compounds in (13) are transparently endocentric while the third is exocentric. In this combination, the deverbal N modifies, directly or indirectly based on their meaning, the first N identifying ‘the kind of N it is’. Given the pattern exemplified in 13(a) and 13(b), many more compounds are formed by substituting V-an forms that can be related to the given first N constituent, e.g. sulat-án ‘writing place’, gawá-án ‘working place’, etc. for the former and tulug-án ‘sleeping place’, làrú-án ‘place for playing’, tanggap-án ‘receiving (visitors)-place’, etc. for the latter. In addition, other Ns may occupy the first position in the compound, for instance, 13(a) and 13(b) may be interchanged, e.g. silíd-kainán ‘dining room’ and hapág/mesa-ng aralan ‘study table; desk’ because the two elements are compatible in making clear their referents. As long as the relevant relationship existing in the combination is maintained, more compounds can be expected to be formed.

A third affix which derives deverbal Ns is **-in**. It is less productive than the first two affixes. The meaning attached to this class of derived Ns is ‘something to V’. With the suffix ending in a dental nasal, the N form requires the linker =ng before the following N, which relates as location or as goal (intended user) of what is expressed by the deverbal N.

---

10 The form hapág is an older form. The current form especially in the urban areas is the Spanish borrowing mesa, thus, mesa-ng kainán.
Since Tagalog synthetic compounds do not present a problem of structural analysis, unlike the contrasting accounts for similar compounds in English, they obviously do not need a separate subclassification.

### 3.1.4 Copulative compounds

When two related words of identical categories form a compound in which not one is head and their joint meanings comprise its composite meaning, the result is known as a **copulative compound**. Some forms of this type are allied to the exocentric type in terms of the unpredictability of their derived meaning. Compare 15(a), whose constituent meanings may be said to represent 'sweetness and deliciousness' which characterize the event referred to as 'honeymoon', with 15(b) which refers to the resulting mixture of the two kinds of liquid specified by the constituent elements of the compound. Similar to 15(b), the compound in 16(a) with the cliticized coordinate conjunction between the Ns expresses clearly a physical state, hence, an A.

15. \( N + N \)N

(a) pulót + gatá?
molasses + coconut.milk

'honeymoon'

(b) patís + mansí?
fish.sauce + kalamansi (a Filipino lemon fruit)

'a mixture of fish sauce and kalamansi juice'

16. \( N + N \)A

(a) butó=(a)t + balát
bone=and skin

'skin and bones'

In the following rare combination of A + A, 17(a) specifically refers to a shape that may be described as falling between the shapes represented by the individual constituents of the compound. Even this single illustration reminds us that given a possible categorial combination, there is no assurance of its productivity nor of its systematic semantic representation.
It will be observed in the succeeding groups of $V + V$ combinations that the interaction between the actions represented by the elements in the compound is more transparent. Both actions are involved individually such that the more common form and meaning is like that found in (20) as $V$ compounds, although the $N$ in (18) is necessarily characterized as being engaged in the actions depicted by the two member $V$s, similar to the descriptive meaning found in the $A$ compound in (19).

(18) $V + V|N$ (very rare)
(a) bantáy + salakay
to.guard to.attack
‘opportunist; one who does an inside-job’

(19) $V + V|A$
(a) urong + sulong
to.retreat to.advance
‘ambivalent; indecisive’

(20) $V + V|V$
(a) balík + aral
return study
‘to review past or previous lessons’
(b) akyát + pana’og
go.up go.down
‘to go up and down (the stairs)’
(c) lumubóg + lunitáw
to.sink to.appear
‘to bob up and down’
(d) paró(n)=(a)t + parito
to.go.there=and to.come.here
‘to go back and forth’
(e) labás + pasok
go.out come.in
‘to go in and out (of a room)’

3.2 Phrasal compounds
Unlike in English where phrasal compounds such as those cited in Lieber (1992: 11), e.g. ‘over the fence gossip’, ‘off the rack dress’, ‘a pleasant to read book’, etc., are
frequently occurring, Tagalog does not exhibit similar compound structures. At best, one form observed is of the type:

\[(21) \quad \text{Neg} + V + N|A\]

(a) \((\text{hin})\text{di} + \text{mahapaya}(n)\text{ng} + \text{gatang}\)

\begin{align*}
\text{not} & \quad \text{can.be.leveled/toppled} \\
\text{‘boastful; showy’ (Lit., ‘a cup that cannot be toppled’)}
\end{align*}

(b) \((\text{hin})\text{di} + \text{mahug}\text{a}(n)\text{ng} + \text{karayom} \)

\begin{align*}
\text{not} & \quad \text{can.be.dropped.in} \\
\text{‘crowded with people’ (Lit., ‘cannot drop a needle in’)}
\end{align*}

(c) \((\text{hin})\text{di} + \text{maliparål}(n)\text{ng} + \text{uwák} \)

\begin{align*}
\text{not} & \quad \text{can.be.floated.over} \\
\text{‘unusually large expanse’ (Lit., ‘cannot be flown over by a crow’)}
\end{align*}

The particle \text{hindi} or \text{di}?, for short, negates what is expressed by the following V, which is usually made up of the prefix \text{ma-} ‘abilitative’ and the verb stem with the voice affix \text{-an} to indicate the prominence of an object as in (a) or a location as in (b) and (c). Evidently, from the meanings of the above examples they belong to the exocentric type of compounds. With the category A, however, the intensive form will merely duplicate the negative particle in its full form rather than the whole compound. Thus,

\[(21') \quad \text{hindi=}\text{ng hindi} \quad \text{mahapaya=ng-gatang si Pedro mula nang} \]

\begin{align*}
\text{not(intensive)} & \quad \text{boastful} \\
tumama & \quad \text{sa loto} \\
\text{‘Pedro could not really stop being boastful ever since he won the lottery.’}
\end{align*}

Further study on phrasal compounds may reveal other interesting constraints.

4. The Role of Semantics and Pragmatics in Tagalog Compounding

It has been shown above that Tagalog compounds cannot be generally and adequately characterized in terms of absolute constraints on category features. If the syntactic rules for compounding as proposed by Selkirk (1982:47) for English is followed here, we will have contradictions in terms of headedness and productivity. (Productivity as used here refers basically to the number of instances observed and the potential of formulating analogously similar forms.) Based on the empirical data given previously, the following lexical rules are supposed to account for forming the productive and moderately productive endocentric compounds:

\[(22) \quad \begin{align*}
\text{N} & \rightarrow \text{N} + \left\{ \begin{array}{c} \text{N} \\ \text{A} \end{array} \right\} & \text{(productive)} \\
\text{A} & \rightarrow \text{A} + \text{N} & \text{(moderately productive)} \\
\text{V} & \rightarrow \text{V} + \text{N} & \text{(productive)}
\end{align*}\]

\[11 \quad \text{gatang ‘a unit of measure equal to about a small cup’}.\]
Note that in terms of headedness, Lieber’s (1983) Percolation Principle and later on, her Licensing Conditions (1992) work as far as this type of compounds are concerned, i.e. the left-hand head category is also the category of the whole compound and the non-head on the right is the modifier or complement of the head. In terms of productivity, the question remains to be why the patterns N + N|N and V + N|V are the only productive rules.

Compared with the forms of exocentric compounds below, most of the rules contradict the predictions contained above, both in terms of headedness and productivity:

(23) \[ \begin{align*}
N & \rightarrow \{N, A\} + N \\
A & \rightarrow N + \{N, A\} \\
V & \rightarrow N + N \\
& \text{(moderately productive)} \\
& \text{(moderately productive)} \\
& \text{(productive)} \\
& \text{(productive)} \\
& \text{(moderately productive)} \\
& \text{(moderately productive)}
\end{align*} \]

Why would V + N|N and N + N|A patterns in this type be more productive than the rest of the patterns?

For the so-called synthetic compounds, all the deverbal Ns combine with N resulting in Ns, thus, this set is to a large extent similar to those already accounted for both in rules (22) and (23) above. Finally, with respect to the copulative compounds, only the following are worth generalizing:

(24) \[ \begin{align*}
N & \rightarrow N + N \\
V & \rightarrow V + V
\end{align*} \]

The other copulative patterns A + A|A, V + V|N, and V + V|A are all rare.

Analyzing the productive patterns above, we can be impressed with the N + N and the V + N compounds, the former resulting in either N or A categories and the latter producing either V or N categories. Also of interest are the moderately productive patterns N + A and A + N that may both be either N or A, based on the word meaning of the resulting compound. Let us take the unlikely N + A|A and A + N|N patterns cited previously:

(25) \[ \begin{align*}
A + N|A/N \\
\text{patáy } + \text{ gutom} \\
\text{dead } + \text{ hunger}
\end{align*} \]

‘famished|A; vagabond|N’ (Lit., ‘dead/dying of hunger’)

This compound is ambiguous. As an A, the left-head A is complemented by the right-nonhead N. Its compositional meaning ‘dead (dying) of hunger’ corresponds to that of the syntactic phrase patáy ng/sa gutom. To show that this compound is an A, we can put it to a test for the category A. Intensification of meaning can be done through reduplication of the A with the particle na/=ng joining the two elements. Thus,

(a) \[ \begin{align*}
\text{patáy-gutom } & \text{ na patáy-gutom } \\
\text{extremely.famished } & \text{ the plural children when}
\end{align*} \]

\[ \begin{align*}
\text{dumatíng } & \text{ galing sa } \\
\text{arrived } & \text{ from the game}
\end{align*} \]

‘The children were extremely famished when they arrived from the game.’
Reduplicating only the first constituent and not the whole compound as follows results in an ungrammatical form, which confirms the cohesiveness of the two elements as a word:

(b) *patay na patay-gutom

Compared with the compound’s lexicalized meaning ‘a vagabond’ which is a N, the A meaning still characterizes the life style of the person referred to, i.e. a bum, and considering what he does NOT do, he must always be suffering from extreme hunger.\(^{12}\) It appears that lexicalization does not have to single out one of the number of possible relations that may hold between the constituents as Downing (1977:819) suggests, but it can merely rely completely on the compound’s meaning itself to form the major and essential semantic features of the lexicalized form.

(c) nahuli na ang patáy-gutom na nagnakaw sa amin at our.(place)

‘The vagabond who broke into our place has been caught.’

Consider the next pattern with the corresponding example in its intensive form below:

(26) \(N + A\)A

\[
\begin{align*}
isip & \quad + \quad matandá? \\
\text{mind/thought} & \quad \text{old/adult} \\
\text{‘mature’ (Lit., ‘old/adult mind’)}
\end{align*}
\]

(a) malīīt pa si Teddy ay isip-matandá=ng isip-matandá? na small yet Teddy very.mature already

‘While Teddy was still (small) young, he was already quite mature.’

It is conceivable that the compound in (26), exemplified in (26)a in the intensive or iterative form of an A, also means ‘a mature mind’ which then fits the category N. However, usage further confirms that as a N, as in the following example, it is unacceptable:

(b) *may isip-matandá? si Teddy have mature.mind Teddy

It appears that in this compound the A element is semantically more dominant, especially when it is contrasted with the opposite compound \(isip-bata?\) ‘immature’ where the second element does not refer to ‘child’ per se but to being ‘young’ as opposed to being old.\(^{13}\)

---

\(^{12}\) Owing to the remarkable interderivational possibilities existing between lexical categories in Tagalog, the N category reflected in this compound could have been derived by conversion from the A form.

\(^{13}\) The word \(bata?\) is ambiguous. Basically, it refers to ‘child’, a N. The other meaning is ‘young’. When compared to another ambiguous word, \(magulang\) ‘N[parent; A[mature; of age’, a potential form such as \(isip-magulang\) will more likely be interpreted as ‘having a mature mind’, analogous to \(isip-matandá?\), rather than ‘having a parent’s mind’. This is true especially in rural areas where \(magulang\) continues to be used independently with the meaning as above. In the same vein, it is reasonable to consider the adjectival meaning of \(bata?\) ‘young’.
Based on the preceding, the left-head principle fails on a number of sets of patterns. To account for those compounds that contradict this principle as exceptions can not be that convincing because of the observed frequency of such formations. It is suggested that the combinations and their resulting categories are more motivated by the types of semantic relations that may exist between two words, the situations that trigger their formation, and how the resulting compounds are used. Following a compound rule, as given above, two items put together at random may not form a compound due to a lack of semantic connection that may be established between them for a given situation. And this is true for compounds of any type discussed above. For example, given N + N, a productive combination, it is not easy to explain why sundalo=ng kanin ‘cowardly/weak soldier’ (Lit., ‘soldier that is (like) cooked rice’) is possible but not sundalo=ng kamote (Lit., ‘soldier that is (like) sweet potatoes’) or sundalo=ng damo (Lit., ‘soldier that is (like) grass’), taking ‘cooked rice’ and ‘cooked sweet potatoes’ are comparable food stuffs, or even considering ‘grass’ which is a tender kind of object. An explanation that may be ventured is that ‘cooked rice’, the staple food in the country, is certainly not enough by itself to maintain strong, robust soldiers. If a soldier turns out to be weak or cowardly, then it may be blamed on his non-nutritious diet of mostly ‘cooked rice’. It may also be that the physical or emotional constitution of the soldier is compared with ‘rice’, which may be viewed as either common, soft, bland, nothing much to boast about. Given the context of soldiers, expected to fight enemies in order to win, the complement may portray weakness/cowardice on the one hand or unusual strength/courage on the other. Thus, it is more likely to find an emergent form such as sundalo=ng kamagóng (Lit., ‘soldier that is (like) kamagong — a popular hardy type of tree’) to refer to ‘a strong/courageous soldier’ than say sundalo=ng niyóg (Lit., ‘soldier like coconut (a common native fruit with a hard shell)’).

Turning to the set of copulative compounds in which the pattern V + V\]V is the most frequent, the choice of the constituent Vs is again semantically significant in terms of their being largely opposite in direction, e.g. paník-pana?og ‘to go up and go down (the stairs)’, lumubóg-lumitáw ‘to bob in and out of the water’, etc. Even in the rare formation of the category A, as in urong-sulong ‘ambivalent/indecisive’ (Lit., ‘to retreat and to advance’) the Vs have to be contradictory in order to convey the desired resultant meaning. A potential form could be: tumayo-umupo (Lit., ‘to stand up and to sit down’) which depicts what one does when he/she is restless or worried over something. As a V, this can be put to test via reiteration which reduplicates the compound V with the particle nang between them. Yet, a form such as tumayo-kumain (Lit., ‘to stand up and to eat’) are two actions that do not seem to make up a related unit that makes sense to be useful, and thus, is marked unacceptable.

I contend that each compound is formed on the basis of some semantic and pragmatic relations that can tie or bind the two elements together as a composite unit, be it endocentric or exocentric in type. The former type, undeniably, is transparent in meaning and generally by analogy similar forms can be constructed. For example, given the compound form silíd-kainán ‘dining/eating room’ indicating the location for eating (in the house), the other room-labels expressing the activity done in each easily follow. Thus, silíd-tulugán ‘sleeping room/bedroom’, silíd-tanggapan ‘reception/receiving room’, etc. In like manner, sukang Iloko ‘Iloco vinegar’ readily falls within identical relations existing in: kapé-ng Batangas ‘Batangas coffee’, ripolyong Baguio ‘Baguio cabbage’, pansít Canton ‘Cantonese noodles’, to identify the items as being the best kind because they come from their respective places of origin. Any similar pair of Ns that does not convey this specific meaning relation does not come into existence. There is a different kind of noodles made from rice sticks which is usually boiled and has a special sauce that goes
with it. The general term that refers to it is *pansít luglóg* ‘rice noodles dipped and shaken (in boiling water)’. But compared with *pansit Malabón* which is a brand of *pansit luglog* prepared in the town of Malabon, anyone who is familiar with both terms will certainly prefer to savor the latter. Sometimes, we find meanings extended or generalized as in the case of *sabóng-Intsík* (Lit., ‘Chinese soap’) which is used specifically for laundry. Later, any kind of ‘laundry soap’ manufactured in the country is called by that name. In this manner, an identical form may begin to deviate from its original semantic content and develop its own unique meaning based on its pragmatic function.

Even with the synthetic type of compounds discussed in the previous section, the prominence of the semantic relations existing between the deverbal N constituent and its cooccurring N cannot be ignored and left to chance. Some forms have constituents that are especially tied together and do not allow any substitution or expansion. Given *pamatíd-uhaw* ‘thirst quencher’ (Lit., ‘used for cutting thirst’), we do not find *pamatíd-gutom* but rather *pantawíd-gutom* ‘something that appeases hunger’ (Lit., ‘used for bridging hunger’) (Schachter and Otanes 1972:109). In contrast, we can have *damít* ‘clothing; dress’ as the first N and append a deverbal N that means ‘used to V in or for Ving’, e.g. *damít-pangkasál* ‘wedding dress’ (Lit., ‘dress to wed in’), *damít-pantrabaho* ‘work clothes’ (Lit., ‘clothes to work in’), *damít-pansimbá* ‘Sunday clothes’ (Lit., ‘dress to worship in’), etc. Depending on the various occasions that one has to wear an appropriate attire, corresponding compounds can potentially be formed.

One interesting case of compounding which illustrates how speakers could use a variety of criteria for creating identical forms is exhibited in pet names of individuals. There is this guy known in our community as *Lilo-ng kabayo* (Lit., ‘Lilo who is a horse’). Because he is known to be in the business of selling horses, no one thinks of him as ‘being a horse’ or ‘looking like a horse’, which is a legitimate interpretation when this name is encountered for the first time. And what would you predict his son to be called? It is *Berto-ng bisiro* (Lit., ‘Lito who is a colt’), of course, not because he deals with ‘colts’ but simply being the offspring of someone whose appellation is ‘horse’. With the identical pattern in these two examples, notice that the interpretation depends largely on the particular circumstances surrounding the formation. This special type of compounding is very common particularly in the rural areas. It is popularly used to identify people readily. And the different relations people associate with individuals generally range from their occupation or expertise, physical feature, attitude, a disability, an unusual characteristic, and the like. For example, *Enyo-ng pulís* ‘Enyo, the policeman’, *Leoncia-ng panót* ‘Leoncia, the bald (one)’, *Regina-ng daldál* ‘Regina, the talkative (one)/gossip’, *Juan=(n)g tamád* ‘Lazy Juan’, etc. It may be added that there are also similar names in which the order of the head and its attribute is reversed such as *Mestra-ng Fidela* ‘Fidela, the teacher’, *Pilosopo-ng Tasyo* ‘Tasyo, the philosopher’, (ma)*Tandá-ng Sora* ‘Sora, the old one’.

With the explanations provided thus far to show the inadequacies of following solely a syntactic account, it is undeniable that both the semantic and pragmatic aspects of compounding must be pursued. Much like an underived word whose category and subcategory features are dictated to a large extent by its meaning, compound words as lexical items are also assigned these syntactic features based on their corresponding meaning. This meaning usually reflects some clear semantic features that carry over from both constituents to the whole compound, even when it is associated with a ‘third’ meaning as in certain exocentric compounds. Pragmatics has a lot to contribute to the semantic interpretation of a given compound and this is no better expressed than in knowing and believing which of the specified feature or features in an element plays a paramount role in a given combination. The resulting compound as a new lexical item
also carries certain features all its own in addition to those contributed by each component.

5. Conclusion

From the preceding exposition of the various aspects of compounding in Tagalog, it may be concluded that in Tagalog:

(a) Compounds are formed not because the categories are free nor constrained by their subcategorizational features. In fact, not any combination of two categories allowed by the “syntactic” rules as shown in section 3 is a potentially acceptable form. Each form of any type is constructed as an individual lexical item expressing a sensible relationship between two words, each constituent bearing its own significant identifying semantic features. Its formation may be triggered not only by the convenience of using just a word rather than a phrase but also by a desire to create a more expressive, colorful, metaphoric construction needed for a given situation;

(b) Many compounds, especially of the endocentric type, are formed extensively by analogy, and each one of these enters the lexicon, as do all other words, as soon as it gets constructed; as a lexical entry, a compound may be subject to other forms of lexical derivation;

(c) More compounds of the categories N and A, as indicated by their meanings, are formed most likely because of their naming and subclassifying, rather than asserting, functions;

(d) Even with a meaning which is viewed as unpredictable from the compositional meanings of its constituents, an exocentric compound is also not often completely free from the influence of some semantic and/or pragmatic features associated with one or both elements; and finally,

(e) If only the formation of endocentric compounds is taken into account because they conform to the principle of headedness, despite some counter-examples presented, we will be leaving out many, no doubt too many, of the other types of compound formation.

\[14\] The semantic relationships, more often observed in endocentric compounds, although not exclusively, are also found in syntactic phrases.
References


