Syntatic and Lexical Factors Conditioning the Diffusion of Sound Change

Mark Donohue

Oceanic Linguistics, Volume 44, Number 2, December 2005, pp. 427-442 (Article)

Published by University of Hawai'i Press
DOI: 10.1353/ol.2005.0032

Syntactic and Lexical Factors Conditioning the Diffusion of Sound Change

Mark Donohue

NATIONAL UNIVERSITY OF SINGAPORE

A sound change may propagate through a language in different ways. Different studies attest sound changes spreading at different rates through different phonological and/or phonotactic environments, diffusing through the speaker population (or through different dialects) in different ways, or simply spreading differentially through the lexicon. In Palu'e there is evidence for a sound change applying at different rates for different grammatical categories, with the sound change advancing in the small set of bound grammatical morphemes perhaps more completely than in free lexemes. This is evidence that syntactic information on parts of speech can affect the diffusion of a sound change through a language, and that bound forms are not necessarily more conservative than free lexemes when it comes to phonological change.

1. WAYS FOR CHANGE TO DIFFUSE.[1] While there are many described instances of the exceptionless application of sound changes throughout a language, following the Neogrammarian model, there are many instances in which a sound change is found to apply to only some subset of the possible target words in a language. In these cases, the application of a sound change may be delimited by the phonological or phonotactic environment of the relevant sound, by different varieties of the language, or it may simply be that a sound change applies lexeme-by-lexeme—which is a confession of a sound change being irregular.

A combination of the first two factors can be seen in the *s > *h* sound change in the Kwerba languages of the Mamberamo and Apauwer rivers in Irian Jaya. Anggreso Kwerba, spoken on a tributary of the Apauwer river east of but close to the Mamberamo, is an example of a Kwerba language that has not been affected by the sound change, and shows *s reflected as *s* in all environments and in all lexemes; it is a control for the rest of the data. The other languages, arranged from south to north along the river, show different degrees to which the *s > *h* rule has applied. The words in (1) show reflexes of *s in syllable-onsets in Anggreso, spoken between the Apauwer and the Mamberamo, and four other Kwerba languages spoken along the Mamberamo river, arranged from south to north. In (2) we can see reflexes of *s in syllable-codas. In both cases Anggreso shows no change, while the Mamberamo languages all show var-

1. Thanks to Bethwyn Evans, Harold Koch, and Andy Pawley for various discussions of points raised in this article.

Oceanic Linguistics, Volume 44, no. 2 (December 2005)

ious degrees of change of *s to *h.* This change has applied most thoroughly in onset positions, showing phonotactic conditioning of the sound change, with only one lexical item in this set (Bagusa 'earth, forest') resisting the change.

(1) Reflexes of Proto-Kwerba *s in syllable-onset position: Geographic (and lexical) variables controlling diffusion

	'EARTH, FOREST'	'GOOD'	'FIRE'	'HOT'
Anggreso	ɪsɪw	e**s**am	**s**er	papa**s**am
Kwerba	ɪ**h**ɪw	(anɪɪn)	**h**er	(ɪhpwam)
Kasonaweja	ɪ**h**ɪw	e**h**ama	**h**er	papa**h**am
Trimuris	ɪ**h**ɪw	e**h**am	**h**er	papa**h**am
Bagusa	ɪsɪw	e**h**am	**h**ɪr	papa**h**am

In coda position, however, the sound change is much less advanced, being resisted completely in Kwerba and Bagusa at the opposite ends of the Kwerba language chain along the river, and applying most strongly to the variety spoken in Trimuris. In Kasonaweja there is free variation (both idiolectal, with some speakers favoring *s* and some *h,* and also simply random, with degrees of variation observed from the same speaker) in the pronunciation of 'cassowary', while the two different pronunciations of the reflexes of *ɪs 'woman' reflect its split into two lexemes in the modern Kasonaweja language, *is* 'woman' and *ih* 'wife', differentiated by the application of the *s > *h* change.

(2) Reflexes of Proto-Kwerba *s in syllable-coda position: Geographic and lexical variables controlling diffusion

	'SNAKE'	'CUSCUS'	'CASSOWARY'	'WOMAN'
Anggreso	ɪs**y**a	qaraa**s**	qama**s**	ɪ**s**
Kwerba	ɪs**y**a	qaraa**s**	qama**s**	ɪ**s**
Kasonaweja	ɪs**y**a	qaraa**s**	qama**s**, qama**h**	ɪ**s** (also ɪ**h**)
Trɪmuris	ɪs**y**a	qaraa**h**	qama**h**	ɪ**h**
Bagusa	(sopɪs)	qaraa**s**	qama**s**	ɪ**s**

The data here can only be described by reference to different environments conditioning the change, to a spread at different rates among the different populations, and to differential lexical diffusion.

Another example of lexical diffusion can be seen in the data concerning the fate of Proto-Austronesian *b in selected Austronesian languages from Sulawesi, shown in (3).

(3) Reflexes of PAN *b in languages of southeast Sulawesi: Geographic and lexical variables controlling diffusion

		WOLIO	KAIMBULAWA	MUNA	TUKANG BESI	CIA-CIA
*be(R)say	'paddle'	**b**	**b**	**b**	**b**	**b**
*beŋel	'deaf'	**b**	**b**	**b**	**b**	—
*baqeRu	'new'	**b**	**b**	**b**	β	β
*bulan	'moon'	**b**	**b**	β	—	β
*batu	'stone'	**b**	β	β	β	β
*bibiR	'lip'	**b**/β	β	β	β	β

Here we can see that the application of the *b > β change is not completely regular and exceptionless in any of the languages concerned, but must be described lexically. The change has spread furthest through the lexicon in Cia-Cia, and has only barely begun

to be applied to Wolio, which belongs to a different subgroup of Celebic (Donohue 2005). While Kaimbulawa and Muna are closely related to each other, they are both more closely related to Cia-Cia than they are to Tukang Besi (van den Berg 2002).

Clearly, as is widely acknowledged, sound changes apply not only in an unexceptional way, but in a variety of unexceptional ways. In the remainder of this article I wish to examine another way in which the application of a sound change is not symmetrical throughout the language, where the division between free words and bound morphemes marks the division between the early application of a sound change and its late application. In other words, in addition to phonologically, dialectally, and lexically conditioned sound change, I examine here a case of grammatically conditioned sound change: the only delimiting factor in describing the spread of the sound change is the syntactic category of the lexical items.

2. THE GENITIVE CLITICS OF PALU'E. Consider the inflected forms of a typical nominal in Palu'e, showing agreement for different possessors by genitive clitics.

(4) rero 'friend'
rero-gu 'my/our (exclusive) friend'
rero-de 'our (inclusive) friend'
rero-mo 'your friend'
rero-n 'her/his/their friend'

It is clear that the forms in (4) consist of a root, *rero* 'friend', which is marked for possession by means of the genitive clitics *-gu, -de, -mo,* and *-n,* which mark 1GEN, 12GEN, 2GEN, and 3GEN possessors, respectively.[2] This exemplifies the regular genitive paradigm, with the forms employed clearly traceable to earlier Austronesian forms, most likely reflecting Proto–Malayo-Polynesian *-ku '1SG.GEN', *-ta '1PL.IN.GEN', *-mu '2SG.GEN', and *-ña '3SG.GEN', with a clear loss of number distinction in the Palu'e forms. While the function of these clitics is not exceptional, their form is unusual for a morpheme in Palu'e, for the following reasons:

1. None of the bound forms contains a low vowel; the vowel *a* is overwhelmingly the most common vowel in the language, but is not found in the genitive paradigm.
2. Despite the fact that voiceless stops are more frequent than voiced stops in all positions in lexical roots, especially initially (where *d* is attested in only one word, and *g* not at all), all of the stops that appear in the genitive paradigm are voiced.
3. The genitive paradigm contains the only instance of a morpheme that lacks a vowel entirely, the third person *-n*. This is also the only morpheme to end in a consonant lexically.
4. The clitics of the genitive paradigm are not within the domain of stress assignment; they can neither attract stress nor affect the placement of stress on a word.

2. Number is not marked on the genitive clitics, but can be distinguished by the use of the free form pronouns: *kau reremo* 'your (SG) friend', *miu reremo* 'your (PL) friend', and *ia reron* 'her/his friend', but *konen reron* 'their friend'. There are some complications for first person possessors, but they do not concern us here. One reason to consider these to be clitics, and not affixes, is that they follow any modifying adjectives: *rero mbolagu* 'my good friend', not **rerogu mbola,* and do not affect the location of stress: *réro* 'friend', *rérogu* (***rerógu*) 'my friend' (see point 4).

We can, thus, establish that the clitics behave in a manner different from free lexemes: they are phonologically distinct. Most important, for the purposes of this discussion at least, are the oral stops in the clitics *-gu* and *-de.* These appear to reflect the Proto–Malayo-Polynesian clitics *-ku and *-ta, but when we examine other items in Palu'e that reflect *k and *t, we do not find voiced stops.

(5) Regular reflexes of *t and *k

PMP *t	PALU'E		PMP *k	PALU'E	
*təlu	tlu	'three'[3]	*kuku	kuku	'fingernail'
*ma-qitəm	mite	'black'	*kulit	loke	'skin'
*tasik	tai	'sea'	*likuD	liku	'back, behind'
*kutu	kutu	'louse'	*siku	iku	'elbow'

Can we claim that *-gu* reflects *-ku when the regular reflex of *k in Palu'e is *k* (with similar doubts about *d* reflecting *t)? In this paper I argue that in fact a *g* reflex of *k is in a sense regular, but that the voicing sound change has only begun to apply, and is in the process of moving through the lexicon in a nonrandom fashion, applying to the (higher-frequency) bound morphemes first. This argues for the differential application of sound changes to different parts of speech. The argument I present is based on the phonotactic distribution of voiced and other stops in modern Palu'e, and the peculiar arrangement and distribution of the bilabial obstruents.

3. HISTORICAL CORRESPONDENCES. We have seen that the regular reflexes of *t and *k in Palu'e are *t* and *k,* regardless of environment, unlike *p (see [7] below). Examine the correspondences in (6) involving Proto–Malayo-Polynesian voiced obstruents; in all cases, the Palu'e reflex shows lenition, to *v, l,* or *r.* (A small number of roots reflect *b as *p: paɲu* 'wake up', for instance. It is not known if this is an irregularity in Palu'e, or reflects a complex history of interaction between Palu'e and other Austronesian languages. I assume that *j was a high voiced obstruent, systemically the voiced correspondent of *k [Ross 1992].)

(6) Regular reflexes of *b, *d/D, *Z, and *j

PMP *b	PALU'E		PMP *d/D	PALU'E	
*batu	vatu	'stone'	*DaRaq	ladʒa	'blood'
*bibiR	vivi	'lip'	*Duha	rua	'two'
*bulu	vulu	'hair'	*t[i/u]D[i/u]R	tuli	'sleep'
*babuy	vavi	'pig'	*dapuR	laʔɓo	'kitchen'
PMP *Z	PALU'E		PMP *j	PALU'E	
*Zalan	lala	'path'	*hua(n)ji	ari	'younger sibling'
*quZan	ʔura	'rain'	*qa(m)pəju	ʔɓru	'spleen'
			*qaləjaw	ʔra	'sun'
			*pija	ʔɓira	'how many'

3. CC sequences such as the *tl* in *tlu* 'three' or *tlo* 'egg' are realized with an epenthetic schwa breaking up the consonant cluster. The regular reflex of *ə is *e* in Palu'e (for instance, *mite* 'black' < *ma-qitəm, *lale* 'fly' < *lalət, but it is tempting to treat some of these schwas as reflexes of earlier schwas, such as in [təlo] < *qatəluR. There are enough schwas that do not have an etymological schwa to require us to posit a process of epenthesis in the language anyway: [ləma] 'tongue' < *maya (with metathesis), [məni] 'mosquito' < *nyamuk. Synchronically, however, the best analysis is clearly to treat these extrametrical schwas as epenthetic, reflecting just one strategy that is part of a broader historical move away from clusters (the other involving NC sequences reducing to a voiced C; see (10) and (11).

The loss of an entire voiced series is exactly the sort of circumstance that can lead to a drag chain forming, replacing the lost series with another, such as Grimm's Law in Germanic. This seems to be what has happened in Palu'e, where the loss of the voiced bilabial stop *b through lenition has created a gap that has resulted in a voicing (and preglottalizing) rule applying to *p, because that "area" of phonetic space had become empty. This is shown in (7).

(7) The drag chain involving bilabial obstruents:

*b > *v*

*p > ʔɓ

Examples of bilabial correspondences illustrating the accretion of voicing (as well as preglottalization and slight implosion) to *p can be seen in (8)—see also 'spleen' and 'how many' in (6). Note, however, that not all the reflexes of *p are voiced; all intervocalic instances of *p are reflected as *ʔɓ,* but the root-initial reflexes are a mix of *p* and *ʔɓ*. There are no medial *p*s in Palu'e.

(8) Regular reflexes of *p

a. Initial

PMP	PALU'E		PMP	PALU'E	
*pusəj	ʔɓuse	'navel'	*puluq	pulu	'ten'
*palu	ʔɓalu	'hit'	*panaw	pana	'walk'
*paRi	ʔɓadʒi	'stingray'	*paqit	paʔi	'bitter'
*puqun	ʔɓuʔu	'tree'	*penuq	pnu	'full'
*paqa	ʔɓaʔa	'thigh'			

b. Medial

PMP	PALU'E		PMP	PALU'E	
*papan	ʔɓaʔɓa	'plank, board'	*apa	aʔɓa	'what'
*hapuy	aʔɓi	'fire'	*hipaR	iʔɓa	'in-law'
*dapuR	laʔɓo	'kitchen'			

Can we ascribe the split between *p > *p* and *p > *ʔɓ* in Palu'e to any conditioning factor? Clearly the voicing of *p has applied completely to instances of intervocalic *p, and is in the process of advancing through the lexicon in initial position. Generalizing further, we can see some general conditions that apply to the phonotactic possibilities for voiced stops (which, other than *ʔɓ,* are rare) in Palu'e.

1. The voiced obstruents *d, dʒ,* and *g* are not found initially in lexical roots.
2. Of the voiced obstruents, only *ʔɓ* is found as the second member of a C_1C_2V root.

The bilabial *ʔɓ* displays quite different behavior from the other voiced stops: it is by far the most frequent of the phonetically voiced stops/affricates, comprising 48 percent of all tokens (the other figures are: *d* 12%, *dʒ* 24%, *g* 15%); it is the only voiced stop to appear initially in a lexical root; and, compared with the voiceless bilabial stop, it shows the following phonotactic peculiarities:[4]

1. **vVʔɓV* is not an attested pattern (*ʔɓVvV* is attested), while *s*V*d*V and *s*V*t*V are both attested.

4. Further restrictions include the lack of roots with the form **bvV* and **vbV,* but these are part of a general ban on fricatives and stops from the same place of articulation appearing in the same monovocalic root.

2. Neither *pˀɓ*V or *pr*V ([pə̆ˈˀɓV] and [pə̆ˈrV] respectively, with epenthetic vowels) are attested in any roots, while *td*V and *tr*V are.
3. *p* is not found in intervocalic position, while both *t* and *k* are.

By contrast, stops from the other places do not show restrictions with respect to their appearance with other consonants of the same place of articulation. The fact that the instances of (intervocalic) *d, ʤ,* and *g* are mostly onomatopoeic words or words that do not appear to be reflexes of known Austronesian reconstructions (see tables 1–4 below) suggests that the appearance of voiced stops in Palu'e is largely the result of their recent and ongoing reintroduction. This has followed the earlier loss of this series through lenition to fricatives, which was partially described in (6) and (7), through paths other than regular sound changes applying to Proto-Austronesian lexemes.

The *p > *ˀɓ* change that was illustrated in (7) and that does apply to Proto-Austronesian roots is following this trend toward the reinforcing of voicing as a functional contrast in the phonological system, but there is only scattered evidence for a more general rule that operates to voice the voiceless stops, such as in (9) (we can also see the merger of *D and *j as *l/r* in Palu'e).

(9) The drag chain resulting in voiced obstruents:

*b, *D, *j > v, l/r, l/r

*p, *t, *k > p/ˀɓ, t/(d), k/(g)

Examples of *d* and *g* in Palu'e can be seen in tables 1–4. Overwhelmingly these words do not reflect Proto-Austronesian reconstructions. Those occurrences of intervocalic *d* that do appear to be reflexes of Austronesian roots are shown in table 1 (drawing on Wurm and Wilson 1975 as a main source). In five instances, *tʃdu, mude, ridi, sudu,* and *tadu,* the *d* reflects *nt/*nd. In four more, it might reflect a *t: *lade katen / lodo*, possibly *pdo*, possibly *tdu* (involving an etymology with metathesis of *R and *t: **tutuR) and *vide*, and in one instance, *todo*, it appears to reflect an unlenited *d. Examples of the loss of the nasal in an NC sequence are attested in modern loanwords, shown in (10). Further instances of intervocalic *d* that do not appear to reflect any lexemes attested beyond Flores are shown in table 2.

Examining the voiced velar stop *g*, we find that there is a smaller proportion of identifiable Austronesian reflexes. Table 3 lists the reflexes of identified non-Flores reconstructions. Two of these, the reconstructions given for *mugu* and *pga*, are highly speculative. Of the others, one, *toge*, appears to reflect an NC sequence; one, *vago*, genuinely (and

TABLE 1. INSTANCES OF *d* IN PALU'E WITH RECONSTRUCTABLE ANCESTRY

PAN		PALU'E	
*munti	'lemon'	mude	'citrus fruit'
*diŋdiŋ	'wall'	ridi	'wall'
*sundul	'hit'	sudu	'punch'
*tanduk	'horn'	tadu	'horn'
*lateŋ	'nettle'	lade katen	'jellyfish'
		lodo	'liana'
*betis	'calf'	vide	'calf'
*tedes	'crush'	todo	'stamp on'
*ta	'we'	de	'our'

irregularly—see [5]) reflects *k, and one, *lgu*, apparently reflects a form reconstructable only for Philippines languages.[5] Instances of *g* without a plausible reconstruction outnumber those with etymologies by eight to one, and are listed in table 4.

In common, these loans show nasal + stop sequences from the source language being regularly borrowed as simple stops, as shown in table 5. Nasals followed by (homorganic) voiceless stops are not permissible sequences in native Palu'e words. The sequence *nt* is not attested in Palu'e, except in the obvious loanword *kntaŋ* 'sweet potato' < Indonesian *kentang* 'potato'. Similarly, *ŋk* is attested only in *kaŋkoŋ* < Indonesian *kaŋkuŋ* 'swamp greens'; there are no known instances of *mp* in Palu'e.

While we have seen that there is an incipient drag chain operating for the bilabial series in Palu'e, when we examine lexical appearances of the other places we do not find much evidence for such a chain. Although we must identify a series of voiced stops in the language, they are not strongly attested in lexical items derived from Proto-Austronesian sources, and for both *t and *k the regular reflexes are voiceless *t* and *k*, respectively. Equivalents of (7) for the other places of stops are shown in (10) and (11). As we can see, there is no evidence for a general drag chain operating, replacing the PAN voiced alveolar and voiced velar stop phonemes. As we have seen from examining a broader range of lexical items, including those not reflecting PAN forms, the

TABLE 2. INSTANCES OF *d* IN PALU'E WITHOUT (KNOWN) HISTORY EXTERNAL TO FLORES

aʔe-de†	'this'	mudu	'betel pepper'
ʔɓde	'straight'	nadu	'dolphin'
ʔɓadu	'resin'	nodo	'sit, stay'
ʔɓidu	'ancestor stone'	pda	'dye cloth black'
keda	'bent'	rede	'red'
kda	'kick'	salavdo	'fish species'
kedi	'kneel'	vatu pudun	'pile of stones'
kide	'taut'	vida	'flat thing with width'
mada	'pen, cage'	vidi	'goat'
mae kidokada	'whirlpool'	voda	'traditional buffalo-killing ceremony'
mde	'sink'		

† The *-de* in this word is probably etymologically the IPL.INCL.GEN clitic (with *aʔede* possibly meaning something like 'near us', 'our vicinity'), so it is not really an independent attestation.

TABLE 3. INSTANCES OF *g* IN PALU'E WITH A (POSSIBLE) RECONSTRUCTABLE ANCESTRY

PAN		PALU'E	
*teŋkuk	'nape of neck'	toge	'nape'
*bakaw	'mangrove'	vago	'mangrove'
*luGuŋ	(Proto-Philippines)	lgu	'sound of thunder'
*qumu	'roast'	mugu	'burn, cook (TR)' (?)
*(i-a)ku	'I'	-gu	'my'

5. I have also found one instance in which there is a Flores-external cognate: Palu'e *taga* 'necklace' might be cognate with Tukang Besi *toko* 'necklace', which, given that Tukang Besi *k* reflects Proto-Austronesian *k (see [5] above), suggests a voiceless origin for the *g* in the Palu'e lexeme (something like *tVkaw). The existence of extensive trade and cultural contacts between Flores and Southeast Sulawesi, where Tukang Besi is spoken, and the presence of a Tukang Besi community on the north coast of Flores near Palu'e, makes it plausible, though by no means certain, that *taga* might be a loan that entered at an early stage in the history of Palu'e.

functional opposition between voiced and voiceless stops has been reintroduced into the language through new lexical items being added to the language, some by borrowing and some with no obvious provenance.

TABLE 4. INSTANCES OF *g* IN PALU'E WITHOUT HISTORY EXTERNAL TO FLORES

ʔɓaga	'handspan'	n.ge†	'fill'
ʔɓiga	'plate'	naga	'receive, get'
tʃaga	'tell, order'	nogo	'strong'
tʃega	'scratch, broom, sweep'	nugu	'tie strings together'
tʃega	'decorate'	oga	'fall, come down'
tʃga	'plant a stake'	raga	'barbed arrow'
tʃoga	'pour'	ragi	'men's sarong'
tʃogo	'fight (of pigs)'	sago	'with results'
gu	'my, our'	sogo	'buy a ceremonial pig'
kagi	'hang off (someone)'	taga	'chain, necklace'
kegi	'step on tippy toes'	tagalai	'dragonfly'
kgu	'handful'	tagalai	'tomato' ‡
kugu	'part of a loom'	tagu	'ring (bell)'
lego	'ladle'	tgi	'chock full'
legu	'ring (bell)'	tgo	'eat something hard'
lge	'pot'	tgolgo	'tickle'
lige	'mud'	tgu	'poke, stab'
lige	'crush, smash'	toga	'hard, strong'
ligu	'back, flat (of blade)'	toge	'poke, stab (pigs)'
loga	'spouthole'	voga	'flooring'
mgi	'flounder'		

† The full stop between *n* and *g* is used to make clear the fact that this is not a digraph denoting [ŋ], but rather a sequence of two phonemes.

‡ This word is found across a wide part of central Flores (Louise Baird, pers. comm.), and so in all likelihood is not a native Palu'e word.

TABLE 5. INSTANCES OF *g* IN PALU'E WITH A (POSSIBLE) RECONSTRUCTABLE ANCESTRY

PAN		PALU'E	
*teŋkuk	'nape of neck'	toge	'nape'
*bakaw	'mangrove'	vago	'mangrove'
*luGuŋ	(Proto-Philippines)	lgu	'sound of thunder'
*qumu	'roast'	mugu	'burn, cook (TR)' (?)
*(i-a)ku	'I'	-gu	'my'

TABLE 6. SUSPECTED INDONESIAN NC SEQUENCES BORROWED AS C OR $C_{[+voice]}$ IN PALU'E

PALU'E		INDONESIAN/MALAY	
tʃaga	'guard'	jaga	
tʃita	'love'	cinta	
leʔɓa	'carry with stick'	lembar	
muta	'vomit'	muntah	(or PAN *muntaq)
naka	'jackfruit'	naŋka	(or PAN *naŋka)
taga	'catch'	taŋkap	
tuʔɓu	'blunt'	tumpul	

(10) Reflexes of dental/alveolar obstruents and clusters:

	*D >	l, r
*n(t, d) >	d	
*t >	t (> d)	

(11) Reflexes of velar obstruents and clusters:

	*j >	r
*ŋk >	g	
*k >	k (> g)	

What, then, of the voicing in the genitive clitics? In the 1GEN *-gu* and 12GEN *-de* clitics we see voicing of the initial segment in the morpheme, which we have seen is not the place we would expect voicing to first appear. Being enclitic forms—see the examples in (4)— these stops are, of course, only realized intervocalically, and so are not found word-initially, even though they are morpheme-initial. Nonetheless, they are voiced reflexes of PMP voiceless stops that are not otherwise found (except for the *ʔb* reflex of *p). Why are the genitive clitics not realized as **-*ku* and **-*te?* Frequency might play a part in the acquisition of voicing in these two morphemes, but other high-frequency words, such as the free pronoun *aku* '1SG' and *kita* '1PL.IN' similarly reflect voiceless stops in intervocalic position (PMP *i-aku, *i-kita) and do not show voicing (***agu, **kida/**gida*).

4. THE VOICING OF *ku AND *ta. Given that the regular reflexes of *k and *t are *k* and *t*, how do we explain the voiced stop reflexes in *-gu* and *-de?* Ranged against the fact that the regular reflexes are voiceless, we have the following factors favoring, or at least supporting, voiced reflexes:

1. The voiced stops in *-gu* and *-de* are found in intervocalic position.
    - Although the stops are morpheme-initial, the fact that Palu'e does not allow roots with final consonants means that these enclitics will always appear between two vowels, an environment that is the first to show voicing in reflexes of the bilabial *p.
2. A rule of $C_{[-voice]} \longrightarrow C_{[+voice]}$, part of a general process of lenition, shows clear evidence of diffusing through the language. The stop voicing rule is advancing differentially, depending on place (the bilabial stop shows a higher rate of voicing compared to the rest) and phonotactic position (intervocalic stops are more likely to be voiced than initial ones). Contemporary Palu'e is in the early stages of the application of this voicing rule.
    - While voicing has applied most completely to intervocalic bilabial stops, it has also spread to initial bilabial stops, and to the genitive clitics.
3. Voiced stops already existed in intervocalic position prior to the application of a rule voicing etymologically voiceless stops in Austronesian etyma, through their appearance in a variety of loanwords and words of uncertain provenance. This provided a motivation for the voicing rule to extend from lenition involving spirantization to simple voicing.

We can see that while we simply need to take into account the word versus grammatical morpheme status of the lexical root when determining whether or not the voicing rule applies to it, this is secondary to the consideration of whether (nonbilabial) voiced

stops can appear in the word. The appearance of voiced stops, other than the bilabial, directly correlates with the status of those words as being native Palu'e words with or without Proto-Austronesian etymologies. Those words that are not native Palu'e, such as *tadu* 'horn', or that do not appear to be descended from any known Proto-Austronesian source, such as *pda* 'dye cloth black', are the only ones that evidence medial voiced stops.

The changes that are posited here are shown in the form of a series of drag chains, shown in (7), (10), and (11), summed up in (9). We can also reconstruct the phonological history of the stops in Palu'e in terms of system-wide changes. The first (relevant) stage is that reconstructed for Proto–Malayo-Polynesian. Here there are (roughly) voiced stops corresponding to all the voiceless stops except *q.

(12) Stage 1: Proto–Malayo-Polynesian

Voiceless	*p	*t	*k	*q
Voiced	*b	*D	*j	

Moving from this stage to that attested in Palu'e, the first step involves the lenition of the voiced stops (as well as the merger of *l with *D, *Z, and *j). The voiceless stops remain unchanged.

(13) Stage 2: Pre-Palu'e 1, post-lenition

Voiceless	**p	**t	**k	**q
Lenited	**v	**r, **l	**r	

The third stage approaches modern Palu'e closely, and sees the reintroduction of a voicing distinction in the bilabial series. The mechanism behind the acquisition of this new voiced (preglottalized, imploded) bilabial stop has been described in the discussion leading up to (9) in section 3.

(14) Stage 3: Pre-Palu'e 2, post-lenition

Voiceless	p	**t	**k	**q
Preglottalized, Imploded	ʔɓ			
Lenited	v	**r, **l	**r	

Note that modern Palu'e does have a pair of voiced stops, *d* and *g*, as well as a voiced affricate (*ʤ,* corresponding to *ʧ*). The etyma with these stops do not in the main have reconstructable Austronesian etymologies, and so must be considered separately to the list of phonemes deriving from Proto-Austronesian.[6]

(15) Stage 4: Modern Palu'e

Voiceless	p	t	k	ʔ
Voiced preglottalized	ʔɓ	(d)	(g)	
Lenited	v	r, l		
	ʔɓ	(d	g)	

(in roots without Proto-Austronesian etymologies)

6. I have recorded a nonpreglottalized, nonimploded bilabial stop in one word, [ˈabi] '*kesambi* fruit', where it contrasts with ʔɓ in the word [ˈaʔɓi] 'fire'. I suspect that *abi* '*kesambi* fruit' is a loanword from another Flores language, identity as yet unknown. No other instances of [b] have been noted.

Stage 5 is a hypothetical development from modern Palu'e, assuming that the trends that we can see in the language in terms of the introduction of preglottalized, imploded stops continues. This language has now reduced the phonetic contrast between the bilabial voiced stop and the voiced stops produced in other places of articulation, and has regularized the phonologies of inherited and more recent words. While this does not represent the phonological system of Palu'e, it closely resembles the phonologies of many of the closely related languages found south and west of Palu'e in other parts of Flores.

(16) Stage 5: Hypothetical post-Palu'e
(attested in the languages south and west on mainland Flores)

Voiceless	p	t	k	ʔ
Voiced preglottalized	ˀɓ	⁽ⁿ⁾d	⁽ˀ⁾g	
Voiced prenasalized	mb	nd	ŋg	
Lenited	v	r, l		

The only difference between this pattern and that which I have described for Palu'e is the presence of the prenasalized stops. While we do not need to recognize a separate series of prenasalized stops for the main part of the Palu'e lexicon, allowing for nasal codas automatically assimilating in place to a following stop to model roots such as *temba* 'cliff', *rendi* 'bag', or *raŋga* 'finger', there are three domains in which prenasalized stops are frequent in initial position. There are some initial prenasalized stops in general etyma as well, but there is a preponderance of them in onomatopoeic words (loosely defined), in animal species names, and in loan words. The known examples of each case are shown in table 6.

There are two possible motivations here. The loan words can be accounted for by assuming that the phonotactic bans on initial *d* and *g* are being enforced by the prenasalization, while at the same time the use of an unusual phonological pattern marks the word as being a loan word. While licensing an illicit phonotactic template to accommodate loan words is not that unusual, the Palu'e solution seems to have been to extend the range of a word-medial licit consonant sequence, which was only found in separate but adjacent syllables, to a word-initial position where it could only be interpreted as being a complex onset. There is further evidence, other than the general restriction to particular lexical domains, to suppose that the initial prenasalized stops are relatively recent additions to the Palu'e phonological system, and are not yet fully phonemic. While all other consonants can occur as the single onset in a CV lexical root, the prenasalized stops *mb, nd,* and *ŋg* are very rarely found in this environment, and then only *mb* and *ŋg* (the [+grave] consonants) are found, in onomatopoeic roots, listed in table 6. On the other hand, roots of the form m.bV and n.dV, such as /mbu/ [mə̆ˈbu] 'hum', and /ndi/ [nə̆ˈdi] 'carry' are reasonably common, and display an epenthetic vowel typical of sequences of two distinct consonants (compare with /kba/ [kə̆ˈˀɓa] 'grope, feel the way' and /kda/ [kə̆ˈda] 'kick'). Clearly there are at least some ways, at least quantitative, in which the [mb] sequence is treated as a sequence of two phonemes, and not a single prenasalized phoneme.[7]

An alternative account of the voicing of the stop in *-gu* could be based on the fact that some Austronesian languages reflect *-n(a)ku as an alternate for *-ku in the first person singular genitive, derived from the genitive *ni combined with the pronominal clitic (Blust 1977, Ross 2002). This alternative form would then acquire voicing through the regular postnasal voicing described by Blust (1993) for Central Malayo-Polynesian, and well-attested in Palu'e (as seen in [10–11], and table 7). This is a possible account for *-gu,* but does not carry through for the first person inclusive form, as there is no evidence for a *n(i)ta form that could provide a voicing environment. Thus, even if the voicing of *-gu* is due to there having been a nasal in PMP, the

TABLE 7. ATTESTED INITIALLY PRENASALIZED ROOTS IN PALU'E

	PALU'E		
ONOMATOPOEIC	mbo	'squeal sound'	
	mbu	'low thudding sound'	
	ŋga, ŋge	'sound used to drive away fowl'	
	ŋgoe	'crooked, bent, bend'	
	ŋgole	'twisted'	
	ŋgora	'crooked'	
	ŋgu	'hum'	
SPECIES	mbaˀɓambou	'centipede'	
	ero.mbu	'honey bee'	
	mbu	'bird that makes a tapping sound'	Indonesian: *burung puyu*
	kudʒu ndˀɓa	'largest crab (species)'	
	ŋgka	'small bird species'	
	ŋgo	'fish with a yellow tail (species)'	
	ŋgudʒu	'fish (species)'	
	ŋguruŋgero	'small crab (species)'	
LOANWORD	mbuku	'book'	Malay *buku*
	mbsi	'iron'	Malay *bəsi*
	mbako	'tobacco'	Malay *(ta)bakau*
	kandera	'chair'	Malay *kadera*
	ndala	'net, thread'	Malay *jala*
	ndoa	'ugly, cruel, evil'	Malay *do'a* 'debt'
	ndosa	'sin'	Malay *dosa*
	ŋgalu	'necklace'	Malay *kalung*
	ŋgari	'matches'	Malay *garis*
	ŋgta	'sap'	Malay *getah*
GENERAL ETYMA	mbola	'good'	
	mbre	'shut (eyes)'	
	ndero	'play'	
	ŋgedʒe	'cucumber, watermelon'†	
	ŋgeˀu	'baby'	

† This word is a probable loan. Palu'e informants report that neither cucumbers nor watermelons are traditional crops on the island, and grow poorly in the rocky volcanic soil.

7. The fact that *mbu* 'hum' is pronounced as [mə̆ˈbu], and not *[mə̆ˈˀɓu], suggests either that [ˀɓ] has the allophone [b] following a (necessarily homoorganic) nasal if not separated by an underlying vowel (compare with /mabu/ [ˈmaˀɓu] , *[ˈmabu] 'grey'), or that the process of phonemicization of the unit prenasalized phoneme /mb/ has progressed some way, without yet unambiguously producing a new phoneme. See also footnote 6.

voicing of *de* can only be accounted for by contagion from the voicing on the 1sg clitic, a contagion that is limited by the morpholexical specifications of the item. This would represent a weaker claim than that advanced in the main body of this article for phonological change in a morphosyntactically limited environment, because it involves a rather abstract form of paradigmatic regularization: all genitive clitics must employ voiced consonants. Nonetheless, it is still clear that morphosyntactic factors condition the spread of a new, secondary, sound change, that of obstruent voicing without a conditioning preceding nasal.

As an aside, for completeness I should note that the only attested cases of nasal-plus-voiceless stop sequences are found medially, and with the loanwords *kntaŋ* [kə̆nˈtaŋ] 'potato' and *kaŋkoŋ* [ˈkaŋkoŋ] 'swamp grass'. This is not a phonological pattern that has shown any great spread in the language, though the fact that a greater number of (presumably older) loans of words with *nt* or *ŋk* clusters in the original language have been borrowed without the nasal (see table 5 for examples). The fact that there are now some loans that *do* allow the cluster might offer a glimpse of the future direction of change in the language.

5. CONCLUSION: GRAMMATICALLY CONDITIONED SOUND CHANGES. Grammatically conditioned sound change is not unknown. Antilla (1972:80), reports that the final vowels of case endings are lost in Western Finnish, but not vowels that are final in lexical roots. Thus [mult] < **mu-lta* 'from me', but [multa] < **multa* 'dirt, earth'. Clearly the loss of the final vowel is conditioned by the nonlexical status of the vowels in case endings, but this example and the others discussed by Antilla all involve the complete loss of a segment, conditioned by various factors. The data that we have seen from Palu'e show a sound change, $C_{[-voice]} \longrightarrow C_{[+voice]}$, diffusing through the language in a complex way, with the grammatical status of the root to which it applies (bound or free) being part of the specification of its "environment" for the purposes of the application of the sound change. An example of differential deletion of a segment in a grammatically conditioned environment can also be found in the treatment of the PMP 3sg.gen clitic *-ña in Palu'e, where it exceptionally appears as *-n*, not the expected but unattested **-*na*, showing the irregular loss of a final vowel. The changes we have seen in the first person genitive clitics in Palu'e are of a different sort than the Finnish ones, or the third person genitive clitic, and critically they involve changes in the phonological units, and not simply deletion.

Another irregular change can be seen in the treatment of PMP *s in Palu'e. While PMP *s is regularly reflected as *s* in Palu'e in lexemes such as *sau* 'dog' (< *asu), *ʔɓuse* 'navel' (< *pusej), *vose* 'boil (on body)' (< *bisul),[8] and occasionally as *t* or Ø (such as *tusu* 'breast' or *iva* 'nine'), we find that *isa 'one' is reflected most frequently (in terms of the number of times this is produced, or its appearance in daily idioms) as *a*, occasionally as *ha*, and only in the most formal or ritual speech as *sa*. The first pronunciation shows the complete loss of *s; the second displays a highly idiosyncratic [h], a sound that is not attested elsewhere in the language. Only the final, and least frequent, pronunciation shows the regular *s* reflex of *s.

8. With the vowels swapping syllables, just as the consonants of *loke* 'skin' in (5) have swapped syllables while leaving the syllable onsets in place.

A similar case is found for the *=su* '1SG.GEN' clitic in Tukang Besi. This clitic reflects *ku, but it is one of only three instances of an *s* reflex of *k (the other two being from the dative paradigm, *=nso* 'for you' and *=nsami* 'for us', < *nV-ko and *nV-kami, respectively); *k* is the regular reflex of *k, appearing in most other reflexes (occasionally *ʔ* reflects *k), as shown in (17) (from Donohue 1999).[9]

(17) Regular reflexes of *k in Tukang Besi

PMP	TUKANG BESI		PMP	TUKANG BESI	
*kahiw	*kau*	'tree, wood'	*kuku	*kuku*	'fingernail'
*ikuR	*iku*	'tail'	*kulit	*kuli*	'skin'
*ikan	*ika*	'fish'	*likuD	*ta/liku*	'behind'
*kutu	*kutu*	'louse'	*siku	*siku*	'elbow'

Other languages in the area exemplify *k > *s* (such as Muna *sau* 'wood'; see van den Berg 1991), but do not apply the change to the first person singular possessive marker (Muna has *-ku* 'my'), ruling out the possibility of Tukang Besi having borrowed this particular morpheme from another language in which *k > *s* is regular.

Similarly, in modern Indonesian we find irregular sound changes applying to a limited part of the lexicon. The loss of *s with no intermediate step (*h*, for instance) might seem unlikely, but is attested: a similar change, with no intermediate *h*, is currently underway in Indonesian functor words: *sudah* > *uda(h)* 'already', not attestedly via **huda(h)*; *satu* > *atu* 'one', with no attested **hatu*; *saja* > *aja* 'just, only', not heard as **haja*.[10] These are not classically free lexical items, clearly delimiting the spread of the *s > Ø sound change to functors. Words such as *sukun* 'breadfruit' is never heard as **ukun*. On the other hand, *sampai* 'arrive, up to' does appear as *hampai* in some people's speech, indicating that a second sound change is in progress following a different path through the language.

Another example can be found in modern Indonesian, where the regular reflex of *n is as *n*—the nasals have shown little tendency to change in Indonesian, as in most languages. Examples of this can be seen in (18).

(18) Regular reflexes of *n in Indonesian

PMP	INDONESIAN		PMP	INDONESIAN	
*naŋuy	bərə/naŋ	'swim'	*Zalan	jalan	'path, road'
*enem	ənam	'six'	*quZan	hujan	'rain'
*ikan	ikan	'fish'	*anak	anak	'child'
*wanan	k/anan	'right'	*tanaq	tanah	'earth, ground'

Despite this, the nonactive voice prefix in Indonesian is *di-*, reflecting Old Malay **ni-*, itself a development of *-in- (see Cumming [1991:42], Teeuw [1959, 1965], and especially van den Berg [2004] for arguments against treating this prefix being a prefixal development of the 3SG pronoun, as has sometimes been assumed). We may observe the following changes in the inflection of a verb.[11]

9. It is possible that the prefix *sa-*, which marks a temporal adverbial clause ('after Xing, …), is related to the commonly attested form *ka-* (as in the Tagalog recent past marker). If so, this would be a second instance of *k > *s* (thanks to Dan Kaufman for discussion leading to this observation).

10. This sound change has probably been inspired by an identical sound change, with an identical restriction, in Jakarta Malay.

(19) The *n > d change in Malay

PMP	*tanem 'plant'	*t-in-anem 'planted'
Old Malay (Talang Tuwo)		ni-tanem
Modern Malay		di-tanam

For one final example of grammatically conditioned sound change, we can note that Bali-Vitu (Austronesian, Oceanic) shows the unexpected retention of Proto-Oceanic final consonants, but only for inalienable nouns (van den Berg, pers. comm.). Here, too, a deletion rule is blocked from applying when a particular grammatical condition is met.[12]

The evidence supports the conclusion, hardly surprising, that irregular change is the regular way for a language to develop historically. Furthermore, in addition to lexical irregularity, there is good evidence that sound changes are often limited by a syntactic category division. In other words, the phonological history of a language depends as much on morphosyntactic information as it does on what is usually considered to be "closer" to the sounds themselves, things such as phonotactic constraints, (phonologically) conditioning environments, or changes in related sounds (drag chains, push chains). But this cannot be the whole story; in a way, the syntactic category differences that we can see here are possibly epiphenomenal, relying on the fact that the bound pronominal suffixes are very high frequency items, and so are the morphemes most likely to participate in any innovative phonological processes (see Bybee 2001 for extensive discussion of the role of frequency and phonological change).

REFERENCES

Antilla, Raimo. 1972. *An introduction to historical and comparative linguistics*. New York: Macmillan.

Blust, Robert. 1977. The Proto-Austronesian pronouns and Austronesian subgrouping: A preliminary report. *Working Papers in Linguistics* 9(2): 1–15. Honolulu: Department of Linguistics, University of Hawai'i.

Bybee, Joan. 2001. *Phonology and language use.* Cambridge: Cambridge University Press.

Cumming, Susanna. 1991. *Functional change: The case of Malay constituent order*. Berlin: Mouton de Gruyter.

Donohue, Mark. 1999. *A grammar of Tukang Besi*. Berlin: Mouton de Gruyter. Grammar Library Series No. 20.

———. 2005. The pretenders to the Muna-Buton group. In *Papers in Austronesian dialectology and subgrouping*, ed. by John Bowden and Nikolaus Himmelmann, 21–35. No. 563. Canberra: Pacific Linguistics.

Ross, Malcolm. 1992. The sound of Proto-Austronesian: An outsider's view. *Oceanic Linguistics* 31:23–64.

Teeuw, A. 1959. The history of the Malay language: A preliminary survey. *Bijdragen tot de Taal-, Land- en Volkendunde* 115:138–61.

———. 1965. Old Balinese and comparative Indonesian linguistics. *Lingua* 14:271–84.

van den Berg, René. 1989. *A grammar of the Muna language.* KITLV, Verhandelingen 139. Dordrecht/Providence: Foris.

11. Similar denasalization is found with **meR-*, appearing as *ber-* in Malay. Again, the regular reflex of *m is *m,* seen in the form for 'six' in (18), and again the grammatical environment is the only conditioning factor behind the different treatment.

12. Modeling this in terms of rerankings of constraints would be difficult.

———. 1991. Muna historical phonology. In *Studies in Sulawesi linguistics,* pt. 2, ed. by J. N. Sneddon, 2–28. NUSA 33.
———. 2002. The place of Tukang Besi and the Muna-Buton languages. Paper presented at the Ninth International Conference on Austronesian Linguistics. Canberra, Australia. January 11.
———. 2004. Some notes on the origin of Malay *di-*. *Bijdragen tot de Koninklijk Instituut voor Taal-, Land- en Volkenkunde* 160(4): 532–54.
Wurm, S. A., and Basil Wilson. 1975. *English finderlist of reconstructions in Austronesian languages (post-Brandstetter)*. Series C-33. Canberra: Pacific Linguistics.

mark@donohue.cc