Filipino Grammar

Stephen Borja, Justin Ching, Zhean Ganituen

February 22, 2025

1 Preliminaries

1.1 The Filipino Alphabet

Let $\mathcal{F}$ the Filipino language and $\mathbb{F}$ be the Alphabet of $\mathcal{F}$, this alphabet is composed of 56 scripts and 11 punctuation marks [2]. The 56 scripts are divided into two, the first half being the capital letters of the modern Latin script with the addition of "Ñ" and "Ng"; while the other half is the lower case variants of each letter.

The 11 punctuation marks in the Filipino language are the: *tuldok* (.), *tandang pananong* (?), *tandang padamdam* (!), *kuwit* (,), *kudlit* ('), *gitling* (-), *tutuldok* (:), *tuldok-kuwit* (;), *panipi* ("), *pambukas na panaklong* ((), *pampasarang panaklong* ()), at ang *tutuldok-tuldok* (...)

In mathematical notation, we can represent $\mathbb{F}$ as the set:

$$\mathbb{F} = \{\mathrm{a, b}, \ldots, \mathrm{z}, \tilde{\mathrm{n}}, \mathrm{ng}, \ \mathrm{A, B, C}, \ldots, \mathrm{Z}, \tilde{\mathrm{N}}, \mathrm{Ng}\} \ \cup \{., ?, !, ,, ', \text{-}, :, \ ;, ", (,), \ldots\}$$

and the size of $\mathbb{F}$, $|\mathbb{F}| = 67$.

From this we can introduce the following sets:

1. $\mathbb{M} = \{.,?,!,,,',\text{-},:,;,",(,),...\}$, the set of punctuation marks

2. $\mathbb{V} = \{\mathrm{a,e,i,o,u,A,E,I,O,I}\}$, the set of upper and lower case vowels

3. $\mathbb{C} = \mathcal{F} - (\mathbb{M} \cup \mathbb{V})$, the set of upper and lower case consonants

4. $\mathbb{V}_{\mathrm{upper}}$ is the set of upper case vowels

5. $\mathbb{V}_{\mathrm{lower}}$ is the set of lower case vowels

6. $\mathbb{C}_{\mathrm{upper}}$ is the set of upper case consonants

7. $\mathbb{C}_{\mathrm{lower}}$ is the set of lower case consonants

8. $\mathbb{L} = \mathbb{F} - \mathbb{M}$, is the set of consonants and vowels

9. $\mathbb{L}*$, is the set of words in $\mathcal{F}$

Remark on the Digraph: ng/Ng or "*en dyi*"   Although the letter
"Ng" or "ng" is a concatenation of two separate graphemes or symbols in $\mathbb{F}$
(since Ng = N · g and ng = n · g), the letter "Ng" is officially recognized as
a grapheme in $\mathbb{F}$ since it represents a distinct Filipino sound. In particular,
it represents the *voiced velar nasal sound*, or in the International Phonetic
Alphabet (IPA), the ŋ sound [1].

For instance, the word "hangin" has 5 letters namely: "h", "a", "ng", "i",
"n", since "ng" is pronounced as a velar nasal sound, not as two separate
sounds n-g. So, "hangin" is pronounced as haŋin ("ha-ngin"). Take for
instance the English word "manger" where "ng" is a substring but is not
pronounced as the velar nasal sound. Instead, its pronunciation is ˈmeɪ ndʒ
ər ("meyn-jer"); not ˈmæŋ ʒ ər ("mang-jer"), ˈmæŋ ər ("mang-er") or ˈmæŋər
("manger").

1.2   Common Errors

1.2.1   "es"-words and "is"-words

Given the Spanish and English roots of Filipino, some *loan* words have rules
for Filipino spelling. Let s be any string, the English language $\mathcal{E}$, the Spanish
language $\mathcal{S}$, and $\mathcal{F}(s)$ is the translation of s in $\mathcal{F}$. Furthermore, let $\mathbb{S}*$ be
the set of words in $\mathcal{S}$ and $\mathbb{E}*$ be the set of words in $\mathcal{E}$.

1. $\left(\forall s \in \mathbb{S}*, s \xrightarrow{\text{regex}} \text{es}(\mathbb{S}*) \right) \rightarrow \left(\mathcal{F}(s) \xrightarrow{\text{regex}} \text{es}(\mathbb{M}|\mathbb{C})* \right)$

2. $\left(\forall s \in \mathbb{E}*, s \xrightarrow{\text{regex}} \text{s}(\mathbb{E}*) \right) \rightarrow \left(\mathcal{F}(s) \xrightarrow{\text{regex}} \text{is}(\mathbb{M}|\mathbb{C})* \right)$

Rule (1) denotes that if s is a Spanish word, translating s to a Filipino
word would use "es" as the prefix to the word to denote that $\mathcal{F}(s)$ is a word
of Spanish origin. On the other hand, for rule (2), if s is an English word,
then $\mathcal{F}(s)$ would use "is" as the prefix to the word to denote that it is of
English origin.

Example 1. *Here are examples of English and Spanish loan words in Fil-
ipino:*

1. Ako ay papasok sa **eskwelahan**. (correct)
   Ako ay papasok sa *iskwelahan*. (incorrect)
   eskwelahan ($\mathcal{F}$) from escuela ($\mathcal{S}$)

2. Aba! Malaki pala ang **espasyo** rito. (correct)
   Aba! Malaki pala ang *ispasyo* rito. (incorrect)
   espasyo ($\mathcal{F}$) from espacio ($\mathcal{S}$)

3. Marami kaming **estudyante** sa Computer Science. (correct)
   Marami kaming *istudyante* sa Computer Science. (incorrect)
   estudyante ($\mathcal{F}$) from estudiante ($\mathcal{S}$)

4. Mahilig sila Turing at Sipser maglaro ng **eskrima**. (correct)
   Mahilig sila Turing at Sipser maglaro ng *iskrima*. (incorrect)
   eskrima ($\mathcal{F}$) from esgrima ($\mathcal{S}$)

5. Marami raw **espiritu** rito. (correct)
   Marami raw *ispiritu* rito. (incorrect)
   espiritu ($\mathcal{F}$) from espiritu ($\mathcal{S}$)

6. Kinausap mo na ba iyung **ispiker**. (correct)
   Kinausap mo na ba iyung *espiker*. (incorrect)
   ispiker ($\mathcal{F}$) from speaker ($\mathcal{E}$)

7. Marami naman daw **isports** na puwedeng pagpilian. (correct)
   Marami naman daw *esports* na puwedeng pagpilian. (incorrect)
   isports ($\mathcal{F}$) from sports ($\mathcal{E}$)

8. Si Dijkstra ay hindi raw **iskolar**. (correct)
   Si Dijkstra ay hindi raw *eskolar*. (incorrect)
   iskolar ($\mathcal{F}$) from scholar ($\mathcal{E}$)

Example 2. *Here are some examples of literature and webpages where the "es-" and "is-" prefixes are misused:*

1. **TODO**

1.2.2 "kump"-words and "kumb"-words

If $s \in \mathbb{S}*$ and the the prefix of s is given by the regular expression (C|c)on(f|v). Then, $\mathcal{F}(s)$ is prefixed with (C|c)um(p|b).

In particular:

- $s \xrightarrow{\text{regex}} (\text{conf})\mathbb{S}* \rightarrow \mathcal{F}(s) \xrightarrow{\text{regex}} (\text{kump})\mathbb{L}*$

- $s \xrightarrow{\text{regex}} (\text{conv})\mathbb{S}* \rightarrow \mathcal{F}(s) \xrightarrow{\text{regex}} (\text{kumb})\mathbb{L}*$

Example 3. *Here are examples of Spanish words in Filipino with the "kump-" and "kumb-" prefixes:*

1. Pumasok si Cormen sa **kumbento**. (correct)
   Pumasok si Cormen sa *konbento*. (incorrect)
   kumbento ($\mathcal{F}$) from convento ($\mathcal{S}$)

2. Tinanong ko si Leiserson bilang **kumpirmasyon**. (correct)
   Tinanong ko si Leiserson bilang *kunpirmasyon*. (incorrect)
   kumpirmasyon ($\mathcal{F}$) from confirmacion ($\mathcal{S}$)

3. Ang **kumpetisyon** sa agham ay ginanap sa paaralan. (correct)
   Ang *kunpetisyon* sa agham ay ginanap sa paaralan. (incorrect)
   kumpetisyon ($\mathcal{F}$) from competición ($\mathcal{S}$)

4. Nagbigay siya ng **kumpisal** sa pari noong Linggo. (correct)
   Nagbigay siya ng *kunpisal* sa pari noong Linggo. (incorrect)
   kumpisal ($\mathcal{F}$) from confesión ($\mathcal{S}$)

5. Sumali ako sa isang **kumbento** sa simbahan natin. (correct)
   Sumali ako sa isang *kunbento* sa simbahan natin. (incorrect)
   kumbento ($\mathcal{F}$) from convento ($\mathcal{S}$)

6. Hindi ko alam ang **kumbinasyon** sa Graph Coloring. (correct)
   Hindi ko alam ang *kunbinasyon* sa Graph Coloring. (correct)
   kumbinasyon ($\mathcal{F}$) from combinación ($\mathcal{S}$)

7. Ipinahayag niya ang kanyang **kumbersasyon** sa kanila. (correct)
   Ipinahayag niya ang kanyang *kunbersasyon* sa . (incorrect)
   kumbersasyon ($\mathcal{F}$) from conversación ($\mathcal{S}$)

8. Ang patakaran ay dumaan sa **kumpigurasyon** bago ipatupad. (correct)
   Ang patakaran ay dumaan sa *kunpigurasyon* bago ipatupad. (incorrect)
   kumpigurasyon ($\mathcal{F}$) from configuración ($\mathcal{S}$)

Example 4. *Here are some examples of literature and webpages where the "kump-" and "kumb-" prefixes are misused:*

1. **TODO**

5

1.2.3 Morphophonemic Alteration with Suffixes

Let $k \in \mathbb{L}*$ be a word formed by the concatenation as follows $k = \alpha \cdot \text{root}_{\text{allomorph}} \cdot \omega$ where α is the prefix of k, ω is the suffix of k, and $\text{root}_{\text{allomorph}}$ is the root allomorph of k.

The morphophonemic alteration of the root word is given by:

If the root ends in "e", then $\text{root}_{\text{allomorph}}$ ends in "i".
If the root ends in "o", then the $\text{root}_{\text{allomorph}}$ ends in "u".

That is to say, that if a Filipino word is the concatenation of a root word that ends in "e" or "o" and a suffix. Then, "e" will change to "i" and "o" will change to "u" [3].

Example 5. *Here are examples of morphophonemic alterations in Filipino:*

1. **Tinakbuhan** niya ang kanyang problema.
   takbo $\rightarrow$ takbu $\cdot$ han $\rightarrow$ tinakbuhan

2. **Pinasukan** niya ang silid-aralan.
   pasok $\rightarrow$ pasuk $\cdot$ an $\rightarrow$ pinasukan

3. **Pinayuhan** siya ng kanyang guro.
   payo $\rightarrow$ payu $\cdot$ han $\rightarrow$ pinayuhan

4. **Babaihan** ang tawag sa lugar kung saan maraming babae.
   babae $\rightarrow$ babai $\cdot$ han $\rightarrow$ babaíhan

5. **Biniruan** niya ang kanyang kaibigan.
   biro $\rightarrow$ biru $\cdot$ an $\rightarrow$ biniruan

6. Gusto naming **laruin** ni Knuth iyan!
   laro $\rightarrow$ laru $\cdot$ in $\rightarrow$ laruin

7. Tinanong ko siya kung hahaluin iyong pagkain...
   halo $\rightarrow$ halu $\cdot$ in $\rightarrow$ haluin

8. Kakayanin kitang **talunin** sa Catan!
   talo $\rightarrow$ talu $\cdot$ nin $\rightarrow$ talunin

Example 6. *Here are some examples of literature and webpages where that have the incorrect root allomorph:*

1. **TODO**

1.2.4 Reduplications

A full reduplication is a word $k \in \mathbb{L}*$ and is defined by the concatenation

$$k = \text{"b"} \cdot \text{"-"} \cdot \text{"b"}$$

or it is a repetition of a root word b joined by the *gitling* "-". In a full reduplication, the repetition of the b has no variation from the first b.

On the otherhand, an ablaut reduplication is a word $j \in \mathbb{L}*$ and is defined by the concatenation

$$j = \text{"x"} \cdot \text{"r"}$$

or it is the repetition of the root word x and the root word with a vowel alteration r. Notice that for ablaut repetitions, there is no *gitling* joining the repetitions. Furthermore, the ablaut repetition of a root word creates a word with a different meaning. The ablaut or vowel alteration is defined by:

If the root ends in "e", then the ablaut ends in "i".
If the root ends in "o", then the ablaut ends in "u".

Example 7. *Here are examples of reduplications in Filipino:*

1. Kung **ano-ano** ginawa ni Hopcroft kanina. (correct)
   Kung *anu-ano* ginawa ni Hopcroft kanina. (incorrect)
   ano-ano is a full reduplication

2. Pupunta kami mamaya sa **sari-sari** store. (correct)
   Pupunta kami mamaya sa *sarisari* store. (incorrect)
   sari-sari is a full reduplication

3. Gagawin natin iyong **kanya-kanyang** gawain mamaya. (correct)
   Gagawin natin iyong *kanyakanyang* gawain mamaya. (incorrect)
   kanya-kanya is a full reduplication

4. Paulit-ulit ka nang tanong! (correct)
   Paulet-ulit ka nang tanong! (incorrect)
   ulit-ulit is a full reduplication

5. Bumili kami ng **haluhalo** kaninang umaga. (correct)
   Bumili kami ng *halo-halo* kaninang umaga. (incorrect)
   haluhalo is an ablaut reduplication

Example 8. *Here are some examples of literature and webpages where the wrong reduplicant is used:*

1. **TODO**

1.2.5 Raw v. Daw (Enclitic Particles); Rin v. Din

Let $p = $ " " the blank symbol or the symbol containing space. And, we have the sentence structure $S = \alpha + p + EP$. The *enclitic particles* ($\mathbb{EP}$) is the set $\mathbb{EP} = \{$"raw", "daw"$\}$ and α is any noun, adjective, verb, or adverb. The proper usage of the enclitic particles are given by:

$$\left(\alpha \xrightarrow{\text{regex}} \left[\text{a-z}\tilde{\text{n}}(\text{ng})\text{A-Z}\tilde{\text{N}}(\text{Ng}) \right] + [\text{aeiou}] \equiv (\mathbb{V} \cup \mathbb{C}) + \mathbb{V}_{\text{lower}} \right) \longrightarrow EP = \text{"raw"}$$

Otherwise,

$$EP = \text{"daw"}$$

In other words, if the preceeding word to the enclitic particle ends in vowel, then the enclitic particle is "raw". If it is a consonant, then it is "daw".

This is the same idea with the adverbs "rin" and "din". If the sentence structure $S = \omega + p + (\text{"rin"}|\text{"din"})$ and ω ends in a vowel then the adverb used is "rin". Otherwise, the adverb is "din".

Here are some examples of literature and webpages where the enclitic particles, and the adverbs "raw" and "daw" are misused:

1. **TODO**

1.2.6 Ng v. Nang

There is confusion with the proper use of "ng" and "nang".

There are five way to use "nang": Let $p = $ " " the blank symbol or the symbol containing space.

1. The string "nang" is used before an adverb when the structure of any sentence S is given by:

$$(\text{S}_1 = \text{v} + \text{p} + \text{pron*} + \text{p} + \text{"nang"} + \text{p} + \text{adv})$$

Where v is a verb, adv is an adverb, and pron is a pronoun.

Ex. 1: Kumain nang mabagal.
Ex. 2: Kumain siya nang mabagal.

2. "Nang" can be used to replace the conjunction "noon" + ligature "-ng".

$$(S_2 = \text{"noong"} + v \mid adv)$$

Ex. 1: (Noong | Nang) pumunta siya sa mall. . .
Ex. 2: (Noong | Nang) maaraw na araw. . .

3. "Nang" can be used to replace a combined Enclictic Particle "na" + ligature "-ng".

$$(S_3 = v + p + \text{"na"} + p + \text{"-ng"})$$

Ex. 1: Ang mamatay na ng dahil sa iyo.
       Ang mamatay nang dahil sa iyo.

4. "Nang" can be used to replace "upang" or "para".

$$(S_4 = \text{"nang"} + p + v)$$

Ex. 1 (from Ortograpiyang Pambansa):
. . . kailangang bitayin si Rizal (para | nang) matakot ang mga Filipino.
Ex. 2 (from Ortograpiyang Pambansa):
Dinala si Pedro sa ospital (para | nang) magamot.

5. "Nang" can be used in between repeated verbs.

$$(S_5 = v_1 + p + \text{"nang"} + p + v_1)$$

Ex. 1: Kain nang kain...

Ortograpiyang Pambansa highlights that "nang" should not be mistaken as the contraction of the words "na" + "ang" and that the correct contraction for them is "na'ng".

The string "ng" is used to denote possession similar to the word "of' in English when the structure of any sentence S is given by:

$$(S_1 = n + p + "ng" + p + n)$$

Where n is a noun.

Ex. 1: Bag ng bata
Ex. 2: Pinto ng bahay
Ex. 3: Hawakan ng pinto

"Ng" can be used to mark possession of actions.

$$(S_2 = v+ + p + "ng" + p + n)$$

Ex. 1: Binalat at kinain ng bata ang saging.
Ex. 2: Tinapon ng basurero...
Ex. 3: Inalis ng lola...

1.2.7   *Gitling* Usage