Quantitative component interaction: data from Tagalog nasal substitution[*]

Kie Zuraw

X.1 Introduction: quantitative component interaction

For diagnosing interaction between syntactic and phonological components, invariant phonology is ideal. For example, Piggott and Newell (2014) show that Ojibwa vowel-vowel sequences are resolved by deletion in inalienably possessed nouns, but by consonant insertion in alienably possessed nouns. Piggott and Newell argue that this supports a particular syntactic analysis of the two types of possession, and a difference between phonology that applies within a phase and phonology that applies across a phase boundary. The lack of variation in the Ojibwa case makes it clear that a word's phonological treatment depends on its syntactic structure.

Much phonology, however, is not invariant. Suppose there were a language like Ojibwa except that vowel deletion was only the *preferred* option in inalienable possession, and consonant insertion the preferred option in alienable? Is this predicted or not by current models of syntax-phonology interaction? If it is attested, should the optionality be in the syntax or the phonology, and how should it work?

This paper aims to make an empirical contribution to this line of research by looking in detail at several prefixing patterns in Tagalog and their phonological behaviour. It won't propose syntactic analyses, nor an overall model of syntax-phonology interaction, but will point out some

[*] Thanks to Heather Goad, Heather Newell, Máire Noonan, Glyne Piggott, Lisa Travis, Michael Wagner, and the SIRG group for their organizing work and feedback.

distinctions that could possibly be accounted for syntactically, as well as challenges that variation poses for such an account.

X.2 Tagalog nasal substitution background

In Tagalog, it is well known that a suffix nasal and a stem-initial obstruent optionally are replaced by a nasal at the same place of articulation as the original obstruent (see (Zuraw 2010) for references and many more details). The table in (1), adapted from (Zuraw 2010), shows one example, for each obstruent, of application and one of non-application. All items are from (English 1986), a dictionary. We can see that prefix *ŋ* combines with *p* or *b* to form *m*, with *t*, *s*, *d* to form *n*, and with *k*, *ʔ*, *g*, to form *ŋ̈* (when there is reduplication, "RED", this occurs in both copies). In the examples where nasal substitution does not occur, the nasal generally assimilates in place to the following consonant.

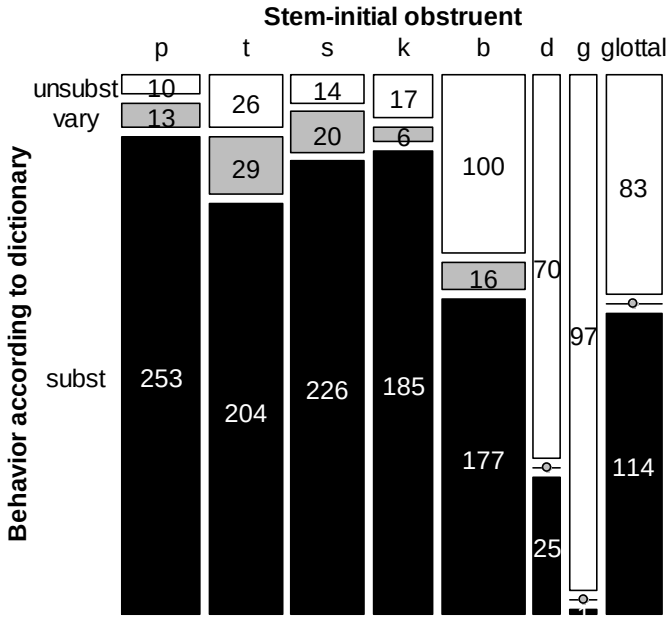
(1)	*stem*		*affixes*	*affixed form showing substitution*	
p	**p**ighatí**ʔ**	‘grief’	pa**ŋ**-RED-X	pa-**m**i-**m**ighatí**ʔ**	‘being in grief’
t	**t**iwála**ʔ**	‘faith’	ka-pa**ŋ**-X-an	kà-pa-**n**iwalá**ʔ**-an	‘traditional belief’
s	**s**úlat	‘writing’	ma**ŋ**-RED-X	mà-**n**u-**n**ulát	‘writer’
k	**k**amkám	‘usurpation’	ma-pa**ŋ**-X	ma-pa-**ŋ**amkám	‘rapacious’
ʔ	**ʔ**isda**ʔ**	‘fish’	ma**ŋ**-X	ma-**ŋ**isdá**ʔ**	‘to fish’
b	mag-**b**igáj	‘to give’	ma**ŋ**-X	ma-**m**igáj	‘to distribute’
d	**d**aláŋin	‘prayer’	pa**ŋ**-X-in	pa-**n**aláŋ̈-in	‘praying’
g	**g**indáj[1]	‘unsteadiness on feet’	pa**ŋ**-RED-X	pa-**ŋ**i-**ŋ**indáj	‘unsteadiness on feet’
p	**p**oʔók	‘district’	pa**ŋ**-X	pa**m-p**oʔók	‘local’
t	**t**abój	‘driving forward’	pa**ŋ**-X	pa**n-t**abój	‘to goad’
s	**s**úlat	‘writing’	pa**ŋ**-X	pa**n-s**úlat	‘writing instrument’
k	**k**úlam	‘sorcery’	ma**ŋ**-RED-X	ma**ŋ-k**u-**k**úlam	‘witch’
ʔ	**ʔ**ulól	‘silly’	ma**ŋ**-X	ma**ŋ-ʔ**ulól	‘to fool someone’
b	**b**igkás	‘pronouncing’	ma**ŋ**-RED-X	ma**m-b**i-**b**igkás	‘reciter’
d	**d**iníg	‘audible’	pa**ŋ**-X	pa**n-d**iníg	‘sense of hearing’
g	**g**áwaj	‘witchcraft’	ma**ŋ**-RED-X	ma**ŋ-g**a-**g**áwaj	‘witch’

We can also see that nasal substitution can be triggered by a variety of affixation patterns, with the rightmost prefix most commonly being *maŋ-*/*naŋ-* or *paŋ-*. Several affixation patterns

[1] Neither the bare root *gindaj* nor any of its derivatives appear in the corpus mentioned below. This is the only instance of substitution of *g* found in English’s dictionary.

are discussed in detail below. The mosaic plot in (2) combines all items to show the rate at which native words in (English 1986)'s dictionary undergo nasal substitution over all, broken down by stem-initial obstruent. For example, 253 words whose stem begins with *p* undergo nasal substitution, 10 do not, and 13 are listed in the dictionary as having both pronunciations. The width of each column is proportional to the number of items that make it up; black tiles represent words that undergo nasal substitution, white tiles word that don't, and grey tiles words that vary. It's evident that the voiceless obstruents undergo substitution more frequently than the voiced, and within the voiced obstruents, there is a place-of-articulation difference:

(2) Dictionary nasal-substitution rates, all prefixation types combined



The above plot reflects only native words. In order to get larger samples for the affixation patterns below, loans from Spanish are also included in counts.

X.3 Morphological differences

Schachter and Otanes (1972) describe nasal substitution generally as applying to *p*, *t*, *s*, *k*; applying optionally to *b*, and not applying to *d*, *g*, and *ʔ*, with exceptions as discussed below, but dictionary data suggest a more variable picture.

This section looks at the most common morphological patterns that can cause nasal substitution. We will see that affixation patterns, even homophonous ones, can differ in their propensity to cause nasal substitution.

X.3.1 Distinctions among *maŋ-X* verbs

The *maŋ-X* pattern forms verbs with a variety of semantic relationships to the root *X*. Schachter and Otanes list five main categories of *maŋ-X* verb:

(3) adversive: "deliberately harmful or destructive activity" (p. 343)
distributive pluractional: "activity directed toward multiple objects" (p. 343)
intensive/iterative pluractional: "intensive or repeated activity directed toward a single object" (p. 343)
incomplete change: "partial or temporary acquisition of [a] quality" (p. 293)
activity: "activities—particularly recreational activities—involving [*X*]" (p. 309)

Examples are given in (4). In the examples chosen here, the root exists as a freestanding word, usually a noun. The basic verb form, *‹um›X* or *mag-X*, is also given; as the glosses show, these verbs generally have a more-transparent semantic relationship to the root.

(4) Five types of *maŋ-X* verb, adapted from Schachter and Otanes, pp. 293, 309, 343, with supplemental data (stress and root glosses) from English (1986).

root		*‹um›X or mag-X verb*		*adversive maŋ-X verb*	
káʔin	‘consumption’	k‹um›áʔin	‘to eat’	ma-ŋáʔin	‘to devour [prey]’
súgat	‘wound’	s‹um›úgat	‘to wound’	ma-núgat	‘to wound (deliberately)’
tagáʔ	‘large incision’	t‹um›agáʔ	‘cut’	ma-nagáʔ	‘to slash (with intent to hurt or destroy)’
walís	‘broom’	mag-walís	‘to sweep’	maŋ-walís	‘to hit with a broom’
root		*‹um›X or mag-X verb*		*distributive pluractional maŋ-X verb*	
kúha	‘act of getting’	k‹um›úha	‘to get’	ma-ŋúha	‘to gather’
pitás	‘picked [flowers, fruits]’	p‹um›itás	‘to pick’	ma-mitás	‘to pick (a number of things)’
tahíʔ	‘stitch’	t‹um›ahíʔ	‘to sew’	ma-nahíʔ	‘to sew (several things, or professionally)’
root		*‹um›X or mag-X verb*		*intensive/iterative pluractional maŋ-X verb*	
tákot	‘fear’	t‹um›ákot	‘to frighten’	ma-nákot	‘to frighten (a number of people)’
dikít	‘glue’	d‹um›ikít	‘to get stuck to’	ma-nikít	‘to get thoroughly stuck to’
sújoʔ	‘ingratiation’	s‹um›újoʔ	‘to curry favour with’	ma-nújoʔ	‘to curry favor with (repeatedly)’
ʔíbig	‘love’	ʔ‹um›íbig	‘to love’	ma-ŋíbig	‘to pay court to’

root		*‹um›X or mag-X verb*		*incomplete change maŋ-X verb*	
kupás	‘colorless’	k‹um›úpas	‘to become colorless’	ma-ŋúpas	‘to fade’
lakí	‘size’	l‹um›akí	‘to grow, enlarge’	man-lakí	‘to enlarge temporarily [eyes]’
pulá	‘red’	p‹um›ulá	‘to become red’	ma-mulá	‘to become reddish, redden temporarily’
putíʔ	‘white’			ma-mutíʔ	‘to blanch’
sakít	‘illness, pain’			ma-nakít	‘to ache’
tigás	‘hardness’			ma-nigás	‘to become stiff’
root		*‹um›X or mag-X verb*		*activity maŋ-X verb*	
baŋkáʔ	‘boat’	mag-baŋkáʔ	‘to go by boat’	ma-maŋkáʔ	‘to go boating’
kabájo	‘horse’			ma-ŋabájo	‘to go horse-riding’
ʔisdáʔ	‘fish’			ma-ŋisdáʔ	‘to go fishing’

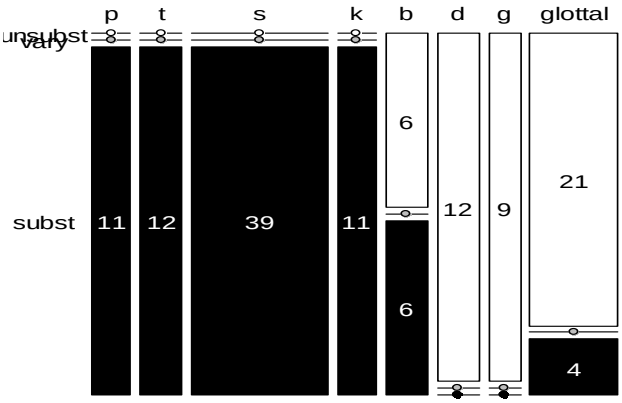
De Guzman (1978) observes that adversive *maŋ-X* verbs show a lower rate of nasal substitution than do other *maŋ-X* verbs. Specifically, she proposes that for adversives, only voiceless obstruents undergo the change, not *b*, *d*, *g*, *ʔ*;[2] whereas for other *maŋ-X* verbs, nasal substitution is obligatory for all obstruents. The dictionary data agree that there is a difference in

[2] Though in some verbs, a [+habitual] or [+professional] feature may override the [+adversive] feature, allowing nasal substitution in *b*, *d*, *g*, or *ʔ* (p. 98).

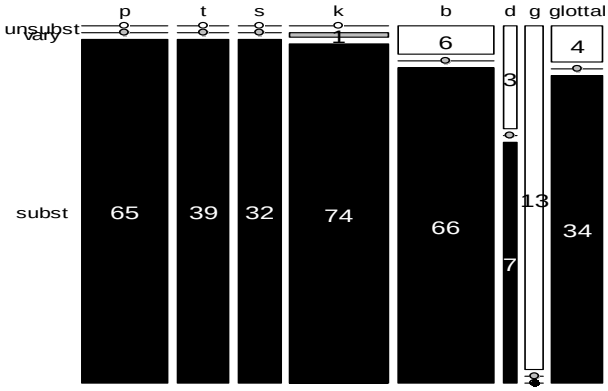
this direction: *b*, *d*, and *ʔ* have lower rates of nasal substitution for adversives than in other *maŋ-X* verbs:

(5) Nasal substitution rates in *maŋ-X* verbs

adversive *maŋ-X* verbs



other *maŋ-X* verbs



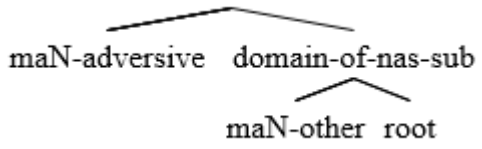
(The counts above reflect my own classification of verbs based on their glosses, and of course these classifications could be incorrect.)

Thus, although *p*, *t*, and *k* are at or near a ceiling of 100% nasal substitution for all the *maŋ-X* verbs, and *g* is at a floor of 0%, in the consonants that show the most variation overall (*b*, *d*, *ʔ*), there is a difference.

X.3.2 Possible mechanisms

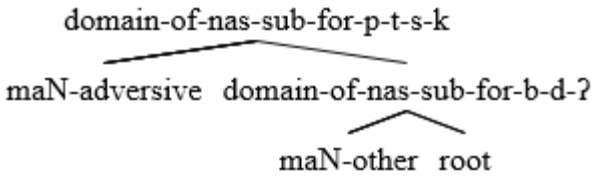
We now how a case of homophonous verbal prefixes with different phonological behavior. If one non-adversive *maŋ-* triggered nasal substitution and adversive *maŋ-* did not, an appealing syntactic explanation would be that adversive *maŋ-* is higher in the tree, outside the domain in which nasal substitution applies:

(6) Hypothetical syntactic scenario: predicts 0% substitution for adversives, 100% for the rest[3]



We could even set up a scenario where nasal substitution is two different rules, applying in two different domains, as shown in (7). Nasal substitution would always apply to *p*, *t*, *k*, because all the prefixes would be within the domain of the rule for those consonants. It would never apply to *g*, because there would be no rule for substitution on *g*. As for *b*, *d*, *ʔ*, the rule for substitution on those consonants would apply only to non-adversives because of its smaller syntactic domain.

(7) Another hypothetical scenario: predicts adversives always substitute *p, t*, *k*, never substitute other consonants; non-adversives always substitute all consonants except *g*.



The problem is that we want nasal substitution to apply to *most* non-adversive words with *b*, *ʔ*, and to a smaller proportion of adversive words with those consonants (leaving aside *d* for

[3] Trees generated by Miles Shang's tree generator, http://mshang.ca/syntree/.

the moment). We could allow some variation in syntactic structure, such that the majority of non-adversive verbs have their *maŋ-* prefixes are low, as shown in (7), but a few are higher, in the *maŋ-* adversive position. This would make no difference to how words with *p*, *t*, *k*, or *g* are pronounced, but would affect *b* and *ʔ*, causing them to be unsubstituted for this minority of non-adversive words. Conversely, while most adversive words would have their prefix high, a minority would have it low; if those words' roots begin with *b* or *ʔ*, this would be detectable as nasal substitution. Turning to *d*, unless we want to take the drastic step of allowing a root's initial consonant to affect the syntactic position of its prefix, we need to ensure that the lower-domain *b-d-ʔ*-substitution rule applies to *d* at a lower rate—so much lower that even the minority of adversive verbs whose prefix is low still fail to undergo nasal substitution.

This leads us to a model where the probability that a word undergoes nasal substitution is predicted by its morphology and its phonology: each type of verb has a certain probability of having its prefix high or low, and if the prefix is low, it has a certain probability of undergoing nasal substitution for each of *b*, *d*, and *ʔ*. The results of this reasoning are shown in (8). Probabilities for the prefix's being low, and for nasal substitution's applying if it is, were fitted using the *optim()* function in R (R Core Team 2014), ignoring the one word in this group that the dictionary lists as variable.

(8) Deriving nasal substitution probabilities

	prob. of prefix low	if prefix low, prob. of nas. sub.	overall predicted prob. of nas. sub.	actual prob. of nas. sub.
p, *t*, *k*	*irrelevant*	*irrelevant*	100%	100%
b, adversive	23%	98%	23%*98% = 23%	50%
b, non-adversive	94%	98%	94%*98% = 92%	92%
d, adversive	23%	66%	23%*66% = 15%	0%
d, non-adversive	94%	66%	94%*66% = 62%	70%
ʔ, adversive	23%	95%	23%*95% = 22%	16%
ʔ, non-adversive	94%	95%	94%*95% = 89%	89%
g	*irrelevant*	*irrelevant*	0%	0%

Given the relatively low numbers overall, the probabilities derived seem like a reasonable match to the observed probabilities.

X.3.3 What kind of variation?

Before moving on to the other constructions, some clarifications are in order as to the nature of the variation reported here. The quantitative data reported are all from a dictionary that was compiled by an English speaker (Leo English) and a Tagalog speaker (Teresita Castillo). Exactly how the dictionary makers chose which pronunciation to give is unknown, but it was probably based on their general experience and perhaps on Castillo's own judgements, as well as on entries in previous dictionaries.

Very few items are listed by the dictionary as having variable pronunciation—these are the tiles in grey in the mosaic plots. For the words that are listed as variable, it is unknown whether a typical speaker would produce both forms interchangeably, or use them in different registers; or whether each speaker would use just one variant. Because there are so few of these cases, knowing this information would not have much effect on the data, though.

The great majority of words are listed as either substituted (black in the mosaic plots) or unsubstituted (white). That is, the rates of substitution are over types, not over tokens. It is unlikely that we are seeing here a mixture of invariant systems from different speakers. For example, suppose that some speakers never nasal-substitute *b*, *d*, *ʔ* in adversives (interpretable as, their adversive *maŋ-* is always high in the syntactic structure) and others always do (all of their *maŋ*-s are low in the structure), so that the apparent variation is illusory. In order to derive the observed dictionary data, some entries' pronunciations would be based on one type of speaker, and others on the other type of speaker. This seems unlikely if one speaker checked all the entries: she would have had to accept many pronunciations counter to her own.

Schachter and Otanes (1972) make no claims about differences in nasal substitution among *maŋ-X* verbs, but as we will see below, they do make claims of different rates for *paŋ-X* adjectives and verbs (see sections X.3.6, and X.3.8 below). Schachter, an English speaker, and Otanes, a Tagalog speaker, state that their grammar is based on "the Manila dialect, as spoken by recent college graduates" (p. 1). This statement is made in the context of describing pronunciation, but probably applies to the rest of the grammar too, and suggests that the patterns they describe represent a coherent lect.

X.3.4 paŋ-RED-X gerunds

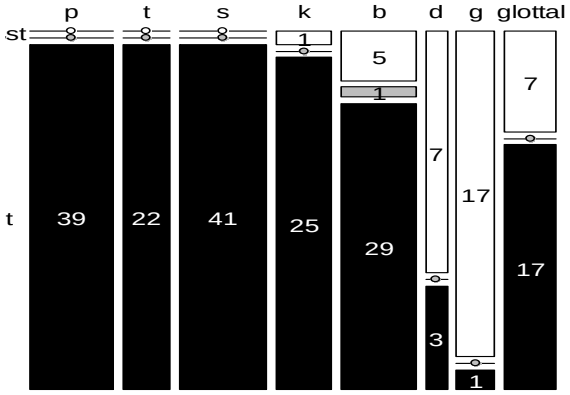
The *paŋ-RED-X* pattern forms gerunds meaning ‘act or fact of doing X’, mainly of verbs that themselves are of the form *maŋ-X* (see (Schachter & Otanes 1972) p. 160):

(9) *paŋ-RED-X* gerund examples

unaffixed		*maŋ-X verb*		*paŋ-RED-X gerund*	
siʔíl	‘oppressed’	ma-niʔíl	‘to oppress’	pa-ni-niʔíl	‘oppressing’
tuksó	‘temptation’	ma-nuksó	‘to tempt’	pa-nu-nuksó	‘tempting’
bulaklák	‘flower’	ma-mulaklák	‘to blossom’	pa-mu-mulaklák	‘blossoming’

Words with this morphology show a high rate of nasal substitution generally, with low rates only for the substitution-resistant consonants *d* and *g*:

(10) Nasal substitution rates in *paŋ-RED-X* gerunds



In the schematic syntactic structure conjectured above, this *paŋ-* would be low, like non-adversive *paŋ-*.

X.3.5 maŋ-RED-X nominals

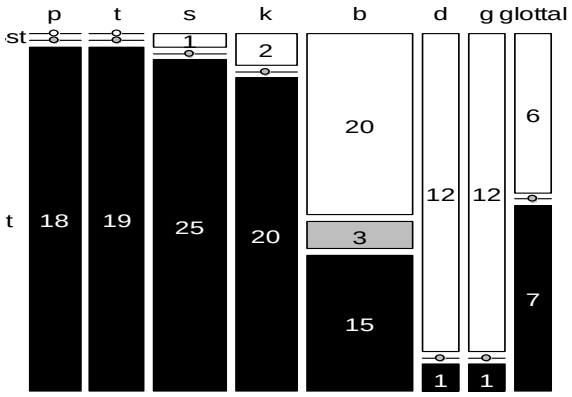
Schachter and Otanes give the meaning of *maŋ-RED-X* as "a person associated—in most cases, professionally—with what the base designates" (p. 103). Examples are shown in (11); because the unaffixed root is usually a noun, a basic verb form is shown also.

(11) *maŋ-RED-X* nominal examples

unaffixed		*verb form*		*maŋ-RED-X nominal*	
kutjáʔ	'mockery'	k‹um›utjáʔ	'to mock'	ma-ŋu-ŋutjáʔ	'cynic'
bigkás	'pronunciation, diction'	b‹um›igkás	'to pronounce, utter, declaim, recite'	mam-bi-bigkás	'declaimer, reciter'
gáʔod	'oar'	g‹um›áʔod	'to row'	maŋ-ga-gáʔod	'oarsman'

Schachter and Otanes describe nasal substitution in this construction as applying to *p*, *t*, *s*, *k*, varying with *b*, *ʔ*, and not applying to *d*, *g*. The dictionary data match this closely. The *maŋ-RED-X* pattern also has a high rate of nasal substitution, though somewhat lower than *paŋ-RED-X*, as seen in the substitution rates for *b* and *ʔ*, which are more sensitive to morphological differences than the voiceless consonants are:

(12) Nasal substitution rates in *maŋ-RED-X* nominals



This *maŋ-* also seems to be fairly low then, although not as consistently as non-adversive *maŋ-*, and note the small number of works with *s* and *k* that resist nasal substitution, suggesting the possibility of an even higher position. To the extent that these small differences in rate of nasal substitution between prefixes are real, they begin to cause discomfort for the syntactic schema in (7)—where in the grammar would we capture different affixes' tendencies to stray from their canonical position?

X.3.6 paŋ-X adjectives, reservational vs. instrumental

Schachter and Otanes divide adjectives formed by *paŋ-X* into 'reservational' and 'instrumental' (pp. 218-221). Reservational adjectives have the meaning of "reserved or intended for use [or wearing] on/in, etc." (p. 218):

(13) Examples of *paŋ-X* reservational adjectives from Schachter and Otanes

root		*paŋ-X reservational adjective*	
kamáj	‘hand’	paŋ-kamáj	‘for the hand (e.g., watch)’
kapé	‘coffee’	paŋ-kapé	‘for coffee (e.g., sugar)’
kúmot	‘blanket’	paŋ-kúmot	‘for making a blanket (e.g., wool)’
ʔopisína	‘office’	paŋ-ʔopisína	‘for the office (clothing, furniture)’
simbá	‘going to church’	pan-simbá	‘for wearing in church’
dalawá	‘two’	pan-dalawá	‘for two (e.g., table)’
ʔikalawáŋ prémjo	‘second prize’	paŋ-ʔikalawáŋ prémjo	‘for use as second prize (e.g., book)’
pijésta	‘fiesta’	mam-pijesta, ma-mijésta	‘for using/wearing at a fiesta’
paléŋke	‘market’	pam-paléŋke, pa-maléŋke	‘for wearing/spending at the market’

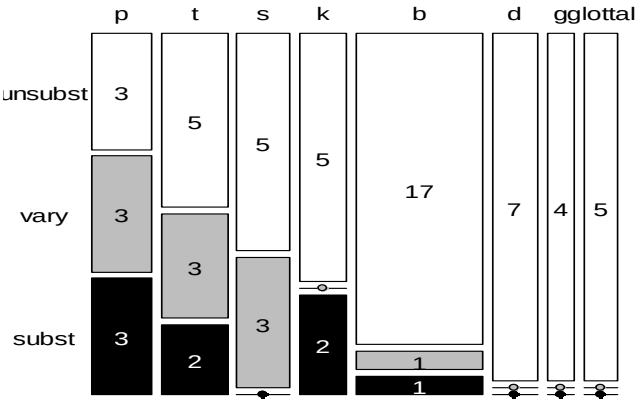
Instrumental adjectives, on the other hand, mean “for use in performing [...] the action designated by the [...] verb formed with the same base” (p. 220):

(14) Examples of *paŋ-X* instrumental adjectives from Schachter and Otanes

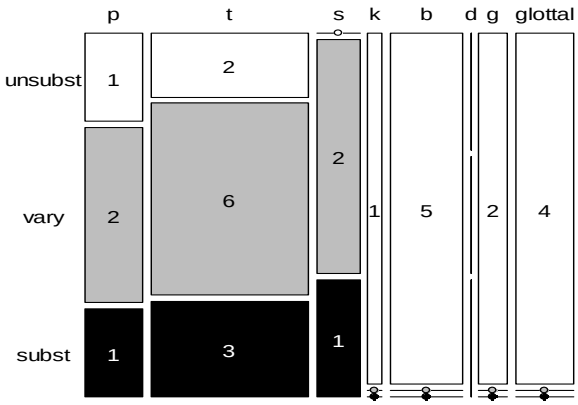
verb form		*paŋ-X instrumental adjective*	
h‹um›ábi	‘to weave’	paŋ-hábi	‘for use in weaving’
mag-lúto?	‘to cook’	pan-lúto?	‘for use in cooking’
mag-púnas	‘to wipe’	pam-púnas, pa-múnas	‘for use in wiping’
mag-nobéna	‘to perform novenas’	paŋ-nobéna	‘for use in performing novenas’
mag-pa-túlog	‘to put to sleep’	pam-pa-túlog	‘for use in putting to sleep’
mag-takíp	‘to cover’	pan-takíp, pa-nakíp	‘for use in covering’

Schachter and Otanes describe nasal substitution as not applying to reservationals, and applying optionally to *p*, *t*, *s*, *k* in instrumentals (see pp. 319-320 for some additional details). Similarly, De Guzman proposes that in instrumental adjectives, nasal substitution is optional for *p*, *t*, *s*, *k*, and forbidden for *b*, *d*, *g*, *ʔ*; in other adjectives, no nasal substitution occurs (except occasional "by analogy" (p. 99)). Unfortunately, dictionary data are too sparse to really compare the two cases, but they confirm the generally lower rate of application than in the prefixes we've seen so far:

(15) *paŋ-X* reservational adjectives

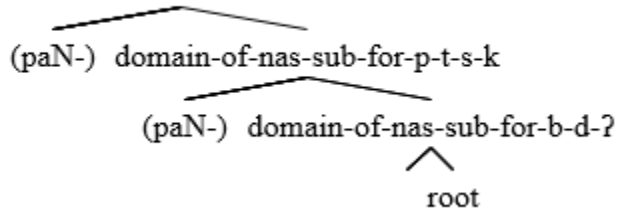


paŋ-X instrumental adjectives



Returning again to the syntactic schema in (7), these *paŋ*-s would have the option of being in a third, even-higher position in the tree, outside the domain of even *p*-*t*-*s*-*k*-substitution:

(16) Schematic syntax for adjectival *paŋ-*



X.3.7 paŋ-X nominalizations

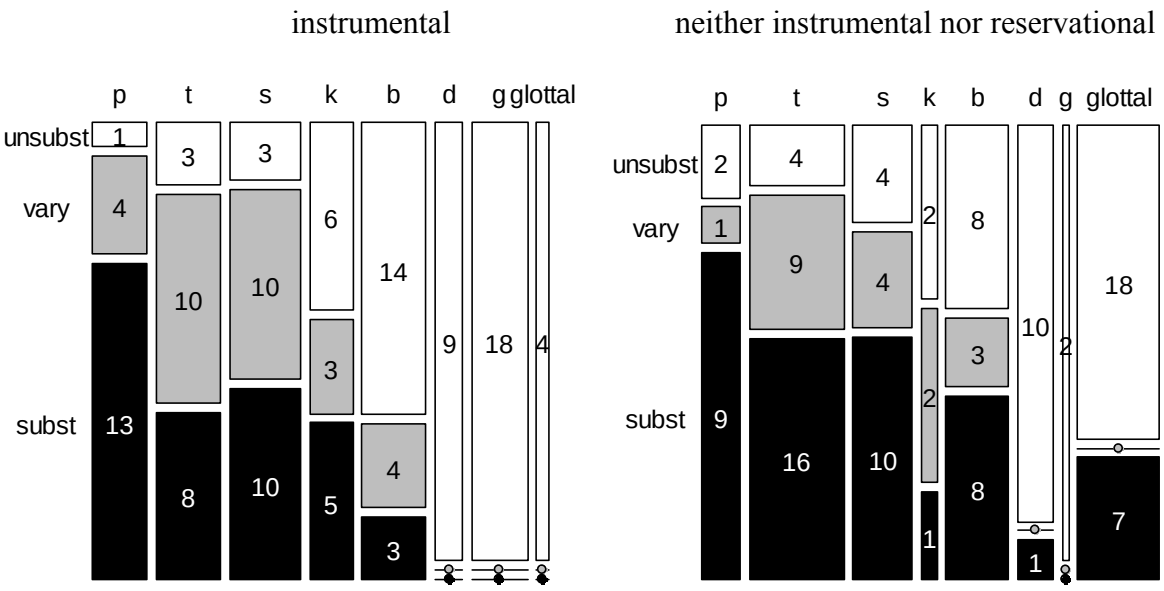
Schachter and Otanes don't discuss *paŋ-X* nominalizations, but De Guzman suggests that they are, at least sometimes, derived from instrumental or reservational adjectives, such as the first three examples in (17); these may be lexicalized ('Christmas present') or semantically transparent. Other types of nominalizations occur too, as illustrated by the rest of the examples in (17).

(17) Examples of *paŋ-X* nouns from De Guzman (p. 89) and (English 1986).

root		*paŋ-X noun*	
paskó	'Christmas'	pa-maskó	'Christmas present'
paskó	'Christmas'	pam-paskó, pa-maskó	'something for use/wear at Christmas time'
taním	'plant'	pan-taním, pa-naním	'something used for planting'
báʔo	'coconut shell'	pa-máʔo	'coconut-shell dipper'
daláŋin	'prayer'	pa-naláŋin	'prayer'
palít	'change'	pa-malít	'substitute'

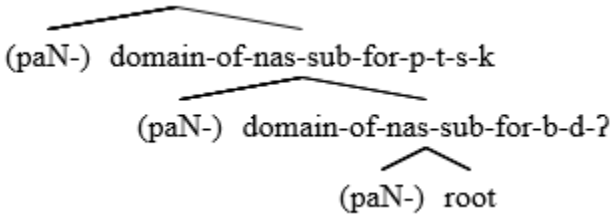
Ideally, we would compare rates of nasal substitution for reservational, instrumental, and other *paŋ-X* nouns, but there were only 15 clearly reservational *paŋ-X* nouns in the dictionary data. Comparing instrumentals to non-reservationals others (though it wasn't always clear how to classify an item), there seems to be little difference, though there is perhaps a trend for the non-instrumental, non-reservational nouns to have a higher rate of nasal substitution, consistent with De Guzman's proposal that nasal-substituted instrumentals may become lexicalized and develop an idiosyncratic meaning:

(18) paŋ-X nouns



It would seem that these *paŋ*-s are able to occur in all three positions in the tree of (16), or at least in the top one and the bottom one:

(19) Schematic syntax for nominal *paŋ-*



X.3.8 paŋ-X verbs

Schachter and Otanes (p. 291) claim that verbs formed with *paŋ-* (together with other affixes, as illustrated below) differ in their rate of nasal substitution. Verbs that have *paŋ-*, possibly with other affixes, but not the prefix *ʔi-*, apply nasal substitution to *p*, *t*, *s*, *k*, show variation for *b*, *ʔ*, and do not apply to *d*, *g*. These include constructions such as *paŋ-X-an*, *paŋ-X-in*, *maka-paŋ-X*, and *maki-paŋ-X*.

In Tagalog, a verb's morphology indicates the theta-role of the clause's externalized DP (see (Pearson 2001)'s analysis of Malagasy). When the prefix sequence *ʔi-paŋ-* is used, Schachter and Otanes state that nasal substitution depends on the theta-role of the externalized DP. Examples are shown in (20). If the external argument is the patient, beneficiary, or causer, then nasal substitution is as for other *paŋ-* verbs (applies to *p*, *t*, *s*, *k*; varies with *b*, *ʔ*). If the external argument is a reservational noun, nasal substitution does not apply. The description on p. 291 groups instrument-external verbs together with reservational-external, but the examples and discussion on p. 320 depict instrument-externals as having variation, like instrumental adjectives: nasal substitution is variable for *p*, *t*, *s*, *k* (other consonants not discussed).

(20) *paŋ-* verbs

n.s. rate	*extern. DP*	*root*		*related verb*		*paŋ- verb*	
high	patient	anák	‘child’	ma-ŋanák	‘to give birth’	ʔi-pa-ŋanák	‘to be given birth to’
	beneficiary	kúha	‘getting’	k‹um›úha	‘to get’	ʔi-pa-ŋúha	‘to be gathered for’
		tahíʔ	‘stitch’	t‹um›ahíʔ	‘to sew’	ʔi-pa-nahíʔ	‘to be sewn for’
	causer	kiníg	‘tremor’	ma-ŋiníg	‘to tremble’	ʔi-ka-pa-ŋinig, ʔi-pa-ŋinig	‘to cause to tremble’
medium	instrument	sukláj	‘comb’			ʔi-pan-sukláj, ʔi-pa-nukláj	‘to be used for combing’
low	reser-vational	báhay	‘house’			ʔi-pam-báhay	‘to be worn in the house’
		sópas	‘soup’			ʔi-pan-sópas	‘to be used in soup’

There aren’t enough of these verbs in the dictionary to count, but the claimed three-way distinction would suggest that the patient-, beneficiary-, and causer-external verbs are in an intermediate position in the tree (below the domain of *p-t-s-k*-substitution but above that of *b-d-ʔ*-substitution), reservational-external verbs are in the highest position, outside the domain of any substitution, and instrument-external verbs vary between the two positions.

X.4 Conclusion

To summarize, rates of nasal substitution in Tagalog are sensitive to some broad distinctions—: *paŋ-RED-X* gerunds, *maŋ-RED-X* nouns, *paŋ-X* adjectives, and *paŋ-X* nouns—but also some fine distinctions, including within homophonous prefixation patterns: whether verbs are adversive or not, whether an adjective is instrumental or reservational, and the theta-role of a verb's externalized DP.

If the syntax drives these phonological differences—for example, if there is a difference in syntactic structure between adversive and non-adversive verbs—then it must do so in a way that is not categorical. It's not that adversive verbs can't undergo nasal substitution; they merely do so at a lower rate than non-adversives. Because all of the prefixation patterns where we had dictionary data do admit nasal substitution, none of them can have a syntactic structure that absolutely prevents the prefix and the stem from undergoing nasal substitution, such as by placing a phase boundary between the two.

A possibility entertained above is that the syntactic structure itself is variable, with prefixes having different propensities to occupy different positions. Some tend to be so low that *b* and *ʔ* usually substitute, and *d* often does too; others tend to occupy an intermediate position where only *p*, *t*, *s*, *k* can substitute, though they can sometimes be lower, so that other consonants substitute too. And yet others are sometimes so high that even *p*, *t*, *s*, *k* fail to substitute. It remains to be seen whether the differing tendencies to be low or high line up with plausible syntactic diagnostics of position in the tree.

A second possibility is that syntactic structure is fixed, and all of the prefixation patterns have a structure that allows nasal substitution, but some of the structures encourage nasal substitution more than others—spelling out how this would work is a challenge.

A rather different possibility is that syntactic structure and lexicalized pronunciations compete. As De Guzman (1978) discusses, some words may become lexicalized as undergoing nasal substitution (and perhaps some as not undergoing), and the lexicalized pronunciation would have to be able to override syntactic structure. For example, following De Guzman, perhaps the structure of *paŋ-* nouns allows *p-t-s-k-* and forbid *b-d-ʔ-*substitution, but this is overridden in a few lexicalized cases. In that case, the challenge shifts to explaining why some prefixation patterns are more likely to generate lexicalized, nasal-substituted words.

This paper leaves all these questions open, and invites syntacticians to investigate these cases more closely.

References

De Guzman, Videa P. (1978). 'A Case for Nonphonological Constraints on Nasal Substitution.' *Oceanic Linguistics* 17(2). 87–106.

English, Leo (1986). *Tagalog-English dictionary*. Manila: Congregation of the Most Holy Redeemer; distributed by Philippine National Book Store.

Pearson, Matthew (2001). 'The clause structure of Malagasy: a minimalist approach.' PhD dissertation, UCLA.

Piggott, Glyne & Heather Newell (2014). 'Interactions at the syntax-phonology interface: evidence from Ojibwe.' *Lingua* 150. 332–362.

R Core Team (2014). *R: A language and environment for statistical computing.* Vienna: R Foundation for Statistical Computing. www.R-project.org.

Schachter, Paul & Fe T Otanes (1972). *Tagalog Reference Grammar.* Berkeley, CA: University of California Press.

Zuraw, Kie (2010). ‘A model of lexical variation and the grammar with application to Tagalog nasal substitution.’ *Natural Language and Linguistic Theory* 28(2). 417–472.