Tagalog

Nikolaus P. Himmelmann

1.   Introduction

Tagalog is perhaps the best-known Philippine language. Its 17 million native speakers make it one of the two largest Philippine languages (along with Cebuano). In 1937 it was chosen as the basis for the national language, Filipino (formerly Pilipino), which was officially declared the national language of the Philippines in 1946 (see STEINHAUER, POLITICS, for a for further details). The differences between Filipino and Tagalog are mainly lexical, the Filipino lexicon being systematically expanded by the Institute of National Language. Because of its prestige as the national language and the fact that it is a medium of education and public discourse throughout the Philippines, Tagalog is also widely known as a second language (estimates vary between 70 and 90 per cent of the population).

Since the beginning of Spanish colonization in 1565, the Tagalog area has been the center of political power in the Philippines, and Tagalog has been the primary medium of communication between the colonial powers and the native population. Not surprisingly, then, Tagalog has been strongly influenced by the colonial languages, Spanish and, from 1898 to 1946, American English (other important and older sources for loan words in Tagalog are Malay (Wolff 1976) and Chinese (Chan Yap 1980)). This influence is most visible in its lexicon and phonology. However, there has also been some normative influence on its morphosyntax (clear examples are comparative constructions and the use of adverbials/conjunctions). The Spaniards began using Tagalog very early as a missionary language, preparing catechisms in Tagalog and writing grammars. A *Doctrina Christiana*, dated 1593, is the oldest preserved Tagalog document and shows that the standard language has changed very little in the last 400 years.

Most Tagalog native speakers live in the southern parts of Luzon, including the capital Manila. They are also found in many other parts of the Philippines, in particular on Mindoro, Palawan and in some parts of Mindanao (cf. McFarland 1983:29, 80). In addition, there are major Tagalog communities outside the Philippines, for example, in the United States and in Canada.

No in-depth study of Tagalog dialects exists. McFarland (1983:80) claims that Tagalog exhibits surprisingly little dialectal diversity. The following eight major regional dialects are mentioned in the Ethnologue (Grimes 1999): Bataan, Batangas, Bulacan, Lubang, Manila, Marinduque, Tanay-Paete, and Tayabas. For Tayabas, Arsenio (1971) provides an impressive dictionary. Furthermore, Lopez (1970) and

Soberano (1980) have drawn attention to some lexical and morphological idiosyncrasies in the Tagalog dialect spoken on the small island of Marinduque.

Of all Philippine languages, Tagalog is the one that has been studied most intensively and that has had the greatest influence in linguistics. There are a number of grammars in Spanish, the most influential one being the one written by Totanes (1865, first published in 1745). It served as the basis for many analyses of Tagalog until the turn of the century (e.g., Humboldt 1838, Müller 1882:87-163, Marre 1901). In the Spanish tradition, Tagalog grammars were cast in the framework and terminology of Latin grammars, the last and most comprehensive one being Blake (1925). Tagalog played an important role in the emergence of American structuralism because it was the first non-Indo-European language Leonard Bloomfield worked on. Bloomfield (1917) is among the earliest and finest pieces of structuralist grammatical analysis, attempting to present Tagalog in its own terms. Tagalog also plays a major role in Bloomfield's classic book *Language* (1933).

Beginning with Humboldt (1838), Tagalog has served as *the* prototypical example of a Philippine-type language in modern comparative grammar and typology. As pointed out in HIMMELMANN, TYPOLOGICAL CHARACTERISTICS, SECTION 6, this choice is somewhat unfortunate in that Tagalog differs in some important regards from its nearest relatives, the Central Philippine languages (not to mention the distinctly different Northern Philippine languages). The ongoing debate surrounding grammatical relations in Tagalog has been sparked by Schachter's (1976) paper on the distribution of subject properties.

A review of linguistic work on Tagalog is given in Constantino (1971:118-145) and Reid (1981). Major grammatical reference sources for Tagalog include Bloomfield (1917) and Schachter & Otanes (1972). Wolff et al. (1991) is the most comprehensive textbook to date, which also contains a wealth of grammatical observations and data. Major dictionaries include Panganiban (1972), English (1977, 1986), Santos (1983), Rubino (1998) and Constantino (1999). Bearing further witness to its prominent position as a national language and as a linguist's favorite, Tagalog is also the object of a substantial number of survey articles and encyclopedia entries (e.g. Schachter 1987, 1995, Zorc 1994, DeGuzman 2001).

Many issues in Tagalog grammar are a matter of considerable controversy. Among the prominent topics are stress, reduplication, voice marking and grammatical relations, and lexical categories (the distinction between nouns and verbs). For reasons of space, no attempt has been made here to provide for a detailed account of these controversies. However, ample references are given for all of these issues, which should make it easy to follow them up if required.

Most examples in this article come from the author's own corpus of spontaneous spoken narratives which includes stories from Wolff et al.'s (1991) textbook (see Himmelmann 1999:245f for details). Other sources are Bloomfield's (1917) text collection and the example clauses found in Father English's (1986) dictionary and

Schachter & Otanes' (1972) reference grammar. The examples from the narrative corpus retain features of the spoken language (in particular common reductions).

2. Phonology and orthography

2.1. Segment inventory, syllable structure and stress

The segment inventory of Tagalog does not include any unusual sounds. The vowels are /a,e,i,o,u/ and the consonants are listed in Table 1. Consonants in parentheses occur only in loans (and in a few colloquial forms such as [can] for *tiyán* 'stomach').

Table 1: Tagalog Consonant Phonemes

	Labial	Dental	Alveolar	Palatal	Velar	Labio-velar	Glottal
Stops voiceless	p	t		(c)	k		ʔ
voiced	b	d		(ɟ)	g		
Nasal	m	n			ŋ		
Fricative	(f)		s				h
Lateral			l				
Tap or trill			r				
Glide				j		w	

The most common syllable types are CV and CVC. That is, there are no vowel-initial syllables. In native words, syllable-internal consonant clusters only arise if unstressed vowels are deleted before glides as when the pronoun *siyá* is realized as [ʃja] or *tuwíd* 'straight' as [twid]. A small number of syllable-internal consonant clusters are allowed in loans, e.g. *trabaho* 'work' and *lipstik*.

Lexical bases are typically disyllabic and any combination of the two basic syllable types is allowed (i.e. CV.CV, CV.CVC, CVC.CV, and CVC.CVC). Clitics and affixes are often monosyllabic (either CV or CVC). Since many Tagalog words consist of a lexical base and at least one affix, words with three or four syllables frequently occur. Moreover, since most affixes can be combined with each other, words with even higher numbers of syllables are also common (e.g. *makipagtáwánan* 'join in laughter' < *tawa* 'laugh').

Stress in Tagalog appears to be closely interrelated with vowel length. For some authors (e.g. Bloomfield 1917, Matsuda French 1988) stress is the primary phenomenon while vowel length is an epiphenomenon. Others (e.g. Schachter & Otanes 1972, Wolff et al. 1991) consider vowel length to be primary. Only one instrumental study exists (Gonzales 1970) and many of the relevant facts are still unclear. Here some of the

important (and mostly uncontroversial) facts are presented as primarily stress-related, for reasons of expository convenience.

Most importantly, stress is distinctive in Tagalog. That is, segmentally identical word forms are distinguished by the fact that stress falls on a different syllable, either the final syllable (e.g. *bukás* 'open') or the penultimate syllable (e.g. *búkas* 'tomorrow'). However, stress cannot fall on just any syllable. CVC-syllables in non-final position are generally not stressed (exceptions are *mínsan* 'once' and *pínsan* 'cousin').

Stressed syllables are made prominent by pitch changes (in isolation it is usually a strong rise), increased duration and/or loudness (intensity). Various factors contribute to the choice among these stress marking devices, including the syntactic context and sociolinguistic variables. When stress falls on a non-final syllable, the vowel in this syllable is distinctly lengthened while a stressed final syllable may be short.

Affixes and reduplicated syllables may bear their own stress, resulting in multiply stressed words, e.g. *má-kíta* 'be seen', *púpuntá* 'will go', *máka-limút-an* 'to forget'. Consequently, not only lexical bases but affixed forms as well can be distinguished solely by a difference in stress. Compare *magkákásáma* 'do things together' with *magkakasáma* 'people who accompany each other'.

Affixation may involve stress shifts. For example, it is common that the addition of a suffix to a lexical base causes the stress to shift one syllable to the right. Hence, *táwag* 'call' + *-an* is *tawágan* 'to call someone', and *laróʔ* 'game' + *-an* is *laruán* 'toy'. Not infrequently, different affixes (more precisely: different derivations) are distinguished by their interaction with stress. Thus, there is another derivation from the lexical base *laróʔ* involving the suffix *-an* which results in the form *lárúan* 'playground'.

Within intonational phrases, lexical stress (i.e. the stress pattern that occurs on words in isolation) is modified in a number of ways as a result of its interaction with pitch changes related to intonation contours. For example, the phrase *kaharíang itó* 'this kingdom' may be realised as *kaharíang ito* or *kahariang itó*, in which one of the two inherently stressed syllables loses its stress (see also Bloomfield 1917:143f and Schachter & Otanes 1972:30-56).

2.2. Orthography

Tagalog orthography, which makes use of the Roman alphabet, matches pronunciation fairly well in most words. That is, most letters in Tagalog represent the same value they have in the phonetic alphabet (for many English loans, however, the conventional English spelling is retained, e.g. *cylinder* instead of *silinder*). Only one native sound, the velar nasal [ŋ], is represented by a combination of two letters, i.e. <ng>. The alveo-palatal stops in loans (/c/ and /ɟ/) are frequently rendered by <ts> and <dy>, respectively. And as in most western Austronesian orthographies, the palatal glide /j/ is represented by <y>.

A major deviation from the close match between spelling and pronunciation results from the fact that glottal stops and stress are not indicated in the standard orthography.

Consequently, words such as *abó* 'ashes' that appear to be vowel-initial in writing actually have an initial glottal stop, i.e. <abó> represents /ʔabó/. Similarly, two vowels which occur adjacent to each other in writing are separated by a glottal, thus *tao* 'person, people' is /táʔo/. And many orthographically vowel-final words such as *sira* 'break, damage' are actually /síraʔ/.

The lack of an indication for final glottal stops in writing may create confusion since two words may differ only with regard to this feature. Compare /báta/ 'robe' with /bátaʔ/ 'child', both conventionally spelled as *bata*. Similarly, stress distinguishes otherwise identical word forms, as illustrated above. In order to make this important information available, the better dictionaries and textbooks (as well as Bloomfield 1917) indicate both stress and final glottal stops. The convention for stress is to indicate it with an acute accent whenever it does not fall on the penultimate syllable, or when two adjacent syllables are stressed. Thus, conventional *abo* is written *abó* in the dictionaries and conventional *laruan* either *laruán* (for 'toy') or *lárúan* (for 'playground'). But *bata* 'robe', which has penultimate stress, is just written *bata* in the dictionaries.

The convention for final glottal stops is to indicate them by a grave accent on the final vowel (conventional *bata* 'child' is *batà* in the dictionaries). Words having both final stress and a final glottal stop are marked by a circumflex on the final vowel, e.g. *hindî* 'no'. In this article, all examples are presented according to the dictionary conventions, except that final glottals are indicated by <ʔ> (thus *hindî* is spelled *hindíʔ*).

Two further orthographic conventions are of practical importance: First, in the Tagalog alphabet the letter <k> takes the place conventionally occupied by <c>. Thus, the beginning of the alphabet reads *a b k d e* etc. Furthermore, the digraph <ng> representing the velar nasal is placed in between <n> and <o>, hence *l m n ng o p* etc. Second, it is conventional to abbreviate the following two frequent grammatical function words: the genitive marker /naŋ/ is written as *ng* and the plural marker /maŋá/ as *mga*. This convention is also adhered to in the present chapter.

2.3. Phonological alternations

There are very few phonologically conditioned alternations in Tagalog. That is, when affixes are added to lexical bases (or clitics to their hosts) neither affix nor base undergo many formal changes. Note in particular that initial glottal stops are not deleted when preceded by a consonant-final formative. For example, *mag+aral* 'to study' is [mag.ʔá.ral] and not *[ma.gá.ral]. However changes in the stress pattern occur fairly regularly, in particular when suffixes are involved (see 2.1 above).

With a few exceptions, prefixes which involve a homorganic nasal follow the general rule given in HIMMELMANN, TYPOLOGICAL CHARACTERISTICS, SECTION 2.1. Initial /s/ is regularly replaced by /n/ (e.g. *maN+sakít → manakít* 'to hurt'). Before glides, *N* regularly becomes /ng/ (e.g. *maN+wikaʔ → mangwikaʔ* 'say sth. insulting'). The deletion of the voiced stops /b/ and /d/ is not uncommon (cf. *maN+ bilí → mamilí* 'to go shopping' and *maN+daʔíg → manaʔíg* 'to prevail'). Somewhat exceptional is the fact

that initial glottal stops are often retained (*maN+apí* → *mang-apí* [maŋʔapí] 'to oppress'). Furthermore, the replacement of voiceless obstruents is optional in some formations (e.g. *paN+takíp* is either *pantakíp* or *panakíp* 'sth. used for covering'; cf. DeGuzman 1978b).

Another fairly general process is the deletion of the vowel in a stressed base-final CVC-syllable when a suffix is added. Thus, *bigáy+an* is *bigyán* 'give sth to someone' and *sunód+in* is *sundín* 'to obey'. Occasionally this deletion is accmpanied by metathesis: *taním+an* → *tamnan* 'to plant on'. Final glottal stops are generally deleted in this context: *gawáʔ+in* → *gawín* 'to make sth'.

When suffixes are added to vowel-final bases, an /h/ is inserted between base and suffix as in *basa+in* → *basahin* 'to read sth' or *sabi+in* → *sabihin* 'say sth'. H-insertion together with base-final vowel deletion account for derivations such as the following: *bilí+in* → *bilhín* 'to buy sth', *sará+an* → *sarhán* 'to shut sth'. Sporadically, there is also the insertion of an /n/ when bases end in a vowel or a glottal stop (e.g. *tawa+an* → *tawanan* 'to laugh at someone'). Completely irregular is the form *kunin* 'to get sth' from *kuha+in* (which occurs along with *kuhanin* (same meaning)).

A final glottal stop is regularly deleted, and the preceding vowel lengthened, when an enclitic follows. Thus, *hindíʔ=na* 'no longer' is pronounced [hindí:na], *gawáʔ=nya* 'his/her deed' [gawá:ɲa]. Note that no such change occurs with suffixes.

A few other automatic changes occur on the level of phonological words and phrases. In native words, a final /o/ is regularly raised to /u/ when not occurring in phrase-final position. Thus, the proximal demonstrative *itó* is [ʔitó] when phrase-final (as in *bundók na itó* 'this mountain') but it is [ʔitú] in other positions (*itúng bundók*). The base for 'game' is *laróʔ*, but the suffixed form for 'toy' is *laruán*. Similarly, an /i/ occurring in a phrase-final syllable may be lowered to /e/ as in the common pronunciation [gabé] for *gabí* 'night' ([gabí] is also possible). Furthermore, /d/ often becomes /r/ in intervocalic position, although this is by no means always the case. Compare *doón* 'there' with *maroón* 'be there', *tawad* 'discount' with *tawaran* 'ask for a discount on sth', etc. Note that /r/ does not occur in initial position or after a consonant in Tagalog native words.

2.4.   Reduplication

From a purely formal point of view, there are three kinds of reduplication processes in Tagalog. In two of these processes the first consonant and vowel of the base are copied (CV-reduplication). The two processes differ with regard to the fact that in one the reduplicated syllable is assigned stress (RDP1), and in the other it is not (RDP): *mang-gá-gamót* 'will practice medicine' vs. *máng-ga-gamot* (or *mánggagamót*) 'physician'. If a base starts with a consonant cluster, only the first consonant is copied, cf. *trabaho* 'work' → *mag-tá-trabaho* 'will work'. Recall from sections 2.1 and 2.2 that all Tagalog syllables are consonant-initial even though initial glottal stops are not indicated in the

orthography (hence a form such as *mag-íígi* /magʔíʔígi/ 'to adjust' (base *igi* /ʔígi/) is not an example for the reduplication of a vowel-initial base).

The third process consists in copying the first two syllables of the base (RDP2). In the case of disyllabic bases, the complete base is copied including its stress pattern, e.g., *mura* 'cheap' → *mura-mura* 'rather cheap', *lakad* 'walk' → *mag-lakád-lakád* 'to stroll, to walkabout'. If bases contain more than two syllables, the second syllable is copied only up to its peak, which is then also assigned stress, cf. *tahimik* 'peaceful' → *tahí-tahimik* 'rather peaceful', *baluktót* 'crooked' → *balú-baluktót* 'variously bent'. RDP and RDP2 generally apply to lexical bases only, while RDP1 freely applies to prefixes as well (*mag+RDP1+pa+putol* → magpápaputol 'will cause to be cut').

A certain kind of reduplication may occur only once in a derivation, but different kinds of reduplication may co-occur within the same derivation. When RDP1 and RDP2 co-occur, RDP2 derivationally precedes RDP1: *magsásakít-sakitan* 'will pretend to be sick'. RDP1 and RDP may also co-occur, in which case RDP derivationally precedes RDP1. Thus from *takbó* 'run' *mág-ta-takbó* 'run wild' is derived by prefixing *mag-* and RDP. From the latter the imperfective aspect *mag-tá-ta-takbó* is derived by RDP1.

In addition to reduplication, there is also the possibility to repeat complete (grammatical) words in order to express various forms of intensification. In this case, the word forms remain unchanged and there is usually a linker (see below 3.2) inserted in between the two repeated forms. Examples: *pagód na pagód* 'very tired', *magandá-ng-magandá* 'very beautiful', etc.

3.   Basic Morphosyntax

3.1.   Basic clause structure and phrase marking

Tagalog clauses are basically of two types. In simple main clauses, the predicate occurs in clause-initial position and the subject comes at the end. If the subject is a common noun (such as *aswáng* in (1)) it is usually preceded by the specific article *ang*:

```
(1)   dumatíng  na   ang   aswáng
      AV:arrive now  SPEC  vampire
            The vampire came.
```

Typically, there is also at least one clitic particle in clauses of this type (in (1), *na* is such a particle). These clitics are second position clitics, i.e. roughly speaking, they occur after the first word of the predicate. If the predicate consists of only one word (such as *dumatíng* in (1)), this is equivalent to saying that the second position clitics come immediately after the predicate (see also section 3.3).

The second clause type is the existential clause which consists of the existential operator *may* (or *mayroon*) or its negative equivalent *walá* and a complement (*bigás* in (2)). Other, optional constituents in this clause type are clitics (*na* and *raw* in (2)) and local adjuncts (*sa tindahan* in (2)):

(2)   may   bigás na raw sa tindahan
       EXIST rice now RPRT LOC store
              There's rice in the store now, they say.

Existential clauses may be expanded by a subject phrase (typically in final position) and then convey a possessive meaning (cf. McFarland 1978a):

(3)   may   kápitbahay naman **silá**
       EXIST neighbor   also   3p
              They had (or have) a neighbor, …

In colloquial Tagalog, clauses often end in an emphatic particle (*e*, *a*, *ha*, or *o*) which may express regret, sympathy, surprise, urgency, sarcasm, reason or the like:

(4)   napakagandáng   dalaga   iyón **e**
       ELATIVE:beauty:LK young.woman DIST EMPH
              She (that one) is a very beautiful girl, you see.

The preceding example also illustrates the fact that all kinds of expressions may serve as the predicate in simple main clauses. In (4) the predicate consists of the complex expression *napakagandang dalaga* 'very beautiful girl'. The typical predicative expression in simple main clauses, however, is a voice-marked word which is roughly equivalent to an English verb (see 3.3 below for some pertinent differences between English verbs and Tagalog voice-marked words). Voice-marking will be discussed in detail in section 4.1. Here it is sufficient to note that there are essentially two basic types of voice constructions. In actor-oriented voices, the predicate is marked for actor voice (by the infix *-um-* in (5)) and the subject expression refers to the actor of the event denoted by the predicate:

(5)   h**um**anap   na ng bahay **ang mga bataʔ**
       um-hanap   na ng bahay ang mga bataʔ
       AV-search now GEN house SPEC PL   child
              The children looked for houses/a house.

In undergoer-oriented voices, the predicate is marked for an undergoer voice (by the infix *-in-* in (6)) and the subject expression refers to the undergoer of the event denoted by the predicate:

(6)   h**in**anap   na ng bataʔ **ang bahay**
       in   -hanap na ng bataʔ ang bahay
       RLS(UG)-search now GEN child SPEC house
              The children looked for the house(s).

That is, a special relation exists between the subject and the predicate in that the voice affix on the predicate marks the semantic role played by the subject in the event expressed by the predicate. Thus, for example, the infix *-um-* in (5) makes it clear that the subject (*ang mga bataʔ*) is the agent of the search and not its undergoer. This special relation between voice-marked predicates and subjects is one of the reasons for

assigning the grammatical relation *subject* to the *ang*-phrase in clauses such as (5) and (6) (see HIMMELMANN, TYPOLOGICAL CHARACTERISTICS, SECTION 3.8.1, for references and discussion regarding the controversy surrounding subjecthood in Tagalog).

Note that the specific article *ang* is not a subject marker since *ang*-phrases may also occur in other syntactic functions. In the following example, the first *ang*-phrase (*ang langgám*) functions as the predicate while the second *ang*-phrase (*ang tumulong sa mga bataʔ*) is the subject (for a different analysis of this construction, see Constantino 1965, DeGuzman 1986, and Kroeger 1993:148-152):

(7) **ang langgám** rin ang tumulong sa mga bataʔ
    SPEC ant   also SPEC AV:help LOC PL child
           The ants also helped the children (lit. The ones who helped the children were also the ants).

Furthermore, subjects (and some, but not all, non-predicative constituents) may occur in pre-predicate position. In the most common type of construction in which the predicate is not initial, the predicate marker *ay* ('*y* after vowels) signals the beginning of the predicate. Thus, instead of *dumatíng ang aswang*, the following is also possible (same meaning as in (1)):

(8) **ang aswáng ay** dumatíng

Although a subject is always implied in a Tagalog simple main clause (except in meteorological expressions and the like), this subject does not have to be overtly expressed. Compare the following sequence of two clauses where the subject of the second predicate (*inilagáy*), which is coreferential with the subject of the first predicate (*kinuha*), is not overtly expressed:

(9) at kinuha   niyá   ang langgám
    at in   -kuha   niyá   ang langgám
    and RLS(UG)-getting 3s.POSS SPEC ant

    at **inilagáy**   niyá   sa pampáng
    at in   -i -lagáy   niyá   sa pampáng
    and RLS(UG)-CV-position 3s.POSS LOC river.bank
           And he got the ant and put (it) on the riverbank.

The preceding examples also show that non-subject arguments and adjuncts are marked with either the genitive preposition *ng* (i.e. [naŋ]) or the general locative preposition *sa*. The choice among these two markers for a given argument or adjunct depends on its semantic role and, in one instance, on its animacy (human vs. non-human). Non-subject actors (including experiencers) are generally marked as genitives (cf. *ng mga bataʔ* (6)). Non-subject patients and themes are also marked as genitives, if non-human (cf. *ng bahay* in (5)). If they are human and specific, they are usually marked by *sa* (cf. *sa mga bataʔ* in (7)). All other semantic roles - except non-subject instruments, which may also be marked as genitives - are expressed by *sa*-phrases or complex prepositional expressions, which in Tagalog always include the general locative marker *sa* (e.g.,

tungkól sa 'about, regarding', *para sa* 'for'). This includes a large variety of temporal and local adjuncts as in (10) as well as recipients/goals (*sa pampang* in (9) above):

```
(10)  at   dun       na   silá tátabunan              sa   lugar na iyón
      at   doón      na   silá RDP1-tabon        -an  sa   lugár na iyón
      and  DIST.LOC  now  3p   RDP1-complete.cover-LV LOC  place LK DIST
            And there they were covered (with earth) at that place.
```

Tagalog argument and adjunct expressions, then, are generally preceded by a grammatical function marker (i.e. *ang*, *ng*, and *sa* in the case of common noun phrases). Their order is not totally fixed but there is a strong tendency for them to occur in the following order: genitives almost always occur immediately after the predicate, subjects tend to be placed in clause-final position, and locative-marked phrases typically precede the subject when expressing an argument. They may either precede or follow the subject when expressing an adjunct.

Not all nominal expressions in Tagalog are marked by one of the three markers *ang*, *ng*, and *sa*. These markers only occur in common noun phrases. For personal names there is a different series of markers, and demonstratives and personal pronouns have their own *ang*, *ng*, and *sa* forms which are not preceded by a marker. Table 2 lists some of the forms.

Table 2: Noun phrase markers and pronouns

	ANG-FORM (SPEC)	NG-FORM (POSS/GEN)	SA-FORM (LOC/DAT)
Noun phrase markers			
COMMON NOUNS	ang	ng	sa
PERSONAL NAMES	si	ni	kay
Pronouns			
1.SG	akó	ko	akin
2.SG	ikáw/ka	mo	iyo/iyó
3.SG	siyá	niyá	kaniyá
1.DU.IN	kitá/katá	nitá	kanitá
1.PL.IN	tayo	natin	atin
1.PL.EX	kamí	namin	amin
2.PL	kayó	ninyó	inyó
3.PL	silá	nilá	kanilá
Demonstratives			
PRX	itó	nitó	dito, rito
MED	iyán	niyán	diyán, riyán
DIST	iyón	niyón, noón	doón, roón

Overall, the syntactic distribution of the forms in each column is very similar. In fact, the distribution of the *ang*, *ng*, and *sa* forms of the demonstratives is completely identical to the distribution of the *ang*, *ng*, and *sa* forms of common noun phrases. In particular in the spoken language, *ang, ng,* and *sa* are often replaced by the corresponding forms of the demonstratives. Compare the following example in which *yung*, a variant of the distal demonstrative *iyón* plus linker, has taken the place of *ang*:

(11) s**um**igáw **yung   anák**
      AV:shout DIST.LK child
            That child shouted: …

With regard to the *ng* and *sa* forms of personal pronouns and personal names, however, there are some minor, but still important distributional differences which warrant slightly different glosses. The *ng*-form of personal pronouns and personal names is used only to express possessors (as in *ang bahay niyá* 'his/her house') and actors in

undergoer-oriented constructions (as in (9) above). It cannot be used to mark the undergoer in actor-oriented constructions (i.e. *ng bahay* in (5) cannot be replaced by the *ng*-form of a personal pronoun or a personal name). To indicate this difference, the *ng*-form of personal pronouns and personal names is glossed here as *possessive* (rather than as *genitive*). Similarly, the *sa*-form of personal pronouns and personal names is glossed as *dative* (rather than as *locative*) because of a number of distributional differences (cf. Schachter & Otanes 1972:91,136).

3.2. Noun phrase structure and linkers

The structure of noun phrases is very straightforward. As just noted, all noun phrases are introduced by a proclitic phrase marker. Possessors follow their head and are genitive-marked:

(12) ang   hari **ng   lamók**
      SPEC king GEN mosquito
            the king of the mosquitos

An exception are pronominal possessors which can also precede their head and then occur in dative rather than in genitive form (i.e. a semantically fully equivalent alternative to *ang bahay niyá* 'his/her house' is *ang kaniyá-ng bahay* (SPEC 3s.DAT-LK house)).

All non-possessive modifying constructions involve the linker *na* (-*ng* after vowels, /n/ or glottal stop) between each major constituent:

(13) ang   ma-liít   **na** hayop
      SPEC ST-smallness LK animal
            the small animal

The order of the constituents in the linker construction is not fixed ('the small animal' could also be rendered by *ang hayop na maliít*). Only numerals generally precede all other constituents, except demonstratives, as in:

(14) ang **apat na** ma-lalim **na** balón
      SPEC four LK ST-depth LK well
            four deep wells

Attributively used demonstratives occur at the very periphery of noun phrases and may in fact be used twice, both at the beginning and the end:

(15) **itó-ng** isá-ng tasa-**ng itó**
      PRX-LK one-LK cup -LK PRX
            this one cup

Relative clauses are also constructed with a linker and may precede or follow their head. In the following example, there are two relative clauses, one preceding, the other following the head (*isáng ibon*):

(16) **kanyáng   nákítang**   isáng ibon **na mayroóng   pugad**
     kanyá -ng na-kita-ng   isá-ng ibon na mayroón:LK pugad
     3s.DAT-LK RLS.POT-seen-LK   one-LK bird LK EXIST:LK   nest
       (There was) a bird he happened to see that had a nest.

The head remains formally unexpressed in the relative clause. With a few exceptions, it has to correspond to the missing subject argument of the relative clause predicate (cf. HIMMELMANN, TYPOLOGICAL CHARACTERISTICS, SECTION 3.10).

    The linker does not only occur in noun phrases but generally links the major constituents of modifying constructions. Thus, it also links adverbial expressions, as in:

(17) **lalo-ng**   malakí ang takot
     surpassing-LK big   SPEC fear
       The fear was very great …

(18) **biglá-ng** dumatíng   yung   utusan
     sudden-LK AV-arrival DIST:LK servant
       Suddenly that servant came …

3.3. Lexical categories (word classes)

With regard to the grammatical categorization of lexical items, a first and very basic distinction can be made between content words and function words (or full words and particles). Content words have a meaning of their own and fall into large classes of items with similar grammatical properties. Function words signal often elusive grammatical or interactional meanings and form small, closed classes with regard to their grammatical properties.

    Most Tagalog function words are clitics (i.e. phonologically they are not independent words). The noun phrase markers (*ang, ng, sa*, etc.) are proclitics, i.e. they form a phonological constituent with the following content word. Apart from the proclitics, there is also a relatively large number of second position clitics. Second position clitics occur after the first content word of the constituent to which they belong (exceptions are discussed in Schachter & Otanes 1972:187-193, 433-435 and Kroeger 1993:118-123, 152-154). That is, clause-level clitics come right after the predicate if the predicate is in clause-initial position (as in examples (5) and (6) above). If another constituent, for example, an adverbial or a negation, precedes the predicate proper, the clitic will occur right after that constituent. In the following example, the politeness clitic *hóʔ* follows the (non-clitic) negation marker *hindíʔ* (note that clitics in Tagalog are orthographically represented as independent words in standard orthography):

(19) nakú hindí **hóʔ** pwede
     my! NEG   FRM possible
       Hey! That's impossible, Sir.

However, there are some clause initial constituents which are ignored for purposes of clitic placement. These include interjections such as *nakú* in the preceding example and *ay*-inverted constituents (cf. (8) above).

If the prepredicate constituent is linked to the predicate proper by a linker, the clitic (*siyá* in the following example) occurs before the linker:

(20) biglá **siyá-**ng nagbangon
     sudden 3s-LK   RLS.AV:rising
            … she got quickly up (from her bed), …

The domain for the placement of a clitic can also be smaller than a clause. In the following examples, it is a noun phrase. Note that the clitic (here *ko*) always comes after the first content word, regardless of whether or not this is also the semantic head of the construction:

(21) ang   anak   **ko-**ng   dalaga   'my daughter'
     SPEC child 1s.POSS-LK young.woman
(22) sa   dalawá **ko-**ng   kasama   'to my two companions'
     LOC two   1s.POSS-LK companion

There are two basic types of second position clitics: pronouns and a fairly heterogeneous set of particles, including the aspectual particles *na* 'already, now' and *pa* 'still', the question marker *ba*, the politeness markers *hóʔ* and *póʔ*, *yataʔ* expressing uncertainty, *palá* expressing surprise, etc. With regard to pronouns, it should be noted that not all forms of the pronouns are second position clitics. The *ng*-forms of the pronouns (such as *ko* in the preceding examples, see also Table 2 above) are always second position clitics. The *ang*-forms of the pronouns can be used as second position clitics (cf. *siyá* in (20)) but they can also be used as phonologically independent words. They may, for example, occur in clause-initial position preceding *ay*. Thus, an alternative order for (20) is *siyá 'y bigláng nagbangon*. No alternative orders exist for 'true' second position clitics such as *ho* in (19) or *ko* in (21) or (22).

When two or more clitics co-occur with each other, the general rule is that monosyllabic pronominal clitics come first, followed by the clitic particles (among which monosyllabic ones precede disyllabic ones), and disyllabic pronominal clitics come last, as in *bigyán **mo namán akó*** (gift-LV 2s.POSS truly 1s) 'give me (some) please' (cf. also Schachter 1973).

Content words fall into a number of subcategories. These subcategories do not correspond very closely to the categories noun, verb and adjective familiar from English and other Indo-European languages. One major distinction between the Tagalog categories and the English ones pertains to the fact that in the English scheme morpho-lexical and syntactic categories closely correspond to each other. That is, an English verb such as *betray* is a verb both on morphological grounds (for example, it forms a past tense with *-ed*) and with regard to its syntactic distribution (it may be used as a predicate but not, without further derivation, as the head of a noun phrase (**the betray*)).

In Tagalog, there is no such close correspondence between morpho-lexical and syntactic categories. In fact, content words do not have to be subclassified with regard to syntactic categories. They all have the same syntactic distribution, i.e. they all may

occur as predicates, as (semantic) heads of noun phrases and as modifiers (cf. Lemaréchal 1982, 1989, Himmelmann 1991, to appear, Gil 1993, Shkarban 1995, or Naylor 1995). For example, voice-marked words, which generally have a verb-like meaning, are not restricted to the predicate function. They may occur in other syntactic functions as well. In the following example, the undergoer voice form *sásabihin* 'will be said' functions as the patient argument of the actor voice predicate *naghintáy* 'waited for' and is marked accordingly by the genitive marker *ng*:

```
(23)  at   ang  pare    at   siyá ay nag-hintáy
      and  SPEC priest  and  3s   PM RLS.AV-wait

      ng  sá-sabih-in          ng   sundalo
      GEN RDP1-statement-PV GEN soldier
```
> And the priest and he waited for what the soldier would say.

Furthermore, voice-marked words may be in construction with quantifiers, including the existential quantifier *may*:

```
(24)  may    i-pá-pakita        akó sa   iyo
      EXIST CV-RDP1-CAU-seen 1s  LOC 2s.DAT
```
> I have something to show you.

The fact that content words have essentially the same syntactic distribution means that phrases marked by *ang, ng*, and *sa* are not noun phrases in the strict sense (i.e. phrases headed by a noun). Instead, they are *determiner phrases* which are headed by a function word. The function word determines the syntactic distribution as well as the basic meaning type of the phrase (they are all referential expressions). Thus, the term *noun phrase* in this chapter is to be understood in its wide sense as a cover term for both determiner phrases and noun phrases in the strict sense.

   As opposed to the syntactic level, Tagalog content words clearly fall into different classes on the morpho-lexical level. That is, they do not all have the same morphological possibilities and they differ with regard to the meaning alternations associated with various formal derivations. Here it is useful to make a distinction between affixed forms and non-affixed forms. Some affixed forms belong to more or less extensive paradigms of formations so that given one form it is possible to predict the availability of a number of other formations involving the same lexical item. This holds true in particular for voice-marked forms, all of which allow for aspect/mood inflection and a number of other regular derivations (see section 4 below). Because of these morphological characteristics, voice-marked words may be considered a morpho-lexical class of their own which, on the basis of their meaning, could also be called 'verbs', keeping in mind that Tagalog 'verbs' are very different from English verbs.

   Non-affixed forms, here called *lexical bases*, belong to different classes according to the kinds of affixes with which they occur. The actual number and kind of classes is a matter of ongoing research and controversy (for different ways of classifying lexical bases with regard to voice marking, see Blake (1925:38f), Schachter & Otanes

(1972:295-310), Cruz (1975), McFarland (1976), Ramos (1974), DeGuzman (1978a), Ramos & Bautista (1986)). Once again, however, it appears to be the case that one cannot distinguish formally between nominal and verbal bases, as is shown by the following lexicographic practice, among other things.

In standard dictionaries, the basic meaning of most lexical bases is rendered by an English noun, adjective or past participle. This may appear to be odd in the case of bases which denote actions or states, such as *lagáy* 'condition, state, location, position', *kuha* 'a helping, act of getting', *hanap* 'quest, object of search', *alís* 'departure', *kita* 'seen, obvious' or *hirám* 'borrowed'. When these bases are affixed with a voice affix, the meaning is clearly verbal (thus, for example, *umalís* is glossed 'to go/go away (actor-oriented)' and *hiramán* as 'to borrow from someone (undergoer-oriented)'). Most of these bases are used with such affixes most of the time. However, if the unaffixed base is used (which is possible for a great many of them), the meaning is clearly nominal (or adjectival). For example:

(25) subali't tuluy-tulóy   pa   rin   ang   **kain**   ni   Matsíng
     but   RDP2-continue still also SPEC eating PN.POSS small.monkey
          But the monkey's eating continued nevertheless.

(26) biglaan   ang   kanyáng   **alís**
     bigláʔ-an ang   kanyá -ng alís
     sudden-?? SPEC 3s.DAT-LK departure
          His departure was sudden.

Because of examples such as these, the lexicographic practice of glossing lexical bases as either nouns or adjectives/participles has some validity. Note that all the glosses for bases in this chapter are taken from a single source (English 1986), for reasons of consistency.

4.   Major verbal alternations

Like many other western Austronesian languages, Tagalog verbs exhibit an elaborate set of morphological distinctions. Recall from the preceding section that overt voice-marking is the defining feature of verbs. So Tagalog verbs are, by definition, voice-marked. Two further categorial distinctions are - explicitly or implicitly - expressed by all verbs. One of these is aspect/mood (cf. section 4.1). The other one is the distinction between dynamic and stative verbs. Dynamic verbs generally refer to the volitional and controlled doing of an action. Stative verbs refer to states of affairs which do not involve any kind of agent. For dynamic verbs there is a special morphology, called *potentive* here, which signals that the volitional and controlling agent typically involved in the state of affairs expressed by a dynamic verb is not in full control of the event. Stative and potentive morphology are similar both semantically and formally, and therefore are dealt with together in section 4.2.

In addition to the three major verbal categories of voice, aspect/mood and

dynamic/stative, a number of other concepts exists which are morphologically marked on the verb. Most of these concepts are in one way or another related to the way in which the actor is involved in a given state of affairs. They include causation, joint and reciprocal action, making a request and various forms of multiple actions. See Lopez (1937) for an extensive illustration of these affixes and Rubino (this volume) for a similar range of derivations from Iloko, often involving cognate affixes.

4.1. Voice and aspect/mood

The Tagalog voice system consists of at least four voices. Actor voice is marked by the infix *-um-* or the prefixes *mag-* or *maN-*, patient voice by the suffix *-in*, locative voice by the suffix *-an*, and conveyance voice by the prefix *i-*. The latter three voices share several semantic and formal characteristics, which makes it convenient to use *undergoer voices* as a cover term for them.

Most lexical bases may be combined directly with one or more of the voice affixes. Some bases, however, require the use of the stem forming-prefixes *pag-* or *paN-* before voice marking can be applied.

For each voice form it is possible to distinguish two aspects and two moods. Imperfective aspect is marked by accented CV-reduplication (RDP1), perfective aspect is unmarked. Realis mood is marked by the infix *-in-* in the undergoer voices, non-realis mood remaining formally unmarked. In actor voice, the expression of mood is somewhat less straightforward. When actor voice is expressed by *mag-* or *maN-*, there is a special *n*-initial form of the prefix, i.e. *nag-* and *naN-*, which signals realis mood while the *m*-initial form signals non-realis mood (the initial nasals in these forms, *n* and *m*, are historically related to the infixes *-in-* and *-um-*, respectively; cf. Reid 1992). When actor voice is marked by *-um-*, a mood distinction is formally expressed only in the imperfective aspect, where *-um-* is lacking. Table 3 illustrates these formations based on the base *bilí* 'purchase, sale'.

Table 3: Aspect/mood paradigm for *bilí* 'purchase, sale'

	AV ('buy')	AV ('sell')	PV	LV	CV
NON-REALIS/ PERFECTIVE	b<um>ilí	mag-bilí	bilh-ín	bilh-án	i-bilí
NON-REALIS/ IMPERFECTIVE	bíbilí	mag-bíbilí	bíbilh-ín	bíbilh-án	i-bíbilí
REALIS/ PERFECTIVE	b<um>ilí	nag-bilí	b<in>ilí	b<in>ilh-án	i-b<in>ilí
REALIS/ IMPERFECTIVE	b<um>íbilí	nag-bíbilí	b<in>íbilí	b<in>íbilh-án	i-b<in>íbilí

In addition to the lack of *-um-* in the non-realis/imperfective form of the actor voice, there is one further idiosyncrasy in this paradigm: The realis forms of the patient voice lack the suffix *-in* (for further comments on the asymmetries, cf. Himmelmann 1987:171-178).

The *m/n*-alternation signaling mood in the actor voice is also found with a number of *p*-initial bases (e.g., the non-realis form of *panoód* 'watch' is *manoód,* realis *nanoód*). For this and other reasons, the actor voice prefixes *mag-* and *maN-* can be analyzed as consisting of the stem deriving-prefixes *pag-* and *paN-*, respectively, plus *m/n*-alternation signaling voice and mood (see DeGuzman 1978a:149ff). The stem-deriving prefixes *pag-* and *paN-* also occur in the undergoer voices, i.e. there are formations involving *pag--in*, *pag--an*, *ipag-*, etc.

Table 3 is somewhat misleading in that it fails to bring out a very important difference between aspect/mood and voice. Aspect/mood-marking is completely general and regular. That is, given a voice-marked word it is clear that it will occur in all four aspect/mood-forms shown in Table 3. In this sense, it is legitimate to speak of aspect/mood-inflection of verbs in Tagalog.

Voice marking is different. It is common to illustrate the workings of Tagalog voice morphology with multiple derivations from the same base, as it is done here in Table 3. This procedure easily leads to the misconception that it is possible to derive a large number of different voices from every base. This is not the case. Instead, even in those instances where several derivations are possible from the same base, there are typically only one or two derivations in common use while the other ones are marginal and marked in that they convey special semantic and/or pragmatic meanings and implications (cf. McFarland 1976). That is, the voice affixes are essentially derivational when viewed from a crosslinguistic perspective (cf. DeGuzman (1997) and Rubino (1998a) for recent contributions to the inflection vs. derivation-controversy).

Before looking more closely at the different voice formations, a few comments on the four aspect/moods are in order. There are a number of alternative terminologies in use for these forms (cf. Werlen in Bader et al. (1994:95-100) and Kroeger (1993:15-18) for an overview). The terminology used here reflects the formal make-up of the paradigm but it is not quite felicitous in all instances with regard to the function of the forms.

The form labeled non-realis/perfective here is formally the most unmarked form (and therefore also called the *basic form* or *infinitive*). It has essentially the distribution of a subjunctive. That is, it is used in imperatives (27) and in various dependent constructions (reduced complement and adverbial clauses, in construction with modal auxiliaries (28), etc.):

```
(27) hindíʔ! tingnán    mo       ikaʔ si Maria!
     hindíʔ  tingín-an mo       ikaʔ si Maria
     NEG     look  -LV 2s.POSS said PN Maria
```
           Don't (panic)! Just look at Maria (she said)!

```
(28) gusto  ni       Pepito na   sagip-in   ang  dahon
     liking PN.POSS Pepito COMP salvage-PV SPEC leaf
```
          Pepito wanted to catch the leaf.

The non-realis/imperfective form (also called *contemplated aspect* or *future*) generally refers to states of affairs which have not yet taken place at reference time:

```
(29) gágawá    kamí ng  kubó
     RDP1-made 1pe  GEN hut
```
           We will make a hut.

The realis/imperfective forms (*imperfective aspect* or *present tense* in alternative terminologies) are used for states of affairs which are ongoing at reference time:

```
(30) di yong    lalaki … nag-íígi               ng  pantalon
     di iyón:LK lalaki  nag  -RDP1-igi          ng  pantalón
     so DIST:LK man      RLS.AV-RDP1-all.right GEN trousers
```
           So the man was adjusting his trousers …

These forms are also used for habitual or iterative states of affairs (cf. (40) below).

The realis/perfective forms (*perfective aspect* or *past tense*) are used for states of affairs which have been completed at reference time (see examples (7) and (9) above, among others).

Turning now to the voice formations, note first that the ensuing discussion primarily concerns dynamic verbs. The meanings of the voices in stative and potentive formations are slightly different, as discussed in 4.2. In dynamic formations, the actor voice forms usually imply that the subject is a volitional and controlling instigator of the event. The major exception to this rule are self-regulated processes (also known as "achievements") and natural phenomena, which are generally marked by the actor voice infix *-um-* as well. The subject of process verbs is a theme rather than an actor:

```
(31) g<um>anda  na   ang  buhay nilá
     <AV>beauty now SPEC life  3p.POSS
            (… and) their life became beautiful.
```

Expressions for natural phenomena are usually subjectless (*umúulán na* 'it's raining'). The choice of one or the other of the three actor voice affixes *-um-*, *mag-* and *maN-* sometimes conveys semantic differences pertaining to reflexivity, the intensity of the action and the like (cf. Pittman 1966, Lopez 1937:46-49, Drossard 1984:87-92; Himmelmann 1987:101f, 185-188). It is rare, however, that one base occurs with all three actor voice affixes, as in: *pumutol* 'cut', *magputol* 'cut several things', *mámutol* 'cut selectively or in quantity'.

Of the three affixes, *maN-* is the least common. It generally indicates intensive, distributive or repeated action, e.g., *bumilí* 'buy' vs. *mámilí* 'shop', *humampás* 'hit with a whip' vs. *mánghampás* 'go whipping'.

The major contrast is the one between *-um-* and *mag-*. Often *mag-* indicates a greater frequency or intensity of an action, cf. *bumasa* 'read' vs. *magbasa* 'to read a lot/study'. *mag-* and *-um-* may even co-occur to indicate a very high degree of intensity, cf. *mág<um>aral* 'study diligently' or *mag-s<um>igáw* 'shout (long and very loud)'. With regard to bases denoting position or motion, there is a regular contrast in transitivity: *t<um>ayóʔ kamí* (<AV>stand upright 1pe) 'we stood up' vs. *nag-tayóʔ kamí ng bahay* (RLS.AV-stand upright 1pe GEN house) 'we erected a house'. That is, the *-um-*-form denotes actors who move themselves while the *mag-*-form denotes actors who move something else. A similar contrast exists with respect to bases denoting qualities, e.g., *um-init* 'become/get hot' vs. *mag-init* 'make hot, heat'. Much quoted, but unique is the contrast between *bumilí* 'buy' and *mágbilí* 'sell'.

In other instances, however, it is difficult to determine what exactly the choice of the actor voice affix contributes to the meaning of the overall formation. That is, it appears to be simply a lexical feature of some bases to occur with *mag-* and others to occur with *-um-*.

Unlike the three actor voice affixes, the three undergoer voice affixes consistently differ with regard to the semantics of the undergoer. Hence it is customary to distinguish at least three undergoer voices in Tagalog. Ignoring several details and complications, it generally holds true that if the predicate is marked with the conveyance voice prefix *i-*, the subject expresses an argument bearing the semantic role of a displaced theme. Compare:

```
(32) ibinalík           nilá     ang  bataʔ
     i -in    -balík    nilá     ang  bataʔ
     CV-RLS(UG)-return   3p.POSS  SPEC child
            They returned the child.
```

Here the subject (*ang bataʔ*) is the displaced theme (i.e. the entity viewed as moving) of the event expressed by the predicate (*ibinalík*). Instruments are also viewed as moving entities and hence marked with the conveyance voice prefix:

(33) **i**pangpúpútol   ko   na lang   **itóng kutsilyo**
     i -paN-RDP1-putol ko   na lamang itó-ng kutsilyo
     CV-INS-RDP1-cut   1s.POSS now only   PRX-LK knife
            I will just cut it with this knife.

In instrumental uses, it is common that the conveyance voice prefix is added to a derived stem (as in the preceding example) rather than directly to the base. For this reason, a distinction is sometimes made between the conveyance voice proper and an instrumental voice.

    A third use of *i-* which is often mentioned is its use for beneficiaries, as in *I-hanap mo akó ng trabaho* (CV-look for 2s.POSS 1s GEN work) 'Look for some work for me'. It is quite difficult to establish a semantic link between this use and the two preceding ones. Note that benefactive use of *i-* is less frequent than the other two uses both in terms of tokens (text frequency) and types (bases for which a benefactive derivation is in common use; see also Himmelmann 1987:141-143).

    In locative voice (marked with *-an*), the subject expresses a locative argument, understood in a very broad sense. This may be the location at which something happened:

(34) tinirh**án**   ko   **ang bahay na itó**
     in   -tirá   -an ko   ang bahay na itó
     RLS(UG)-dwelling-LV 1s.POSS SPEC house LK PRX
            I stayed at this house.

Or the location to which (or from which) motion occurred:

(35) pinuntah**án**   na namán nilá   **ang bataʔ**
     in   -puntá -an na namán nilá   ang bataʔ
     RLS(UG)-direction-LV now also 3p.POSS SPEC child
            They went to the child.

Locative voice is also used for recipients, addressees, and beneficiaries inasmuch as these are also (more or less direct) recipients or goals of the action as in (36):

(36) títir**án**   ninyó **akó**
     RDP1-tirá   -an ninyó   akó
     RDP1-leftover-LV 2p.POSS 1s
            Will you (please) set some aside for me.

Even more generally, locative voice may be used for all kinds of undergoers which are not directly affected by the action denoted by the verb, as in (27) above and (37):

(37) tulung-**an** ninyó   **akó**
     help -LV 2p.POSS 1s
         (If) you help me, …

The different uses of locative voice as illustrated here are sometimes considered to be different voices. So a distinction is made between directional (or goal) voice,

benefactive voice, and locative voice proper (which often involves stem-formation via
pag- or *paN-*).

The suffix *-in*, marking patient voice, is the unmarked member of the undergoer-
voice-marking affixes and is used for a broad variety of undergoers, including
prototypical patients, i.e. entities directly affected or effected by the event denoted by
the predicate:

(38)   patay-**ín** natin   **itó-ng dalawá-ng Hapón**
      dead -PV 1pi.POSS PRX-LK two   -LK Japan
          Let's kill these two Japanese!

As already mentioned above, in realis mood patient voice is simply marked by the realis
undergoer voice infix *-in-*:

(38') p<**in**>atáy   natin   itó-ng dalawá-ng Hapón
      <RLS(UG)>dead 1pi.POSS PRX-LK two-LK   Japan
          We killed these two Japanese.

The choice among the different voices is determined by considerations of semantic role,
referentiality/definiteness and some other discourse pragmatic factors which are still
poorly understood (cf. Naylor 1975, McFarland 1978b, Adams & Manaster-Ramer
1988, and the discussion in HIMMELMANN, TYPOLOGICAL CHARACTERISTICS, SECTION
5). One fairly robust rule is the following: If there is a definite patient or theme
participant in the state of affairs to be expressed by a simple main clause, then patient or
conveyance voice are chosen (cf., for example, (32) and (38) above). That is, definite
patients or themes are extremely rare in actor voice constructions. But it is not
completely impossible for a formally definite patient or theme to occur in an actor voice
construction in a simple main clause. It can happen if the patient or theme participant is
either in possession of the subject of the construction (39) or it receives a partitive
interpretation (40):

(39)   nag-dá-dalá   silá **ng sarili nilá-ng**
      RLS.AV-RDP1-bring 3p   GEN own   3S.POSS-LK
      banda ng músika
      band GEN music
          They bring their own band.

(40)   namúmúlot   silá **noón**
      naN -RDP1-pulot silá noón
      RLS.AV-RDP1-pick.up 3p   DIST.GEN
      at ipinagbíbili   nilá
      at i -in   -pag-RDP1-bilí nilá
      and CV-RLS(UG)-SF -RDP1-sale 3p.POSS
          (Okay, their means of living was to pick fruit, those fruits which are used in
          chewing betel nuts …) They would pick (some of) those and then sell them,
          …

Note that if one continues to speak about such a formally definite patient or theme it has

to be made the subject of the following predicate (thus *ipinagbíbilí* is chosen here rather than *nagbíbilí*). Similarly if an indefinite patient or theme is referred to again in a subsequent predication, it has to be made the subject. Thus, *ibenta* rather than *magbenta* has to be chosen in the following context:

(41) nangúngúha   ngá   ng baging **para ibenta**
     naN   -RDP1-kuha   ngáʔ   ng bagin para i -benta
     RLS.AV-RDP1-getting indeed GEN vine   for CV-sales
             (My husband hasn't arrived yet) he is collecting vine for selling it.

It is also common to make indefinite patients and themes the subject if they are going to be major participants in the ongoing discourse, in particular if they are animate.

(42) doón   ay ná-kita   nilá   **ang isá-ng ma-lakí-ng higante**
     DIST.LOC PM RLS.POT-seen   3p.POSS SPEC one-LK ST-size-LK giant
             There they saw a great giant ... [the giant is going to be the main
             protagonist of the ensuing episode]

For other undergoer roles such as goals, beneficiaries, locations and instruments, similar tendencies exist. That is, whenever a participant bearing such a role is identifiable for speaker and hearer it tends to be made the subject, unless there is also a definite patient or theme participant involved. Actor voice, then, is a marked choice in all states of affairs which involve undergoers, in particular if the undergoers are definite. However, this does not mean that actor voice constructions are very infrequent in natural discourse. Since many semantically intransitive constructions involve actor voice marking, its overall text frequency is at least as high as the overall frequency of all undergoer voices combined.

The whole issue of choosing a voice is complicated by the fact that the preceding observations only hold for simple main clause constructions in which the verb functions as predicate. Different rules obtain in cleft constructions, relative clauses, and existential constructions, in which the verb does not occur in predicate position. Most importantly, in these constructions there is no restriction on definite patients or themes for actor voice verbs. Compare the following example where the actor voice verb *mangibig* occurs in a relative clause and takes a definite undergoer argument:

(43) sa mga lalaki na **mangibig nung**   **kanyáng   anák**
     sa mga lalaki na maN-ibig noón   :LK kanyá -ng anák
     LOC PL man   LK AV-love DIST.GEN:LK 3s.DAT-LK child
             (So he held a contest) between the men who courted his child …

That is, in these constructions there is no inherent preference for the undergoer voices. Instead, the syntax of the overall construction determines which voice has to be chosen.

Thus, a complete account of voice choice also has to include a discussion of the factors which determine the choice of a construction in which the verb is not a main clause predicate. There are in particular two such constructions, the existential construction and the cleft construction.

Existential constructions can be used whenever a state of affairs involves an indefinite participant, especially participants for which no basic lexical designation exists. In this instance, it is common to use the verb itself as a description for the participant. Thus, in the following example, *nagáalaga* is used as a descriptive label for the person looking after the birds:

```
(44) ay mayroón paláng  nagáalaga              doón      sa  ibun
     ay mayroón palá-ng nag  -RDP1-alagaʔ      doón      sa  ibon
     PM EXIST   so! -LK RLS.AV-RDP1-cared.for DIST.LOC LOC bird
              (But he didn't know) there was already someone looking after those birds.
```

In cleft constructions, focal participants (and in particular, contrastive ones) are put in predicate position while the remainder of the clause, including the verb, occurs in a noun phrase. This construction is regularly chosen for most question words (cf. also example (7) above):

```
(45) ano   ang gágawín      ng langgám sa dahon
     anó   ang RDP1-gawáʔ-in ng langgám sa dahon
     what SPEC RDP1-act  -PV GEN ant       LOC leaf
          (He thought:) What will the ant do on the leaf?
```

To summarize: In choosing a voice form, a Tagalog speaker has to decide first which construction will be appropriate. If a construction is chosen in which the verb does not occur in predicate position, voice choice is determined by the syntax of the construction. If a simple main clause construction is chosen for expressing a transitive state of affairs, there is a strong preference for undergoer voices, in particular if definite patient or theme participants are involved.

Despite the large number of options existing in principle for the expression of a particular state of affairs, the freedom of choice remains limited. For most everyday situations there are conventional ways in which these situations can be referred to. In terms of voice affixation this means that for each base there are typically only one or two voice affixes it frequently combines with.

4.2. Stative and potentive

With a few exceptions, the verbs discussed in the preceding section are only used for states of affairs which involve a volitional and controlling actor (regardless of whether or not this actor is overtly expressed). States of affairs which do not involve a controlling actor are expressed by a related but different set of formations. Two different scenarios have to be distinguished here. On the one hand, the state of affairs may be such that it excludes the involvement of an actor for principled conceptual reasons. This is the case for *stative* expressions such as 'be hungry', 'be angry', 'be adrift', and the like. On the other hand, the state of affairs may be such that in principle it allows for controlling actors but in the specific instance at hand the actor is not in full control of the event. This is the case when someone happens to do something without

having the intention to do it. Formations expressing this second possibility are called *potentive*.

A completely regular correspondence exists between dynamic and potentive formations. That is, for each dynamic verb form there is a corresponding potentive form. The major potentive formatives are *maka-* for actor voice and *ma-* for the undergoer voices. For details compare Table 3 with Table 4.

Table 4: Potentive aspect/mood paradigm for *bilí* 'purchase, sale'

	AV ('buy')	AV ('sell')	PV	LV	CV
NON-REALIS/ PERFECTIVE	maka-bilí	maka-pagbilí	ma-bilí	ma-bilh-án	ma-i-bilí
NON-REALIS/ IMPERFECTIVE	maka-bíbilí	maka-pagbíbilí	ma-bíbilí	ma-bíbilh-án	ma-i-bíbilí
REALIS/ PERFECTIVE	naka-bilí	naka-pagbilí	na-bilí	na-bilh-án	na-i-bilí
REALIS/ IMPERFECTIVE	naka-bíbilí	naka-pagbíbilí	na-bíbilí	na-bíbilh-án	na-i-bíbilí

The typical use of potentive forms is for involuntary actions:

(46) biglá   niyá-ng   **ná-**bigkás   iyón
     sudden 3s.POSS-LK RLS.POT-enunciation DIST
          (Terrified) she suddenly exclaimed: …

This includes actions done accidentally, i.e. the actor may be in control of the action but did not really intend its outcome:

(47) **na-i-**luto   ko   na
     RLS.POT-CV-cooked 1s.POSS now
          I happen to have cooked it already (by mistake).

It also includes perceptions over which the actor (= experiencer) has no control (cf. (42) above).

In a second, somewhat different use potentive forms express the ability of an actor to perform the action in question:

(48) kung inyóng   **ma**pagtiis**án**   iyán
     kung inyó -ng ma -pag-tiís-an   iyán
     if   2p.DAT-LK POT-SF -suffer-LV MED
          If you are able to endure this …

```
(49) at  hindí  makabaríl   sa  kanyá
     at  hindí? maka-baríl   sa  kanyá
     and NEG    POT.AV-gun   LOC 3s.DAT
```
        (The man got bitten by the ants) and wasn't able to shoot at him.

Turning now to stative verbs, these also come in four different voices, two of which are formally identical to the potentive voice forms. But despite these formal similarities and the fact that essentially the same glosses are used for dynamic and stative voices, the syntax and semantics of the stative voices differs quite clearly from the potentive formations so that the category labels are not really indicative of their functions.

In the basic voice form for statives, which is simply called *stative* here, the subject is a theme, i.e. an entity which is in, or currently is undergoing, a given state. The forms are identical to the patient voice forms of the potentive paradigm, i.e. base plus prefix *ma-*. The forms of the stative actor voice are morphologically identical to the potentive actor voice forms, i.e. base plus *maka-*. The stative locative and conveyance voices are marked by the prefix *ka-* to which the basic voice affixes *-an* and *i-*, respectively, are added. See Table 5 for an overview of the forms.

Table 5: Voice and aspect/mood paradigm for stative verbs (base *galit* 'anger')

	ST	ST.LV	ST.CV	ST.AV
NON-REALIS/ PERFECTIVE	ma-galit	ka-galit-an	i-ka-galit	maka-galit
NON-REALIS/ IMPERFECTIVE	ma-gágálit	ka-gágalit-an	i-ka-gágalit	maka-gágálit
REALIS/ PERFECTIVE	na-galit	kina-galit-an	i-kina-galit	naka-galit
REALIS/ IMPERFECTIVE	na-gágálit	kina-gágalit-an	i-kina-gágalit	naka-gágálit

Once again it is essential to note that the forms listed in Table 5 differ substantially with regard to their productivity. The only really productive formation is the basic stative formation. Almost every Tagalog content word base (and many derived stems) can be prefixed with *ma-* and express a state:

```
(50) na-galit       siyá
     RLS.ST-anger 3s
```
        She was/got angry.

With bases such as *basag* 'crack, break', which allow both a state and an action reading, the form *nabasag* is ambiguous: It can mean 'be in a broken state' (stative) or 'happen

to break/able to break' (potentive patient voice). In context, these readings are generally distinguished by the presence of an overt actor expression in the potentive use (*nabasag niyá* 's/he happened to break it/was able to break it').

The stative locative voice is common with bases expressing emotions. The subject expresses the person or thing at which the emotion is directed:

(51) **kina**galit**an**   siyá ng Nanay
     in   -ka-galit-an siyá ng nanay
     RLS(UG)-ST-anger-LV 3s   GEN mother
             Mother was angry with him/her.

Frequently, stative locative voice derivations take on some more specialized meanings. Thus, *kagalitan* also means 'to reprove, to scold, to rebuke'. In addition, stative locative voice derivations are possible with a (relatively small) number of stative expressions which do not pertain to emotions. They then denote the place at which a given state occurs (e.g., *ka-hulug-an* 'place where someone falls' (< *hulog* 'fall')).

The stative conveyance voice is also most common with bases expressing emotions. Stative conveyance voice formations always have the connotation of causation, that is, the subject specifies the reason for the emotion:

(52) **ikina**galít   niyá   akó
     i -in   -ka-galit niyá   akó
     CV-RLS(UG)-ST-anger 3s.POSS 1s
             She got angry at me (I was the reason for her being angry).

The stative conveyance voice is found with a somewhat broader range of bases than the stative locative voice. These include *ikabasáʔ* 'get wet on account of', *ikabagsák* 'fall on account of', *ikatawa* 'laugh on account of', *ikaiyák* 'cry on account of', etc.

The stative actor voice is very similar in meaning to the stative conveyance voice since it also specifies the cause for a given state. But the two formations differ in their grammar and productivity. In the stative conveyance voice construction (as in the stative locative voice construction), the theme argument (i.e. the one who experiences an emotion in the case of emotion verbs) is grammatically coded as a genitive argument. In the stative actor voice construction it is a locative argument:

(53) lahát ng kanyáng sabihin   ay nakagágálit   **sa**
     lahát ng kanyá -ng sabi   -in ay naka   -RDP1-galit sa
     all   GEN 3s.DAT-LK statement-PV PM RLS.ST.AV-RDP1-anger LOC
     akin
     akin
     1s.DAT
             Everything he says irritates me.

The subject expression in the stative actor voice construction usually refers to an inanimate cause (some state of affairs or a thing). With regard to productivity, the stative actor voice forms are the least common of all stative formations and whenever they occur they often take on somewhat specialized meanings (thus *makagalit* is

'irritate, antagonize, give offense' rather than a plain 'make angry'). Furthermore, the stative actor voice derivations are often conventionalized in one of the four aspect/mood forms, for example, *naka-áawaʔ* 'arousing pity, pitiable' (< *awaʔ* 'mercy, compassion'), *nakáka-litó* (or *naka-lílitó*) 'confusing' (< *litó* 'confused, at a loss'), or *nakáka-gandá* (or *naka-gágandá*) 'beautifying' (< *gandá* 'beauty').

5.   Nominalizations and nominal morphology

As pointed out in section 3.3, it is somewhat misleading to speak of nouns and verbs in Tagalog if these terms are understood to have a similar meaning as in English. Strictly speaking, we are dealing in this section with formations which are not voice-marked. Inasmuch as voice-marking is taken to be the defining characteristic of Tagalog 'verbs', these formations are clearly non-verbal. Whether it is useful to call them nominals or nominalizations is a moot question which will be ignored here. Considerations of space prevent the review of all relevant formations. The major focus is on gerunds.

5.1.   Gerunds

Gerunds provide a way to refer to actions or states without orienting them towards one of the participants involved in the action or state via voice-marking. They may denote a type of activity as in:

(54) **pag-lu-luto**ʔ ng pagkain ang trabaho niyá
      GER-RDP-cook GEN food   SPEC work   3s.POSS
            His/her job is cooking food.

Gerunds are regularly derived from actor voice forms, usually by replacing the voice affix with *pag-* or *paN-* (in some instances, unstressed CV-reduplication is also part of the process). Stative gerunds are derived from the basic stative voice by replacing *ma-* with *pagka-*.

   Since gerunds are not voice-marked, all participants involved in the state of affairs denoted by the gerund have to be expressed as non-subject arguments (i.e. either as genitive or locative arguments). This, however, does not mean that gerunds cannot be used as predicates, as shown by example (54) above. As with all Tagalog content words, gerunds can be used in any syntactic function, provided their meaning fits. In the following example, a (stative) gerund functions as the subject of the potentive patient voice predicate *maáalis*:

(55) maáalís   lámang ang pagkaahas   mo
      ma   -RDP1-alís lamang ang pag-ka-ahas   mo
      POT-RDP1-departure only   SPEC GER-ST-snake 2s.POSS
            (You will become a snake and) you will only be able to leave your
            snakehood (your state of being a snake) …

The most common use of gerunds, however, is their use in the function of reduced

temporal clauses. In this use, they designate specific instances of the action or state denoted by the base:

(56) **pag-datíng** namin doón
    GER-arrival 1pe.POSS DIST.LOC

    in-iwan namin don ang bangka?
    RLS(UG)-abandon 1pe.POSS DIST.LOC SPEC boat
        When we arrived there we abandoned the boat, …

(57) at **pag-ká-sabi** niyá nitó
    and GER-ST-say 3S.POSS PRX.GEN
        … and when he had said this ...

A gerund preceded by *sa* may also be used for reduced complement clauses:

(58) pagbawalan mo ang batang iyón **sa**
    pag-bawal -an mo ang bata?-ng iyón sa
    SF -forbidden-LV 2s.POSS SPEC child-LK DIST LOC

    paglalaró? sa lansangan
    pag-RDP-laró? sa lansangan
    GER-RDP-play LOC street
        Forbid that child to play in the street.

5.2. Abstracts, instruments, locations, etc.

Among the other very productive 'nominalizing' processes are the formations for abstracts, instruments and locations. Most of these are segmentally identical, or at least very similar, to the verbal formations discussed in the preceding section. Note, however, that inasmuch as both verbal and 'nominal' derivations are possible from the same base they are usually distinguished by different stress patterns.

- *ka--an* derives abstracts and collectives: *kamurahan* 'cheapness' (< *mura* 'cheap'), *katámáran* 'laziness' (< *tamád* 'lazy'), *kasabihán* 'proverb' (< *sabi* 'statement'), *káratnán* 'possible result, outcome' (< *datíng* 'arrival'), *kapuluán* 'archipelago' (< *puló?* 'island'), *kabundukán* 'mountains' (< *bundók* 'mountain').

- *-an* derives designations for places associated with the base: *bílíhan* 'market, trading place', *aklatan* 'library' (< *aklat* 'book'), *híráman* 'place for borrowing' (< *hirám* 'borrowed', cf. locative voice *hiramán* 'to borrow from someone'), *pintuan* 'doorway' (< *pintó?* 'door'), *itlugan* 'place where eggs are laid' (< *itlóg* 'egg').

- *maN-* plus unaccented CV-reduplication derives a designation for someone professionally occupied with what the base designates: *mámimili* 'shopper, buyer, customer', *mambabasa* 'reader' (< *basa* 'read'), *manlalaro?* 'player' (< *laró?* 'play'), *manggagamot* 'physician' (< *gamót* 'medicine, cure').

- *paN-* (with, in some instances, optional replacement of the base-initial consonant)

derives designations for instruments/means: *pambilí* 'means for buying', *panghampás* 'sth. for hitting' (< *hampás* 'blow, strike'), *pansará/panará* 'fastener, lock' (< *sará* 'closed'), *pampatabáʔ* 'sth. which makes fat, fertilizer' (< *tabáʔ* 'fat, grease').

Acknowledgments

References

Adams, K.L. and Manaster-Ramer, A. (1988) 'Some questions of topic/focus choice in Tagalog', *Oceanic Linguistics* 27:79-101.

Arsenio, M.E. (1971) *A lexicographic study of Tayabas Tagalog of Quezon Province*, Quezon City: University of the Philippines Press.

Bader, T., Werlen, I. and Wymann, A. (1994) *Towards a typology of modality*, Arbeitspapier 32, Bern: Institut für Sprachwissenschaft.

Blake, F.R. (1906) 'Expression of case by the verb in Tagalog', *Journal of the American Oriental Society* 27:183-189.

---- (1925) *A grammar of the Tagalog language*, New Haven: American Oriental Society.

Bloomfield, L. (1917) *Tagalog texts with grammatical analysis*, Urbana: The University of Illinois.

---- (1933) *Language*, London: Allen and Unwin.

Chan Yap, G. (1980) *Hokkien Chinese borrowings in Tagalog*, Canberra: Pacific Linguistics.

Constantino, E. (1965) 'The sentence patterns of twenty-six Philippine languages', *Lingua* 15:71-124.

---- (1971) 'Tagalog and other major languages of the Philippines', in T.A. Sebeok (ed.) *Current Trends in Linguistics*, vol. 8/1: 112-154, The Hague: Mouton.

---- (1999) *The Contemporary English-Filipino Dictionary*, Quezon City: Center for Philippine Languages.

Cruz, E.L. (1975) *A subcategorization of Tagalog verbs*, The Archive Special Monograph No.2, Quezon City: University of the Philippines.

DeGuzman, V.P. (1978a) *Syntactic derivation of Tagalog verbs*, Oceanic Linguistics Special Publication 16, Honolulu: University Press of Hawaii.

---- (1978b) 'A case for nonphonological constraints on nasal substitution', *Oceanic Linguistics* 17:87-106.

---- (1986) 'The subject of identificational sentences', in B.F. Elson (ed.) *Language in global perspective*, 347-362, Dallas: Summer Institute of Linguistics.

---- (1997) 'Verbal affixes in Tagalog: Inflection or derivation?', in C. Odé and W. Stokhof (eds) *Proceedings of the Seventh International Conference on Austronesian Linguistics*, 303-325, Amsterdam: Rodopi.

---- (2001) 'Tagalog', in J. Garry and C. Rubino (eds) *Facts about the world's languages*, 703-707, New York: New England Publishing Associates.

Drossard, W. (1984) *Das Tagalog als Repräsentant des aktivischen Sprachbaus*, Tübingen: Narr.

English, L.J. (1977) *English- Tagalog dictionary*, Manila: National Book Store.

---- (1986) *Tagalog-English dictionary*, Manila: National Book Store.

Gil, D. (1993) 'Tagalog semantics', *Berkeley Linguistic Society* 19:390-403.

Gonzales, A.B. (1970) 'Acoustic correlates of accent, rhythm, and intonation in Tagalog', *Phonetica* 22:11-44.

Grimes, B.F. (1999), *Ethnologue - Languages of the world*, 13th edition, Dallas: Summer Institute of Linguistics.

Himmelmann, N.P. (1987) *Morphosyntax und Morphologie - Die Ausrichtungsaffixe im Tagalog*, München: Fink.

---- (1991) *The Philippine challenge to universal grammar*, Arbeitspapier Nr. 15, Köln: Institut für Sprachwissenschaft.

---- (1998) 'Regularity in irregularity: Article use in adpositional phrases', *Linguistic Typology* 2:315-353.

---- (1999) 'The lack of zero anaphora and incipient person marking in Tagalog', *Oceanic Linguistics* 38: 231-269.

---- (to appear) 'Lexical categories and voice in Tagalog', in P. Austin and S. Musgrave (eds) *Voice and grammatical functions in Austronesian languages*, Stanford: Center for the Study of Language and Information.

Humboldt, W. von (1838) *Über die Kawi Sprache*, Bd. II, Berlin: Dümmler.

Kroeger, P.R. (1993) *Phrase Structure and Grammatical Relations in Tagalog*, Stanford: Stanford University Press.

Lemaréchal, A. (1982) 'Sémantisme des parties du discours et sémantisme des relations', *Bulletin de la Société de Linguistique de Paris* 77:1-39.

---- (1989) *Les parties du discours. Sémantique et syntaxe*, Paris: P.U.F.

Lopez, C. (1937) 'Preliminary Study of the Affixes in Tagalog', in E. Constantino (ed.) (1977) *Selected writings of Cecilio Lopez in Philippine linguistics*, 28-104, Diliman: University of the Philippines Press.

---- (1970) 'On the Boak Tagalog of the island of Marinduque', *The Archive* (University of the Philippines) I/2:1-53.

Marre, A. (1901) 'Grammaire tagalog, composé sur un nouveau plan', *Bijdragen tot de Taal-, Land- en Volkenkunde* 53:547-92.

Matsuda French, K. (1988) *Insights into Tagalog. Reduplication, infixation, and stress from nonlinear phonology*, Dallas: Summer Institute of Linguistics.

McFarland, C.D. (1976) *A provisional classification of Tagalog verbs*, Tokio: Institute for the Study of Languages and Cultures of Asia and Africa.

---- (1978a) 'Tagalog existentials', *Philippine Journal of Linguistics* 9:1-13.

---- (1978b) 'Definite objects and subject selection in Philippine languages', in C. Edrial-Luzares and A. Hale (eds) *Studies in Philippine Linguistics*, vol. 2/1:139-182, Manila: Linguistic Society of the Philippines.

---- (^{2}1983) *A linguistic atlas of the Philippines*, Manila: Linguistic Society of the Philippines.

Müller, F. (1882) *Grundriss der Sprachwissenschaft*, II. Band II. Abt., *Die Sprachen der Schlichthaarigen Rassen*, Wien: Alfred Hölder.

Naylor, P.B. (1975) 'Topic, focus, and emphasis in the Tagalog verbal clause', *Oceanic Linguistics* 14:12-79.

---- (1980) 'Linking, relation-marking, and Tagalog syntax', in P.B. Naylor (ed.) *Austronesian syntax. Papers from the Second Eastern Conference on Austronesian Languages*, 33-49, Ann Arbor: Center for South and Southeast Asian Studies.

---- (1995) 'Subject, topic, and Tagalog syntax', in D. Benett, T. Bynon and G.B. Hewitt (eds) *Subject, voice and ergativity*, 161-201, London: School of Oriental and African Studies.

Panganiban, J.V. (1972) *Diksyunario-tesauro Pilipino-Ingles*, Quezon City: Manlapaz.

Pittman, R. (1966) 'Tagalog -um- and -mag-. An interim report', in *Papers in Philippine Linguistics* 1, 9-20, Canberra: Pacific Linguistics.

Ramos, T.V. (1974) *The case system of Tagalog verbs*, Canberra: Pacific Linguistics.

Ramos, T.V. and Bautista, M.L.S. (1986) *Handbook of Tagalog verbs. Inflections, modes, aspects*, Honolulu: University of Hawaii Press.

Reid, L.A. (1981) 'Philippine linguistics: the state of the art: 1970-1980', in D.W. Hart (ed.) *Philippine Studies: political science, economics, and linguistics*, 223-230, Northern Illinois University: Center for Southeast Asian Studies.

---- (1992) 'On the Development of the Aspect System in Some Philippine Languages', *Oceanic Linguistics* 31:65-91.

Rubino, C.R.G. (1998) *Tagalog standard dictionary*, New York: Hippocrene Books.

---- (1998a) 'The morphological realization and production of a nonprototypical morpheme: the Tagalog derivational clitic, *Linguistics* 36:1147-1166.

Santos, V.C. (1983) *Pilipino-English dictionary*, Metro Manila: National.

Schachter, P. (1973) 'Constraints on clitic order in Tagalog', in A.B. Gonzales (ed.) *Parangal kay Cecilio Lopez*, 214-231, Quezon City: Linguistic Society of the Philippines.

---- (1976) 'The subject in Philippine languages: topic, actor, actor-topic or none of the above', in C.N. Li (ed.) *Subject and topic*, 491-518, New York: Academic Press.

---- (1987) 'Tagalog', in B. Comrie (ed.) *The world's major languages*, 936-958, London: Croom Helm.

---- (1995) 'Tagalog', in J. Jacobs et al. (eds) *Syntax*, vol. II:1418-1430, Berlin: Mouton de Gruyter.

Schachter, P. and Otanes, F.T. (1972) *Tagalog reference grammar*, Berkeley: University of California Press.

Shkarban, L.I. (1995) *Grammatičeskij stroj tagal'skovo jazyka*, Moskva: Nauka.

Soberano, R. (1980) *The dialects of Marinduque Tagalog*, Canberra: Pacific Linguistics.

Totanes, S. de (1865) *Arte de la lengua tagala y manual tagalog*, 4th edition, Binondo: Miguel Sanchez y Ca.

Wolff, J.U. (1976) 'Malay borrowings in Tagalog', in C.D. Cowan and O.W. Wolters (eds) *Southeast-Asian history and historiography: essays presented to D.G.H. Hall*, 345-367, Ithaca: Cornell University Press.

Wolff, J.U. with M.T.C. Centeno and D.V. Rau (1991) *Pilipino through self-instruction*, 4 vols., Ithaca: Cornell Southeast Asia Program.

Zorc, R.D. (1994) 'Tagalog', in D.T. Tryon (ed.) *Comparative Austronesian dictionary*, vol. 1:335-341, Berlin: Mouton de Gruyter.