Some data on nasal substitution in several languages

Kie Zuraw, UCLA; kie@ucla.edu

McSIRG, 17 July 2009

0. Introduction

(1) Tobin Skinner's talk last week

Discussed Timugon Murut data in which a prefix may or may not trigger nasal substitution [Data orig. from Prentice 1971].

		(a) no substutition	(b) yes subst.	
/maN+buli/	→	[ma**m-b**uli]	[ma**m**uli]	'*Topic/Subject* will keep'
/maN+tutu/	→	[ma**n-t**utu]	[ma**n**utu]	'*Topic/Subject* will pound

Interpretation (see Gref 2008): In (a), the vP consisting of the root (V) is spelled out before the prefix attaches, disfavoring subsequent loss of the stem-initial consonant's obstruent-hood. In (b), the root moves out of vP and raises to *maN-* before being spelled out, so destructive phonological operations are possible.

(2) Today: some additional data

I've been looking at *phonological* conditions on nasal substitution in this language family, so I'd like to share some data (original and otherwise) and see what you think about how it fits in to the syntactic picture.

This handout is mostly cut and pasted from a submitted article, "A model of lexical variation and the grammar with application to Tagalog nasal substitution", where you can see details of the phonological analysis, the statistics, etc.

1. Nasal substitution in Tagalog

(3) The basic alternation (data from English 1986)
Triggered fairly productively by prefixes *paN-*, *maN-*, *naN-*.

Here they are with sonorants, which can't undergo nasal substitution:

	stem		*affixes*	*affixed form*	
h	**h**ukbó	'army'	paŋ-	paŋ-**h**ukbó	'military'
m	**m**arká	'mark'	paŋ-	paŋ-**m**arká	'marker'
n	**n**egósjo	'business'	paŋ-	paŋ-**n**egósjo	'for business'
ŋ	**ŋ**álit	'grinding of teeth'	paŋ-RED-	paŋ-**ŋ**a-**ŋ**álit	'grinding of teeth'
w	**w**isik-án	'to sprinkle on'	paŋ-	paŋ-**w**isík	'sprinkler'
j	**j**amót	'annoyance'	maŋ-	maŋ-**j**amót	'to annoy'
l	**l**abás	'exterior'	paŋ-	pa**n**-**l**abás	'external'
r	**r**ehjón	'region'	paŋ-	pa**n**-**r**ehjón	'regional'

Here they are with obstruents:

	stem		*affixes*	*affixed form*	
p	**p**o?ók	'district'	paŋ-	pa**m**-**p**o?ók	'local'
	pighatí?	'grief'	paŋ-RED-	pa-**m**i-**m**ighatí?	'being in grief'
t	**t**abój	'driving forward'	paŋ-	pa**n**-**t**abój	'to goad'
	tiwála?	'faith'	ka-paŋ- -an	kà-pa-**n**iwála?-an	'traditional belief'
s	**s**úlat	'writing'	paŋ-	pa**n**-**s**úlat	'writing instrument'
	súlat	'writing'	maŋ-RED-	mà-**n**u-**n**ulát	'writer'
k	**k**úlam	'sorcery'	maŋ-RED-	ma**ŋ**-**k**u-**k**úlam	'witch'
	kamkám	'usurpation'	ma-paŋ-	ma-pa-**ŋ**amkám	'rapacious'
ʔ	**ʔ**ulól	'silly'	maŋ-	ma**ŋ**-**ʔ**ulól	'to fool someone'
	ʔisdá?	'fish'	maŋ-	ma-**ŋ**isdá?	'to fish'
b	**b**igkás	'pronouncing'	maŋ-RED-	ma**m**-**b**i-**b**igkás	'reciter'
	mag-**b**igáj	'to give'	maŋ-	ma-**m**igáj	'to distribute'
d	**d**iníg	'audible'	paŋ-	pa**n**-**d**iníg	'sense of hearing'
	daláŋin	'prayer'	i-paŋ- -in	?i-pa-**n**aláŋinin	'to pray'
g	**g**áwaj	'witchcraft'	maŋ-RED-	ma**ŋ**-**g**a-**g**áwaj	'witch'
	gindáj	'unsteadiness on feet'	paŋ-RED-	pa-**ŋ**i-**ŋ**indáj	'unsteadiness on feet'

(4) Other affixes

- Impressionistically unproductive, can trigger: *taŋ-*, *tuŋ- siŋ-*, *hiŋ-*, *kaŋ-*, and *kuŋ-*:

taŋ-	*(no substituting examples found)*			
	bílaŋ	'number'	ta**m-b**ílaŋ	'digit'
tuŋ-	*(no substituting examples found)*			
	balík	'upside-down'	tu**m-b**alík	'return'
siŋ-	púno?	'leader'	si-**m**úno?	'grammatical subject'
	tábi?	'move aside!'	pa-si**n-t**ábi?	'respect; asking pardon'
hiŋ-	kúto	'louse'	hi-**ŋ**utú-han	'to pick out lice'
	túlot	'permission'	pa-hi**n-t**úlot	'permission'
kaŋ-	patáj	'corpse'	ka-**m**àtáj-an	'death'
	gatá?	'coconut milk'	kà-ka**ŋ-g**atá?	'first extraction of coconut milk; essence'
kuŋ-	*(no substituting examples found)*			
	babá?	'descent'	mag-pa-ku**m-b**abá?	'humble'

- Productive but never triggers: *mag-kaŋ-RED*, for verbs of accidental result
  dapá? 'face down' *mag-kan-da-rápa?* 'to fall on one's face'

- Never trigger, compound-like: *waláŋ-* 'not exist', *(ʔi)sáŋ-* 'one', *(ka)síŋ-* 'as X as', *pagigíŋ-* 'becoming', *magíŋ-* 'become'

bájad	'payment'	waláŋ-bájad	'free'
dáliʔ	'finger-width'	san-dáliʔ	'one finger width'
ʔitím	'black'	kasíŋ-ʔitím	'as black as'
táʔo	'person'	pagigíŋ-táʔo	'becoming a person'
ʔabogádo	'lawyer'	magíŋ-ʔabogádo	'to become a lawyer'

(Q: Why compound-like? A: At least two syllables long in their full forms, can bear their own stress, produce semantically transparent words; *waláŋ-* and *(ʔi)sáŋ-* are presumably derived from freestanding *waláʔ* 'does not have/exist' and *ʔisáʔ* 'one', plus "linker" *-ŋ-*. Forms with *magíŋ-* are usually spelled as two separate words.)

(5) Reduplication

/paŋ-RED-pighatíʔ/ → *pa-**m**i-**m**ighatíʔ* , not **pa-**m**i-**p**ighatíʔ*

Various explanations for this "double application"
- nasal substitution precedes reduplication, in a counterbleeding order (Bloomfield 1917; Carrier 1979; Raimy 2000)
- both reduplicant and base select a nasal-substituted allomorph because of the morphological context (Marantz 1982; Inkelas and Zoll 2000, 2005)
- a special relationship between base and reduplicant forces nasal substitution to apply to both (Wilbur 1973; McCarthy and Prince 1995)

Reduplicated forms suggest that when nasal substitution applies (and only then), the nasal belongs to a stem that serves as the base of reduplication:

*pa-**m**ighatíʔ* > *pa-**m**i-**m**ighatíʔ* not *pa**m**-ighatíʔ* > **pa**m**-i-(ʔ)ighatíʔ*
*ma**m**-bigkás* > *ma**m**-bi-bigkás* not *ma-**m**bigkás* > **ma-**m**bi-**m**bigkás.*

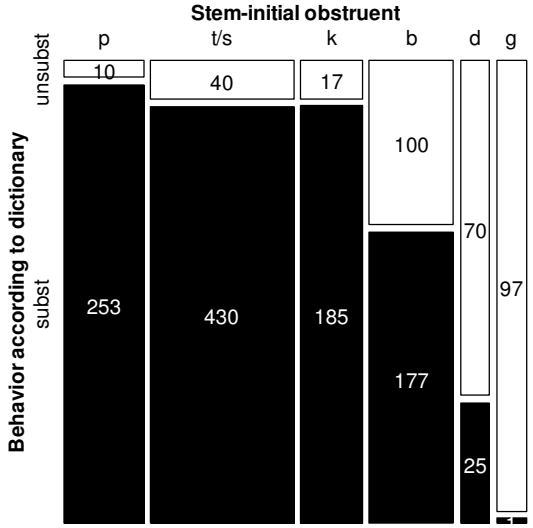
(6) Lexical statistics from dictionary: overall trends

Dictionary used: English's (1986), non-loans only, with an obstruent-initial stem and a potentially nasal-substituting prefix

Phonological trends
- Substitution is more likely if stem-initial consonant is voiceless than if voiced.
- Among the voiced consonants, substitution is most likely with *b* and least likely with *g*.[1]

[1] Previous accounts of the lexical distribution of nasal substitution have stated, mostly in passing, that *g* never substitutes (Bloomfield 1917; Schachter & Otanes 1972); that *d* and *g* rarely substitute (Blake 1925); that voiceless consonants substitute more than voiced ones (De Guzman 1978); and that morphology matters (Schachter & Otanes 1972; De Guzman 1978, who gives detailed claims about various morphological constructions).

Figure 1: Rates of nasal substitution for entire lexicon—dictionary data[2]



(7) Broken down by construction
Dictionary data for six most common affix patterns, accounting for 1,670 of the 1,736 words in the dictionary.

Breakdown by affix based mostly on De Guzman (1978):

- adversative verbs (hostile to the patient): *bató* 'stones' *ma-mató* ~ *mam-bató* 'to throw stones at'
- other verbs: inchoative (*paját* 'thin', *ma-maját* 'to become thin'), stative (*butiktîk* 'teeming with', *ma-mutiktîk* 'to teem with'), professional (*gamót* 'medicine', *maŋ-gamót* 'to practice medicine'), habitual (*sigarîljo* 'cigarette', *ma-nigarîljo* 'to be a smoker'), distributive (*k-um-úha* 'get', *ma-ŋúha* 'to gather things'), repetitive verbs (*bintána* 'window', *ma-mintána* 'to keep looking out a window'), and others.
- instrumental adjectives: *títik* 'writing', *pa-nítik* 'used for writing'
- reservative adjectives: *baŋkéte* 'banquet' *pam-baŋkéte* 'appropriate for a banquet'
- gerunds (*tahîʔ* 'stitch', *pa-na-nahîʔ* 'sewing') and less-transparent nominalizations

[2] Mosaic plot, made using the *mosaic()* function of the *vcd* package (Meyer et al. 2006, 2007) of the statistical computing program R (R Core Development Team 2007).

Figure 2: Rates of substitution for *paŋ-RED-* construction [mainly gerunds (*tahí?* 'stitch', *pa-na-nahí?* 'sewing'), but also some less transparent nominalizations]

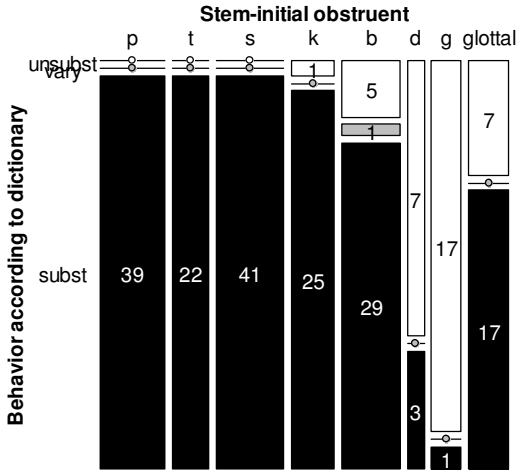


Figure 3: Rates of substitution for *maŋ-RED-* construction [professional or habitual nouns (*bátas* 'law', *mam-ba-batás* 'legislator')]

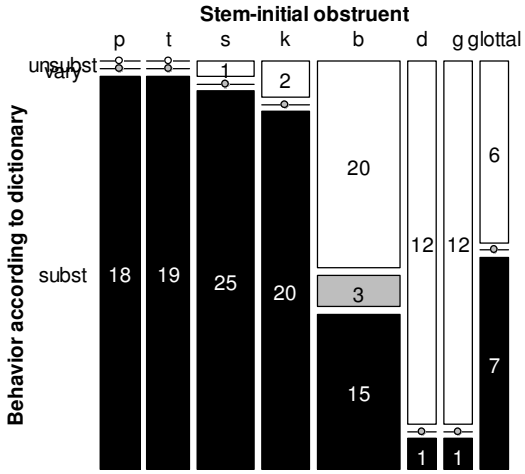


Figure 4: Rates of substitution for *maŋ-* (adversative) construction

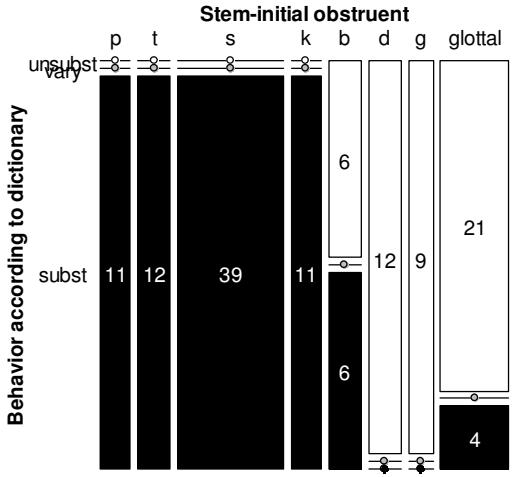


Figure 5: Rates of substitution for *maŋ-* (other verbs) construction

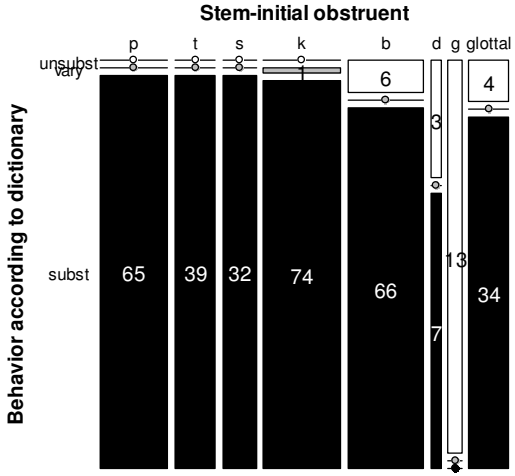


Figure 6: Rates of substitution for *paŋ-* (noun) construction [instrumentals, gerunds, and other nominalizations (*gúgol* 'expense', *paŋ-gúgol* 'spending money')]

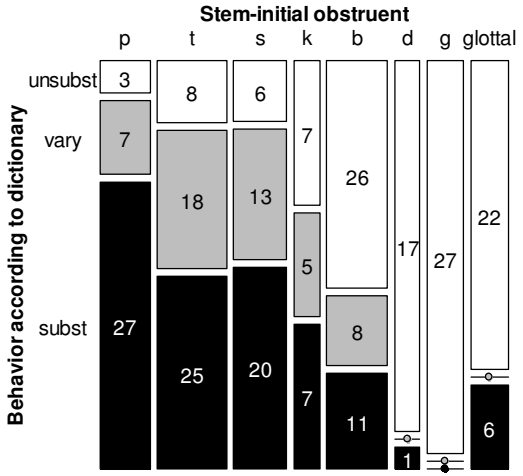
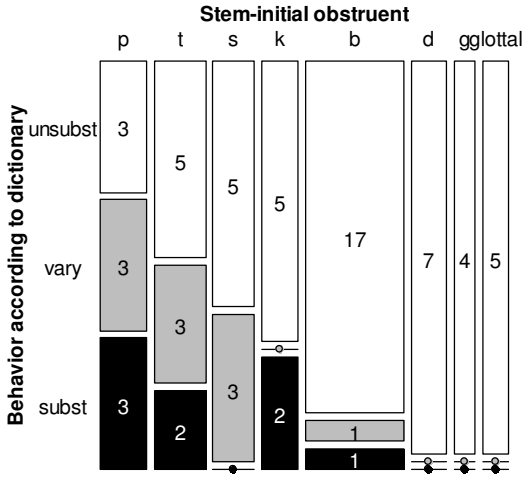


Figure 7: Rates of substitution for *paŋ-* (reservative) construction

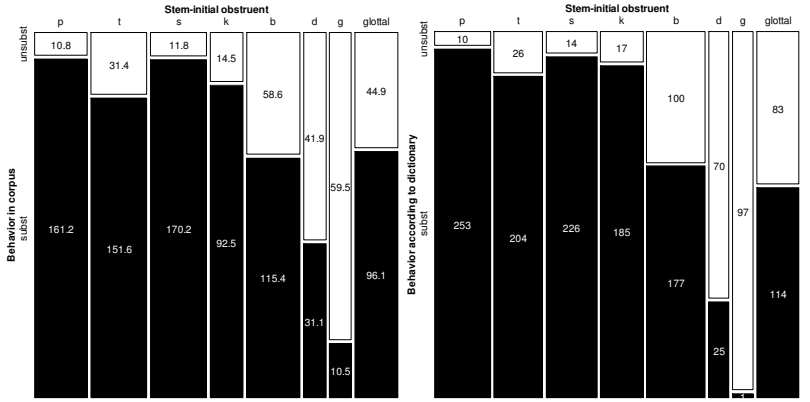


2. Reasons to take these patterns seriously

(8) Confirmation of dictionary data by written corpus

- Web derived, using software written by Ivan Tam and a seed corpus generously supplied by Rosie Jones (derived from Ghani, Jones and Mladenić 2004); see Zuraw 2006 for details.
- Out of 1,715 dictionary words probed in the corpus, 1,107 were attested in at least one variant, for a total of 195,513 tokens.
- Plot on left shows token-weighted type frequencies.

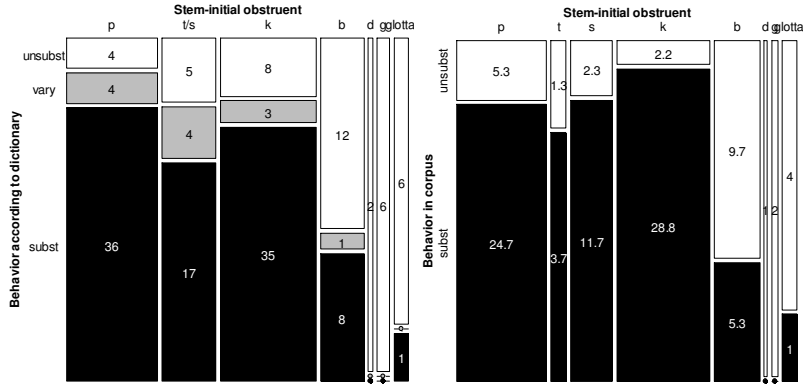
Figure 8: Rates of nasal substitution in corpus vs. dictionary, native words only



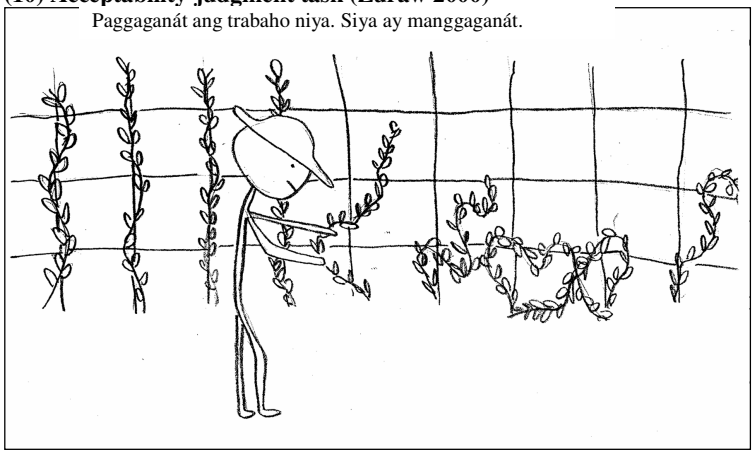
(9) Pattern is extended onto loans

- Tagalog in contact with Spanish from about mid-1500s to early 1900s (and with English since then)
- Loans from Spanish very well integrated, often not recognized as loans.
- A decent number enter into nasal-substituting constructions.
- Voicing effect clearly perpetuated; place effect less clear, but *b* seems to have higher rates than *d/g* (which are rare, at least in these constructions, for some reason).

Figure 9: Substitution rates for Spanish stems, all affixal patterns combined

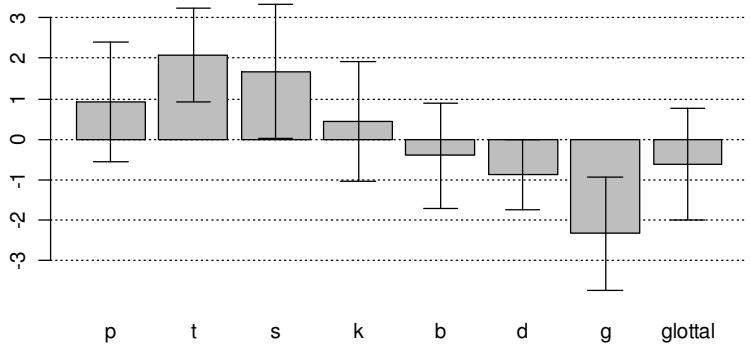


(10) Acceptability-judgment task (Zuraw 2000)



- Participant rates from 1 to 10; sees both substituted and unsubstituted version of each stem.

Figure 10: Acceptability judgments: substituted – non-substituted; error bars indicate 95% confidence interval



- Positive numbers (voiceless): substituted version rated higher
- Negative numbers (voiced): unsubstituted version rated higher
- Except for *p*, consistent with place effect

(11) Binary choice task

- Participants recruited over web
- Each participant sees a different set of items, to avoid item-specific distortions

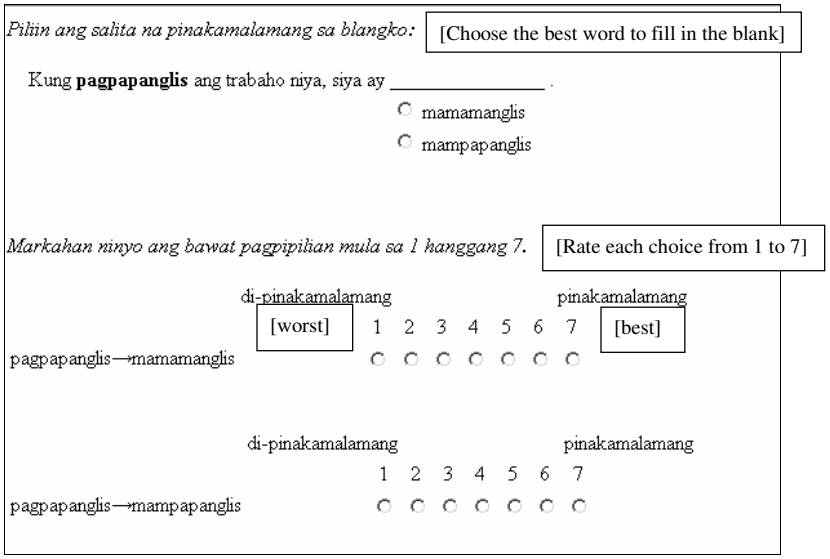
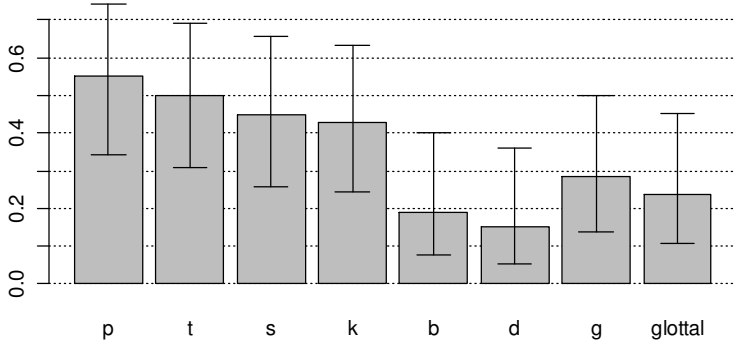


Figure 11: Rates at which subjects selected nasal-substituted option in web survey; error bars indicate 95% confidence intervals.[3]



- Voicing effect robust
- Suggestive place differences, except surprising results for *g*—possible visual confusion between <ng>=[ŋ] and <ngg>=[ŋg].

3. Lexical idiosyncrasy

Several ways in which words that take *paŋ-* and *maŋ-* prefixes can be idiosyncratic, suggesting that the lexical entry must be consulted to determine behavior.

[3] Using the *binconf()* function of the *Hmisc* package of R (Harrell 2008), default Wilson method.

(12) Variation seen above

- Despite the lexical trends, it's not completely predictable which words will undergo substitution (especially among *b*- and *d*-initial stems)

=> At least the words that buck the trend for their segment and morpheme must be listed as exceptions.

(13) Idiosyncrasy is whole-word

Substitution isn't even consistent among derivatives of the same stem:

prefixes	*nas. sub.?*	*(freq. no ; freq. yes)*		
			búhaj	'life'
paŋ-	no	(10 ; 0)	pam-**b**úhaj	'vivifying'
maŋ-	yes	(0 ; 652)	ma-**m**úhaj	'to live'
paŋ-RED-	yes	(1; 1975)	pa-**m**u-**m**úhaj	'manner of living'
			batás	'law'
paŋ-	no	(30 ; 0)	pam-**b**atás	'legal'
paŋ- -an	yes	(1 ; 47)	pa-**m**àtás-an	'legislative'
maŋ-RED-	no	(766 ; 0)	mam-**b**a-**b**atás, mam-**b**a-**b**átas	'legislator'

=> It doesn't suffice to access the stem's lexical entry. We (at least sometimes) need the whole word's entry, or at least need to know which stem *and* which affix are being used.

Walther and Wiese (1999) suggest that each stem and each prefix get a diacritic: if both are [+substitute], substitution occurs; otherwise no.

- But we'd still have to list some exceptional combinations. E.g., if *paŋ-* is [-substitute], the 3 cases where substitution does occur with *paŋ-* must be listed as exceptions; if *paŋ-* is [+substitute], cases like *pam-búhaj* are exceptions.

(14) Semantic idiosyncrasy

Semantic connection is typically apparent, but exact meanings can be unpredictable.

ʔabáŋ	'watcher'	ma-ŋabáŋ	'to wait near people who are eating, hoping to get food'
babáʔe	'woman'	mam-babáʔe	'to have a mistress'
siʔíl	'oppressed by ruler'	ma-niʔíl	'to strangle to death'
ʔibábaw	'surface'	paŋ-ʔibábaw	'veneer'
kíta	'visible'	pà-ŋitáʔ-in, pà-ŋitaʔ-ín	'apparition, omen'
túbig	'water'	ma-nubíg	'to urinate'
balík	'return'	pa-malík	'hand rudder'
gánùʃo	'hook'	maŋ-ga-gánùʃo	'con man'

=> Lexicon must sometimes specify the meaning of a potentially nasal-substituted word

(15) Idiosyncratic stress/length shifts

More often for substituted, but occurs for unsubstituted too.

		maŋ-RED- with stress shift	
tahíʔ	'sewing'	mà-na-náhiʔ	'seamstress'
		maŋ-RED- without stress shift	
puná	'remark'	mà-mu-muná	'critic'
ʔáwit	'song'	maŋ-ʔa-ʔáwit	'singer'
		maŋ- with stress shift	
túbig	'water'	ma-nubíg	'to urinate'
		maŋ- without stress shift	
kíkil	'carpenter's file'	ma-ŋíkil	'to chisel; to ask for money'
		paŋ- with stress shift	
sípit	'claws'	pan-sipít	'(type of) rat-trap'
		paŋ- without stress shift	
túkoj	'mention'	pan-túkoj	'article [grammar]'

=> Need to know what both stem and affix are in order to determine stress.[4]

(16) Three-way distinction

- words that are lexicalized as undergoing nasal substitution *mà-ma-mahála?* 'responsibility' (< *bahála?* 'manager')
- words lexicalized as not undergoing nasal substitution *mam-ba-básа* 'reader' (< *básа* 'reading')
  [*b*-initial stems in the *maŋ-RED-* construction; both are frequent in the corpus and consistent in their behavior there).
- words not yet having a lexicalized behavior, such as these presumably nonce or recent coinages, sampled from the *maŋ-* construction.

spelled form	*frequency*	*presumed English source*
mangkonjugation	1	conjugation
mam-bird-repel	1	bird-repel
mamblog	11	blog
man-takeover	1	takeover
ma-mrospect	1	prospect
mangareer	5	career (<ng> = [ŋ])

[4] Why do certain words (and not others) have idiosyncratic properties? The answers are presumably different from the diachronic point of view and from the point of view of the learner who must replicate the ambient pronunciations and meanings. Hay (2003) discusses a two-way relationship between lexical representation/access and idiosyncrasy: lexical idiosyncrasy can cause learners to treat the affected words more as wholes than as morphologically composed; conversely, if some other factor causes learners to treat certain words as wholes, then those words have a better chance of developing idiosyncrasies over time.

4. Why I thought this was worth talking about here

(17) If variation in substitution vs. not is driven by variation in the syntactic derivation...

- Can words be idiosyncratically marked for which syntactic derivation they take?? (seems strange since this information would be needed before lexical insertion, right?)
- It makes sense that different morphology would be associated with different rates of each derivation type (I think?), but does it make sense that different obstruents would be associated with different rates of each syntactic derivation?

If the variation is just encoded in lexical entries, with some words requiring nasal substitution and others forbidding it...

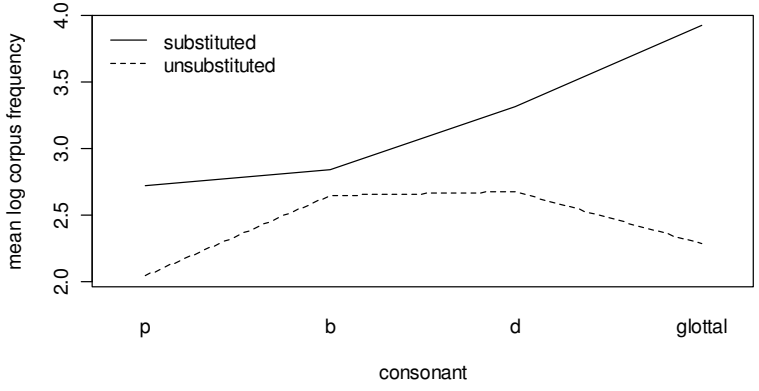
- How do we enforce a lexical nasal-substitution requirement if the spell-out-before-raising option is taken?

5. Nasal substitution and boundary strength

(18) Nasal substitution seems to be negatively correlated with degree to which a word is transparently prefixed

- Meaning: I haven't conducted a systematic study of semantic opacity, we can see some trends in the morpheme-by-morpheme breakdown above.
  - Lowest rates of substitution found in noun-forming *paŋ-* (often transparent) and reservative-adjective-forming *paŋ-* (almost always transparent).
  - Highest rates of substitution found in miscellaneous-verb-forming *maŋ-* construction (semantics very unpredictable) and nominalizing *paŋ-RED-*, (transparency varies).
- Frequency: Seems plausible that higher-frequency words are more likely to be treated by speakers as whole units rather than as prefix-stem combinations (see e.g. Hay 2003).
  - Nasal-substituted words have higher frequency on average. Figure 12 shows the four consonants that had at least 20 words in each of the two categories (substituted and not).

Figure 12: Corpus frequencies for substituted and unsubstituted words



(19) What do you guys think about semantic transparency/frequency vs. derivation type?

Is spell-out-before-raising more consistent with a transparent, compositional treatment?

6. Analysis/model [we can skip this part]

(20) Representations and constraints

I'll spare you the details—and am not that committed to these specific constraints anyway—but here are the key points in case you're curious.

- Prefix is treated as having a floating nasal feature /ma[+nas]/
- There is conflict between a constraint against inserting a segment to support the feature, and one against docking the floating feature to the initial segment of them stem, since it's a different morpheme (*ASSOCIATE$_{\text{hetero-morphemic}}$).

	/p$_1$a$_2$[+nas]$_3$/+/b$_4$i$_5$g$_6$a$_7$j$_8$/	MAX(+nas)	DEP-C	*ASSOCIATE$_{\text{hetero-morphemic}}$
☞ *a*	pa-m$_4$igaj | [+nas]$_3$			*
b	pam$_9$-b$_4$igaj | [+nas]$_3$		*!	
c	pa-b$_4$igaj	*!		

- A constraint MORPHEMECOHESION bans coalescence across a "big" morpheme boundary (shown as # below, though not literally Chomsky & Halle's #). Could also think in terms of derivational levels—or phases?

	/pa[+nas]$_3$/ # /p$_4$ulítika/	MORPHEME COHESION	MAX(+nas)	DEP-C	*ASSOCIATE
a	pa # m$_4$ulítika | [+nas]$_3$	*!			*
☞ *b*	pam$_{12}$ # p$_4$ulítika | [+nas]$_3$			*	
c	pa # p$_4$ulítika		*!		

- Phonetically motivated constraint against non-substitution on voiceless obstruents (Pater's *NÇ$_{\text{ɔ}}$).
- Phonetically motivated constraints against creating stem-initial nasals (through substitution): *[ŋ (or fronter) > > *[n (or fronter) >> *[m

(21) Model of variation

This is the bulk of the paper. Again I'll spare you the details, but here are the key points in case curious.

- Whole words can have lexical entries; high-ranked faithfulness constraints enforce the substitution behavior encoded in them.
- If a word is a new coinage, these faithfulness constraints are irrelevant, and lower, variably-ranked constraints determine the outcome.
- Boersma's Gradual Learning Algorithm (1997, 1998), when trained on the lexicon, is able to learn the rankings of the "subterranean" constraints before the

faithfulness constraints climb to the top and cause (error-driven) learning to cease. [This works even if the constraint set is symmetrical, e.g. including both the constraint for the voicing effect and its opposite.]

7. Cross-linguistic data

- See Newman (1984) for a survey that points out the typological trends below (and has other interesting nuggets), and Blust (2004) for a bigger survey that replicates Newman's findings, with the exception of Kapampangan (see below).
- Depending on the pattern's starting point, nasal substitution has either retreated from less susceptible segments in some daughter languages, spread to more susceptible segments in some daughter languages, or some of each.
- Whatever the case, we'll see that there seems to be great cross-linguistic consistency in what stem-initial consonants are more or less susceptible.
- Suggests that the voicing and place effects shape the diachronic development of a language's lexicon.[5]

(22) Simple factorial typology (ignoring possibility of variation)

		substituted?						
	languages	p	t	k	b	d	g	*sample ranking*
a	Da'a, Wolio, Bugis	–	–	–	–	–	–	*[ŋ, *[n, *[m, *ASSOC >> DEP-C, *NC̥
b	similar to Balantak	+	–	–	–	–	–	*[ŋ, *[n >> *NC̥ >> *ASSOC, *[m >> DEP-C
c	?	+	–	–	+	–	–	*[ŋ, *[n >> *NC̥, DEP-C >> *ASSOC, *[m
d	similar to Yami	+	+	–	–	–	–	*[ŋ >> *NC̥ >> *[n, *[m, * ASSOC >> DEP-C
e	sim. to Toba Batak	+	+	–	+	–	–	*[ŋ >> *NC̥ >> *[n >> DEP-C >> * ASSOC, *[m
f	?	+	+	–	+	+	–	*[ŋ >> *NC̥, DEP-C >> *[n, *[m, * ASSOC
g	Malay/Indonesian & others	+	+	+	–	–	–	*NC̥ >> *[ŋ, *[n, *[m, * ASSOC >> DEP-C
h	Sama-Badjau, Dibabawon Manobo	+	+	+	+	–	–	*NC̥ >> *[ŋ, *[n >> DEP-C >> *[m, * ASSOC
i	Cebuano, Isnag, Sarangani Manobo	+	+	+	+	+	–	*NC̥ >> *[ŋ >> DEP-C >> *[n, *[m, * ASSOC
j	Kalinga	+	+	+	+	+	+	*NC̥ or DEP-C >> *[ŋ, *[n, *[m, * ASSOC

(notation adapted from Newman 1984: 10)

[5] Malay/Indonesian presents one of the few cases where change can be observed in the written record. Currently, Malay/Indonesian has a system in which nasal substitution applies to all the voiceless obstruents and none of the voiced (Lapoliwa 1981; though see Delilkan 2002 for prosodic and morphological complications). But, as Newman (1984) points out, Brakel (1973) claims that substitution can be found on voiced obstruents in 16th and 17th-century Malay manuscripts, with some such words "maintain[ing] themselves as archaic forms till well into the 19th C." (Brakel: 4). It is not clear from Brakel's discussion whether substitution was the norm on (at least some) voiced obstruents in these manuscripts, but we can at least say that the lexicon of Malay has been reshaped over the last few hundred years to reflect a different grammar of nasal substitution.

(23) Some details (language-family designations from Gordon 2005)

- Da'a (Sulawesi, Barr 1995) and Wolio (Sulawesi, Anceaux and Grimes 1995): descendants of nasal-substituting prefixes induce prenasalization, not substitution
- Bugis (Sulawesi; Abas and Grimes 1995): gemination (/maŋ-tunu/ → [mattunu] 'burn s.th., bake s.th.') instead of nasal substitution.
- If we accept Ross's (1988) evidence that "[c]ases of nasal substitution are preserved sporadically in Oceanic languages" (p. 41), then nasal substitution has also died out in the entire Central/Eastern Malayo-Polynesian branch of Malayo-Polynesian (rather than being an innovation confined to Western Malayo-Polynesian).
- Balantak (Sulawesi; Busenitz and Busenitz 1991, Busenitz 1994): nasal substitution applies to *p*-initial stems, unless the next syllable also begins with *p* (Busenitz 1994: 3).
- Yami (Northern Philippine; West 1995) almost exemplifies pattern (d): it distinguishes *p*, *t* from the rest, but the difference is that *p*, *t* are reported to undergo nasal substitution uniformly, and the other stops vary.
- Indonesian/Malay (Sundic; Lapoliwa 1981)
- Sama-Bajau (Sama-Bajaw; Verheijen 1986)
- Dibabawon Manobo (S. Philippine; Forster 1970).
- Cebuano (Meso-Philippine; Wolff 1962): Van Odijk 1959, a description aimed at missionaries, appears to claim that application of nasal substitution is variable in Cebuano, but the passage (p. 44) is difficult to interpret because Odijk appears to be describing the distribution of the prefixes *maŋ-/naŋ-/paŋ-* vs. *mag-/nag-/pag-* rather than the distribution of nasal substitution vs. non-substitution.
- Isnag (N. Philippine; Vanoverbergh 1972): very few *g*-initial stems in Vanoverbergh's (1972) dictionary, and none takes a relevant prefix. Word-initial *g* of other Philippine languages seems to correspond to Isnag orthographic <x>, ([h] in some dialects and [ɣ] in others). When this consonant takes *maŋ-* or *paŋ-*, it behaves as a non-substituted *g*: <xabí> 'night', <ma[ŋ]-gabí> 'to abstain from rice and taro while in mourning' (p. 245), with one exception, <ma[ŋ]-xakkí> 'to have one's skin open piecemeal' (p. 248).
- Limos Kalinga (N. Philippine; Ferreirinho 1993)
- Ginaang Kalinga (N. Philippine; Gieser 1970)
- Sarangani Manobo (S. Philippine; DuBois n.d.)

(24) Cases with variation

- "~" = variation reported
- "+~" = variation reported but with preference for substitution
- "–~" = variation reported but with preference for non-substitution

	p	t	k	b	d	ɖ	g	*sample ranking*
Yami	+	+	~	~		~	~	*NC̥ > *[ŋ ~ DEP-C > *ASSOC >> *[n ~ *[m
Sasak	+	+	+~	~	–		–	*NC̥ > *[ŋ >> *[n >> *[m ~ DEP-C >> *ASSOC
T. Batak	+	+~	–	–~	–		–	*[ŋ >> *NC̥ > *[n >> *[m > DEP-C >> *ASSOC
K. Batak N^2-	+	–~	–~	–~	–		–	*[ŋ ~ *[n > *NC̥ >> *[m > DEP-C >> *ASSOC
K. Batak N^1-	+	+	~	+~	–		–	*NC̥ ~ *[ŋ >> *[n >> DEP-C > *[m >> *ASSOC
K. Batak $N^{3/5}$-	+	+	+	–	–		–	*NC̥ >> *[ŋ ~ *[n ~ *[m ~ *ASSOC >> DEP-C
Palawan	+	+	+	~	–		–	*NC̥ >> *[ŋ ~ *[n >>*[m ~ DEP-C >> *ASSOC
Kapampangan	+	+	+	+	–~		+~	*NC̥ >> *[ŋ ~ *[n ~ DEP-C >> *[m ~ *ASSOC

(25) Details

- Toba Batak (Sundic; Nababan 1981, Percival 1981, Van der Tuuk 1867/1971): reported that *p* always substitutes, *t* (and *s*) usually do, *b* usually doesn't, and *d* and *g* never do.
- Karo Batak (Sundic): Woollams (1996) reports that nasal substitution **applies differently with three different prefixes**
  - N^1- marks active voice
  - N^2- forms intransitive verbs
  - $N^{3/5}$- forms certain adjectives.
- Sasak (Bali-Sasak, Goris 1938)
- Palawan (Meso-Philippine language, Revel 1995)
- Kapampangan (N. Philippine language, Forman 1971a, 1971b; del Corro 1980): looking at Forman's (1971b) dictionary, both *d* and *g* vary, but with *d* non-substitution is more common, while with *g* substitution is more common. See Kaufman (2005) for a treatment of this case in terms on contrast preservation.

In all the languages included in Tryon (1995), the languages surveyed by Newman (1984) and by Blust (2004), and others whose descriptions I have encountered, Kapampangan is the only clear exception to Newman's implicational generalizations about voicing and place.

(26) Additional phonological regularities in Newman and Blust

- special treatment for pseudo-reduplicated stems
- special treatment for monosyllabic stems
- special treatment for stems that contain a nasal+obstruent sequence.
- Also Nomoto (2009), intriguing example from Malay: stem-initial *t͡ʃ* shows variation between substitution and not (*mən-t͡ʃinta* ~ *mə-ɲinta* 'to love', from *t͡ʃinta*). Nomoto uses web data to show that substitution is much more frequent when the stem contains a nasal+obstruent cluster than when it does not.

=> Phonological regulation of the distribution of nasal substitution in the lexicon is cross-linguistically common. Even if these factors now have categorical effects, these languages **must have gone through stages in which what are now regularities were merely tendencies**.

E.g. to get from Malay [variable and limited effect of nasal+obstruent cluster within stem] to Timugon Murut (Prentice 1971) [non-substitution is forbidden if the stem contains a nasal+obstruent sequence—the prefix nasal must either substitute or delete], probabilistic phonological effects on nasal substitution must not be a mere artifact of the lexicon, but must be learned and able to shape the treatment of new and even existing words.

=> Supports idea that lexical regularities can become encoded in the grammar.

==> More important for today's purposes, whatever problems are raised by variation in Tagalog probably are widespread (or have been).

Acknowledgments

This paper is a greatly revised version of portions of my 2000 UCLA dissertation. It has benefited from the comments and input of many people, who of course should not be held responsible for its errors or assumed to agree with its claims: Adam Albright, Marco Baroni, Katherine Crosswhite, Nenita Pambid Domingo, Bruce Hayes, Michael Kenstowicz, András Kornai, Peggy MacEachern, Donka Minkova, Joe Pater, Carson Schütze, Dan Silverman, Donca Steriade, and several reviewers. Ivan Tam wrote most of the software used in creating the web corpus, which took as its seed a corpus generously supplied by Rosie Jones from her own project, and Kevin Ryan did post-processing on the corpus. Thanks to Xiao Chen, Philip B. Ender, Michael Mitchell and Christine Wells for their advice on statistics for the somewhat unusual design of the survey. And thanks to all the anonymous survey participants who gave their judgments of novel words.

References

Abas, Husen and Charles E. Grimes. 1995. Bugis. In *Comparative Austronesian dictionary: An introduction to Austronesian studies*, ed. Darrell T. Tryon, Part 1: Fascicle 1, 549-561. Berlin/New York: de Gruyter.

Allen, Margaret. 1971. *Morphological investigations.* University of Connecticut Ph.D. dissertation.

Anceaux, J.C. and Charles E. Grimes. 1995. Wolio. *Comparative Austronesian dictionary: An introduction to Austronesian studies*, ed. Darrell T. Tryon, Part 1: Fascicle 1, 573-584. Berlin/New York: de Gruyter.

Anttila, Arto. 1997. Deriving variation from grammar: a study of Finnish genitives. In *Variation, change, and phonological Theory*, ed. F. Hinskens, R. van Hout and L. Wetzels, 35-68. Amsterdam: John Benjamins.

Anttila, Arto. 2002. Morphologically conditioned phonological alternations. *Natural Language and Linguistic Theory* 20: 1-42.

Aronoff, Mark. 1976. *Word formation in generative grammar.* Cambridge: MIT Press.

Archangeli, Diana, Laura Moll and Kazutoshi Ohno. 1998. Why not *NÇ? In *Proceedings of the 34th Annual Meeting of the Chicago Linguistic Society, Part I: The main session*, ed. M. Catherine Gruber, Derrick Higgins, Kenneth S. Olson and Tamra Wysocki, 1-26. Chicago: Chicago Linguistic society.

Baayen, R. Harald. 2008. languageR: Data sets and functions with "Analyzing Linguistic Data: A practical introduction to statistics". R package version 0.953.

Baayen, R. Harald and R. Schreuder. 1999. War and peace: morphemes and full forms in a non-interactive activation parallel dual route model. *Brain and Language* 68: 27–32.
Baroni, Marco. 2001. The representation of prefixed forms in the Italian lexicon: Evidence from the distribution of intervocalic [s] and [z] in northern Italian. In *Yearbook of Morphology 1999*, ed. Geert Booij and Jaap van Marle, 121-152. Dordrecht: Kluwer.
Barr, Donald F. 1995. Da'a. In *Comparative Austronesian dictionary: An introduction to Austronesian studies*, ed. Darrell T. Tryon, Part 1: Fascicle 1, 529-537. Berlin/New York: de Gruyter.
Bates, Douglas, Martin Maechler and Bin Dai. 2008. lme4: Linear mixed-effects models using S4 classes. R package version 0.999375-28. http://lme4.r-forge.r-project.org/
Becker, Michael. 2009. *Phonological trends in the lexicon: the role of constraints*. University of Massachusetts, Amherst Ph.D. dissertation.
Berko, Jean. 1958. The child's learning of English morphology. *Word* 14: 150-177.
Bhandari, Rita. 1997. Alignment and nasal substitution strategies in Austronesian languages. In *Recent papers in Austronesian linguistics. UCLA Occasional Papers in Linguistics* 21, ed. Matthew Pearson, 59-69. Los Angeles: UCLA Department of Linguistics.
Blake, Frank. 1925. *A grammar of the Tagalog language*. New Haven, CT: American Oriental Society.
Bloomfield, Leonard. 1917. *Tagalog texts with grammatical analysis*. Urbana, IL: University of Illinois.
Blust, Robert. 2004. Austronesian nasal substitution: a survey. *Oceanic Linguistics* 43: 73-148.
Boersma, Paul. 1997. How we learn variation, optionality, and probability. *Proceedings of the Institute of Phonetic Sciences of the University of Amsterdam* 21: 43–58.
Boersma, Paul. 1998. *Functional phonology: formalizing the interactions between articulatory and perceptual drives*. The Hague: Holland Academic Graphics.
Boersma, Paul and Bruce Hayes. 2001. Empirical tests of the Gradual Learning Algorithm. *Linguistic Inquiry* 32: 45-86.
Brakel, Lode F. 1973. Some notes on Malay morphology. *Linguistic Communications 11. Working Papers of the Linguistic Society of Australia*: 1-10.
Busenitz, Robert L. 1994. Marking focus in Belantak. In *Studies in Sulawesi Linguistics III*, ed. René van den Berg, 1-15. Jakarta: Universitas Katolik Indonesia Atma Jaya.
Busenitz, Robert L. and Marilyn J. Busenitz. 1991. Balantak phonology and morphophonemics. In *Studies in Sulawesi Linguistics II*, ed. James N. Sneddon, 29-47. Jakarta: Universitas Katolik Indonesia Atma Jaya.
Bybee, Joan. 2001. *Phonology and language use*. Cambridge: Cambridge University Press.
Carrier, Jill. 1979. *The interaction of morphological and phonological rules in Tagalog*. Massachusetts Institute of Technology Ph.D. dissertation.
Chomsky, Noam and Morris Halle. 1968. *The sound pattern of English*. Cambridge, MA: MIT Press.
Clements, George N. 1990. The role of the sonority cycle in core syllabification. In *Papers in Laboratory Phonology I: Between the grammar and the physics of speech*, ed. John Kingston and Mary Beckman, 283-333. Cambridge: Cambridge University Press.
Coetzee, Andries. 2000. Post-nasal neutralization phenomena in Tswana: more on *N... constraints. Ms., University of Massachusetts, Amherst.
Coetzee, Andries, Susan Lin and Rigardt Pretorius. 2007. Post-nasal devoicing in Tswana. In *Proceedings of the 16th International Congress of Phonetic Sciences*, ed. Jürgen Trouvain and William J. Barry, 861-864.
Coetzee, Andries and Joe Pater. 2008. The place of variation in phonological theory. Ms., University of Michigan and University of Massachusetts, Amherst. Draft chapter for 2nd edition of the *Handbook of Phonological Theory*, ed. by John Goldsmith, Jason Riggle, and Alan Yu.
Cohn, Abigail and John McCarthy. 1994/1998. Alignment and parallelism in Indonesian phonology. Ms., University of Massachusetts, Amherst. Published in *Working Papers of the Cornell Phonetics Laboratory* 12, 53–137.
De Guzman, Videa. 1978. A case for nonphonological constraints on nasal substitution. *Oceanic Linguistics* 17: 87-106.
de Lacy, Paul. 2001. Markedness in prominent positions. In *HUMIT 2000, MIT Working Papers in Linguistics* 40, ed. Ora Matushansky, Albert Costa, Javier Martin-Gonzalez, Lance Nathan and Adam Szczegielniak, 53-66.
de Lacy, Paul. 2002. The formal expression of markedness. University of Massachusetts Amherst Ph.D. dissertation..
del Corro, Anicia. 1980. *Kapampangan morphophonemics*. Quezon City: Cecilio Lopez Archives of Philippine Languages and the Philippine Linguistics Circle. Publication 1 of *The Archive*.

Delilkan, Ann. 2002. Fusion and other segmental processes in Malay: the crucial role of prosody. New York University Ph.D. dissertation.
Dell, François and Mohamed Elmedlaoui. 1985. Syllabic consonants and syllabification in Imdlawn Tashliyt Berber. *Journal of African Languages and Linguistics* 7: 105-130.
DuBois, Carl. D. n.d. *Sarangani Manobo: an introductory guide*. Manila: Linguistic Society of the Philippines.
English, Leo. 1986. *Tagalog-English dictionary*. Manila: Congregation of the Most Holy Redeemer. Distributed by. Philippine National Book Store.
Ferreirinho, Naomi. 1993. *Selected topics in the grammar of Limos Kalinga, the Philippines*. Pacific Linguistics Series B No. 109. Canberra: Australian National University.
Flack, Kathryn. 2007. *The sources of phonological markedness*. University of Massachusetts, Amherst Ph.D. dissertation.
Forman, Michael L. 1971a. *Kapampangan grammar notes*. Honolulu: University of Hawaii Press.
Forman, Michael L. 1971b. *Kapampangan dictionary*. Honolulu: University of Hawaii Press.
Forster, Jannette. 1970. Morphophonemic changes in Dibabawon. *Pacific Linguistics A* 24: 63-70.
French, Koleen Matsuda. 1988. *Insights into Tagalog: reduplication, infixation and stress from nonlinear phonology*. Dallas, TX: Summer Institute of Linguistics and University of Texas at Arlington.
Fujimura, Osamu. 1962. Analysis of nasal consonants. *Journal of the Acoustical Society of America* 34: 1865-1875.
Ghani, Jones and Mladenić. 2004. Building minority language corpora by learning to generate Web search queries. *Knowledge and Information Systems* 7: 56-83.
Gieser, C. Richard. 1970. The morphophonemic system of Guininaang Kalinga. *Philippine Journal of Linguistics* 1: 52-68.
Goldwater, Sharon and Mark Johnson. 2003. Learning OT constraint rankings using a Maximum Entropy model. In *Proceedings of the Stockholm workshop on variation within Optimality Theory*, ed. Jennifer Spenader, Anders Eriksson and Östen Dahl, 111-120.
Gordon, Raymond G., Jr., ed. 2005. *Ethnologue: Languages of the World*. 15th edition. Dallas, Texas: SIL International. Online version: http://www.ethnologue.com/.
Goris, R. 1938. *Beknopt Sasaksch–Nederlandsch woordenboek*. Singradja, Bali: Kirtya Liefrinck-van der Tuuk.
Halle, Morris. 2001. Infixation versus onset metathesis in Tagalog, Chamorro, and Toba Batak. In *Ken Hale: a life in language*, ed. Michael Kenstowicz, 153-168. Cambridge, MA: MIT Press.
Harrell, Frank E. Jr.. 2008. Hmisc: Harrell Miscellaneous. R package version 3.5-2. http://biostat.mc.vanderbilt.edu/s/Hmisc.
Hay, Jennifer. 2003. *Causes and consequences of word structure*. New York and London: Routledge.
Hay, Jennifer and Harald Baayen. 2005. Shifting paradigms: gradient structure in morphology. *Trends in Cognitive Sciences* 9: 342-348.
Hayes, Bruce. 1999. Phonetically-driven phonology: the role of Optimality Theory and inductive grounding. In *Functionalism and formalism in linguistics, Volume I: General papers*, ed. Michael Darnell, Edith Moravcsik, Frederick Newmeyer, Michael Noonan and Kathleen Wheatly, 243-285. Amsterdam: John Benjamins.
Hayes, Bruce and Margaret MacEachern. 1998. Quatrain form in English folk verse. *Language* 74: 473-507.
Hayes, Bruce and Tanya Stivers. 1995. The phonetics of postnasal voicing. Ms., University of California, Los Angeles.
Hayes, Bruce, Bruce Tesar, Colin Wilson and Kie Zuraw. 2005. OTSoft 2.3. Software package, www.linguistics.ucla.edu/people/hayes/otsoft/.
Herbert, Robert K. 1980. The role of perception in restructuring and relexicalization: two case histories. In *Papers from the 4th International Conference on Historical Linguistics*, ed. Elizabeth Closs Traugott, Rebecca Labrum and Susan Shepherd, 211-220. Amsterdam: John Benjamins.
Hyman, Larry. 2001. On the limits of phonetic determinism in phonology: *NC revisited. In *The role of speech perception in phonology*, ed. Elizabeth Hume and Keith Johnson, 141-185. San Diego, CA: Academic Press.
Inkelas, Sharon, Orhan Orgun, and Cheryl Zoll. 1997. The implications of lexical exceptions for the nature of grammar. In *Derivations and Constraints in Phonology*, ed. Iggy Roca, 393-418. New York, Oxford University Press.
Inkelas, Sharon and Cheryl Zoll. 2000. Reduplication as morphological doubling. Ms., University of California, Berkeley and Massachusetts Institute of Technology.
Inkelas, Sharon and Cheryl Zoll. 2005. *Reduplication: Doubling in morphology.* Cambridge: Cambridge University Press.

Inkelas, Sharon and Cheryl Zoll. 2007. Is grammar dependence real? A comparison between cophonological and indexed constraint approaches to morphologically conditioned phonology. *Linguistics* 45: 133-171.
International Phonetic Association. 1999. Handbook of the International Phonetic Association: A guide to the use of the International Phonetic Alphabet. Cambridge: Cambridge University Press.
Itô, Junko. 1986. *Syllable theory in prosodic phonology*. University of Massachusetts, Amherst Ph.D. dissertation.
Itô, Junko and Armin Mester. 1995. The core-periphery structure of the lexicon and constraints on reranking. In *University of Massachusetts Occasional Papers in Linguistics 18: Papers in Optimality Theory*, ed. Jill Beckman, Suzanne Urbanczyk, and Laura Walsh Dickey, 181-209.
Itô, Junko and Armin Mester. 1999. The structure of the phonological lexicon. In *The Handbook of Japanese Linguistics*, ed. Natsuko Tsujimura, 62-100. Malden, MA and Oxford: Blackwell.
Johnson, Keith. 1997. *Acoustic and auditory phonetics*. Cambridge, MA: Blackwell.
Kaufman, Daniel. 2005. Austronesian nasal substitution and contrast neutralization. Ms., Cornell University.
Kiparsky, Paul. 1982. Lexical phonology and morphology. In *Linguistics in the Morning Calm*, ed. In-Seok Yang, 3-91. Seoul: Hanshin.
Kiparsky, Paul. 1993. Blocking in non-derived environments. In *Studies in lexical phonology*, ed. Sharon Hargus and Ellen Kaisse, 277-313. San Diego, CA: Academic Press.
Labov, William. 1969. Contraction, deletion, and inherent variability of the English copula. *Language* 45: 715-762.
Lapoliwa, Hans. 1981. *A generative approach to the phonology of Bahasa Indonesia*, Canberra: Department of Linguistics, Research School of Pacific Studies, Australia National University.
MacBride, Alex. 2004. *A constraint-base approach to morphology*. University of California, Los Angeles Ph.D. dissertation.
Marantz, Alec. 1982. Re reduplication. *Linguistic Inquiry* 13: 435-482.
Martin, Andrew. 2007. *The evolving lexicon*. University of California, Los Angeles Ph.D. dissertation.
McCarthy, John. 1983. Consonantal morphology in the Chaha verb. In *Proceedings of the Meeting of the West Coast Conference on Formal Linguistics* 2, ed. Michael Barlow, Daniel P. Flickinger, and Michael T. Wescoat, 176-188. Stanford, CA: Stanford Linguistics Department.
McCarthy, John. 2002. Comparative markedness [long version]. In *University of Massachusetts Occasional Papers in Linguistics 26: Papers in Optimality Theory II*, ed. Angela C. Carpenter, Andries W. Coetzee and Paul de Lacy, 171-246.
McCarthy, John. 2003. Comparative markedness [short version]. *Theoretical Linguistics* 29: 1-51.
McCarthy, John and Alan Prince. 1993.Generalized alignment. In *Yearbook of Morphology 1993*, ed. Geert Booij and Jap van Marle, 79-153. Dordrecht: Kluwer.
McCarthy, John and Alan Prince. 1995. Faithfulness and reduplicative identity. In *University of Massachusetts Occasional Papers in Linguistics 18: Papers in Optimality Theory*, ed. Jill Beckman, Suzanne Urbanczyk and Laura Walsh Dickey, 249–384.
Merriam-Webster. 1994. *Merriam-Webster's Collegiate Dictionary*, 10th edition. Springfield, MA: Merriam-Webster. Frederic C. Mish, Editor in Chief
Meyer, David, Achim Zeileis and Kurt Hornik. 2006. The Strucplot Framework: Visualizing Multi-Way Contingency Tables with vcd. *Journal of Statistical Software* 17: 1-48.
Meyer, David, Achim Zeileis and Kurt Hornik. 2007. vcd: Visualizing Categorical Data. R package version 1.0-6. Software program.
Nababan, P. W. J. 1981. *A grammar of Toba-Batak*. *Materials in languages of Indonesia No. 6*. Pacific Linguistics Series D No. 37. Canberra: Australian National University.
Nagy, Naomi and Bill Reynolds. 1997. Optimality theory and variable word-final deletion in Faetar. *Language Variation and Change* 9:37-56.
Narayan, Chandan. 2006. Acoustic-perceptual salience and developmental speech perception. University of Michigan Ph.D. dissertation.
Newman, John. 1984. Nasal replacement in Western Austronesian: an overview. *Philippine Journal of Linguistics* 15-16: 1-17.
Nomoto, Hiroki. 2009. Distantly and prosodically conditioned nasal substitution in Austronesian languages. Talk presented at AFLA [Austronesian Formal Linguistics Association] XVI, University of California, Santa Cruz.
van Odijk, Antonio. 1959. *Elementary grammar of the Bisayan language*. Cebu: Convento Opon.

Ohala, John J. and Carol J. Riordan. 1979. Passive vocal tract enlargement during voiced stops. In *Speech communication papers*, ed. Jared J. Wolf and Dennis H. Klatt, 89-92. New York: Acoustical Society of America.
Pater, Joe. 1999. Austronesian nasal substitution and other NC effects. In *The prosody morphology interface*, ed. René Kager, Harry van der Hulst, and Wim Zonneveld, 310-343. Cambridge: Cambridge University Press.[Edited version in *Optimality Theory in Phonology: A Reader*, ed. John McCarhy. Oxford and Malden, MA: Blackwell.]
Pater, Joe. 2001. Austronesian nasal substitution revisited: what's wrong with *NC̥. and what's not. In *Segmental phonology in Optimality Theory: constraints and representations*, ed. Linda Lombardi, 159-182. Cambridge: Cambridge University Press.
Pater, Joe. 2006. The locus of exceptionality: Morpheme-specific phonology as constraint indexation. In *UMOP: Papers in Optimality Theory III*, ed. Leah Bateman and Adam Werle, 1–36.
Pater, Joe. 2008. Gradual learning and convergence. *Linguistic Inquiry* 39: 334-345.
Percival, W. K.. 1981. *A grammar of the urbanised Toba-Batak of Medan.* Pacific Linguistics Series B No. 76. Canberra: Australian National University.
Prentice, D.J.. 1971 *The Murut languages of Sabah*. Pacific Linguistics Series C No. 18. Canberra: Australian National University.
Prince, Alan. 1997. Stringency and anti-Paninian hierarchies. Handout from lecture given at the LSA institute, 6/26/1997.
Prince, A. and P. Smolensky. 1993/2004. *Optimality Theory: constraint interaction in generative grammar*. Malden, MA: Blackwell. Originally circulated. 1993 as Technical Report TR-2. Rutgers Center for Cognitive Science/Technical Report CU-CS-696-93. University of Colorado at Boulder Department of Computer Science.
R Development Core Team. 2007. R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. Software package, www.R-project.org.
Raimy, Eric. 2000. *The phonology and morphology of reduplication*. Berlin: Mouton de Gruyter.
Revel, Nicole. 1995. Le palawan: phonologie, catégories, morphologie. Paris: SELAF.
Reynolds, Bill and Naomi Nagy.. 1994. Phonological variation in Faetar: an Optimality account. In *Chicago Linguistic Society* 30-II: *Papers from the Parasession on Variation and Linguistic Theory*, ed. Katharine Beals, Jeannette Denton, Robert Knippen, Lynette Melnar, Hisami Suzuki and Erica Zeinfeld, 277-292. Chicago: Chicago Linguistic Society.
Ross, Kie. 1996. Floating phonotactics: infixation and reduplication in Tagalog loanwords. University of California, Los Angeles M.A. thesis.
Ross, Malcolm. 1988. *Proto Oceanic and the Austronesian languages of Western Melanesia.* Canberra: Pacific Linguistics.
Ross, Malcolm. 1995. Some current issues in Austronesian linguistics. In *Comparative Austronesian dictionary: An introduction to Austronesian studies*, ed. Darrell T. Tryon, Part 1: Fascicle 1, 45-120. Berlin/New York: de Gruyter.
Schachter, Paul and Fe Otanes. 1972. *Tagalog reference grammar*. Berkeley: University of California Press.
Siegel, Dorothy. 1974. *Topics in English morphology.* Massachusetts Institute of Technology Ph.D. dissertation.
Tranel, Bernard. 1987. French schwa and nonlinear phonology. *Linguistics* 25:845-66
Tranel, B. 1996. Exceptionality in Optimality Theory and final consonants in French. In *Grammatical Theory and Romance Languages*, ed. Karen Zagona, 275-291. Amsterdam: Benjamins.
Tryon, Darrell T., ed. 1995. *Comparative Austronesian dictionary: an introduction to Austronesian studies*. Berlin/New York: de Gruyter.
van der Tuuk, H. N. 1867/1971. *A grammar of Toba Batak.* Translation, with incorporation of van der Tuuk's manuscript notes and corrections, of *Tobasche spraakunst*. The Hague: Martinus Nijhoff.
Vanoverbergh, Morice. 1972. *Isneg-English Vocabulary*. Honolulu: University Press of Hawaii.
Verheijen, Jilis A.J. 1986. *The Sama/Bajau language in the Lesser Sunda Islands*. Pacific Linguistics Series D No. 7; Materials in Languages of Indonesia No. 32. Canberra: Department of Linguistics, Research School of Pacific Studies, Australian National University.
Walther, Markus and Richard Wiese. 1999. Optimization versus lexical specification. Handout from the workshop Conflicting Rules in Phonology and Syntax, University of Potsdam.
Weinreich, Uriel, William Labov, and Marvin Herzog. 1968. Empirical foundations for a theory of language change. In *Directions for historical linguistics*, ed. Winfred Lehmann and Yakov Malkiel, 97-195. Austin: University of Texas Press.

West, Anne. 1995. Yami. In *Comparative Austronesian Dictionary: An introduction to Austronesian studies*, ed. Darrell T. Tryon, Part 1: Fascicle 1, 315-320. Berlin/New York: de Gruyter.
Wilbur, Ronnie Bring. 1973. *The phonology of reduplication.* Bloomington, IN: Indiana University Linguistics Club.
Wolff, John. 1962. *A description of Cebuano Visayan*. Ithaca: Division of Modern Languages, Cornell University.
Woollams, Geoff. 1996. *A Grammar of Karo Batak, Sumatra*. Canberra: Department of Linguistics, Research School of Pacific Studies, Australian National University.
Yip, Moira. 2002. *Tone*. Cambridge: Cambridge University Press.
Yip, Moira. 2007. Tone. In *The Cambridge handbook of phonological theory*, ed. Paul de Lacy, 229-251. Cambridge University Press.
Zhang, Jie, Yuwen Lai and Craig Sailor. To appear. Effects of phonetics and frequency on the production of Taiwanese tone sandhi. In *Proceedings of Chicago Linguistic Society* 43.
Zonneveld, Wim. 1978. *A formal theory of exceptions in generative phonology*. Dordrecht: Foris.
Zoll, Cheryl. 1996. *Parsing below the segment in a constraint based framework*. University of California, Berkeley Ph.D. dissertation.
Zorc, R. David Paul. 1979. *Core etymological dictionary of Filipino.* Batchelor, Australia: Darwin Community College.
Zsiga, Elizabeth, Maria Gouskova and One Tlale. 2006. On the status of voiced stops in Tswana: against *ND. In *Proceedings of the North Eastern Linguistic Society* 36, ed. Christopher Davis, Amy Rose Deal and Youri Zabbal, 721-734. Amherst, MA: GLSA.
Zuraw, Kie. 2000. *Patterned exceptions in phonology.* University of California, Los Angeles Ph.D. dissertation.
Zuraw, Kie. 2006. Using the web as a phonological corpus: a case study from Tagalog. In *EACL-2006: Proceedings of the 11th Conference of the European Chapter of the Association for Computational Linguistics/Proceedings of the 2nd International Workshop on Web As Corpus*, 59-66.