Novel Avenues to the Study of Language Structure

Novel Avenues to the Study of Language Structure

Edited by

Mihaela Tănase-Dogaru, Alina Tigău,
Ioana Stoicescu and Ruxandra Drăgan

Cambridge
Scholars
Publishing

Novel Avenues to the Study of Language Structure

Edited by Mihaela Tănase-Dogaru, Alina Tigău, Ioana Stoicescu
and Ruxandra Drăgan

This book first published 2025

Cambridge Scholars Publishing

Lady Stephenson Library, Newcastle upon Tyne, NE6 2PA, UK

British Library Cataloguing in Publication Data
A catalogue record for this book is available from the British Library

TABLE OF CONTENTS

CHAPTER ONE

SYLLABLE RESTRUCTURING IN BISLAMA

ANDREI A. AVRAM

The paper looks at syllable restructuring in the English-lexifier pidg-
in/creole Bislama, spoken in Vanuatu. The empirical evidence from Early
Bislama covers a period ranging from 1867 to 1935 and includes trave-
logues, memoirs, letters, descriptions of the New Hebrides, and a play.
The sources for the modern variety are: two corpora of spoken Bislama;
two corpora of written Bislama; dictionaries; a reference grammar; text-
books and workbooks of Bislama; textbooks in Bislama; religious texts. It
is first shown that Early Bislama tends to disallow complex syllable mar-
gins and employs three types of vowel epenthesis as well as consonant
deletion as repair strategies for the resolution of etymological onset and
coda clusters. Also, it is argued that the epenthesized vowels are lexical in
Early Bislama. Forms attested in spoken and written Modern Bislama,
however, reflect the interplay of several factors: fossilization, relaxation of
phonotactic constraints on onsets, and anglicization. Repair strategies are
mostly restricted to the resolution of illicit coda clusters. Also, the status of
the epenthesized vowels in Modern Bislama has been affected, as in most
reflexes of etymological complex onsets these are no longer lexical.

1 Introduction

Bislama is an English-lexifier language spoken in the Republic of Vanua-
tu, formerly known as the New Hebrides Condominium. Today it is the
national language of the country (Crowley 1990). For most of its users,
Bislama is a second language, i.e. it functions as an expanded pidgin.
However, according to Crowley (2008a, 145), "it has gradually been ac-
quiring small numbers of first-language speakers", up to "possibly ten
percent of the population", i.e. it has been creolized.

Historically, Bislama is an offshoot of Melanesian Pidgin English, whose two other modern descendants are Tok Pisin, spoken in Papua New Guinea, and Pijin, spoken in the Solomon Islands. Bislama emerged in the southern islands of Vanuatu and the Loyalty Islands of New Caledonia by the 1860s (Crowley 1990, 60–5). It became stabilized in the first decades of the 20th century (Tryon and Charpentier 2004, 1908).

Vanuatu is a highly multilingual nation, with some 80 languages spoken by its population of approximately 200,000, having "the most complex linguistic demography of any country in the world in terms of languages per head of population" (Crowley 2008a, 144). The local vernaculars, i.e. the substrate languages of Bislama, belong to the Southern Oceanic linkage of the Oceanic branch of the Austronesian languages.

Generally, as mentioned by Lynch *et al*. (2011, 34), Oceanic languages "are frequently phonologically less complex than those of many other linguistic groupings in the world" and their "syllable structures tend to approximate a simple CV type". According to Lynch (1998, 84), "probably the majority of Austronesian languages, however, allow both open and closed syllables (syllables ending in a consonant)", but "there are few consonant clusters, and they mainly occur across morpheme boundaries".

With respect to the substrate languages, Charpentier (1979, 185) states that "dans la plupart des langues mélanésiennes des Nouvelles Hébrides, les lexèmes se présentent sous la forme C̲onsonne + V̲oyelle + C_ + V_". Similarly, Tryon (1987, 9) writes that "there is a restriction which is common in Polynesian languages, namely that sequences of two or more consonants within the same syllable cannot occur together". Tryon (1987, 9) further notes that "in many Vanuatu languages these sequences of consonants must be separated by a vowel". Under the circumstances, the substrate languages are expected to exert considerable influence on the syllable structure of Bislama, *a fortiori* in its earlier stages.

2 Early Bislama

2.1 Corpus

The corpus of Early Bislama consists of pre-World War II records, covering a period ranging from 1867 to 1935. The textual evidence includes travelogues (Giles 1968 [1877], Wawn 1893, Grimshaw 1907, Speiser 1913, Johnson 1922, Titaÿna and Lugeon 1931), memoirs (Paton 1894, Watt 1896, Alexander 1927, Marshall 1937), letters (Fletcher 1923, 1924), descriptions of the New Hebrides (Kay 1872, Imhaus 1890, Baker 1929, Harrisson 1937), a play (Jacomb 1929), grammatical and lexical sketches

of Early Bislama (Pionnier 1913, Jacomb 1914, 90–104), and the samples of Early Bislama in Tryon and Charpentier (2004).

As noted by Crowley (1998, 63–4), "actual [Early] Bislama attestations appear in a variety of guises". These include "straightforward statements about how to say a particular thing in Bislama", e.g. in Pionnier (1913) and Jacomb (1914, 90–104), "quoting the words of ni-Vanuatu speaking Bislama, or of Europeans addressing ni-Vanuatu", and "lexical items, or perhaps phrases, embedded within a sentence in English". Moreover, many of the earlier records of Bislama are often rendered in an anglicized version, as in most early records of English-lexifier pidgins and creoles. The forms collected and transcribed by Père Pionnier, a French missionary, are an important exception in this respect. According to Crowley (1993, 210), "there is no indication that he had any speaking knowledge of English at all". Père Pionnier or "the person who actually prepared the article for publication, knew some written English" (Crowley 1993, 218), but "the level of English was clearly not high, as there are mistakes" (Crowley 1993, 224, n. 7). On the whole, therefore, "normalizing", i.e. "etymologizing" transcriptions are less likely in this case.

All examples in this section are reproduced in the orthography used in the sources. The entries include the date of the attestation, the English gloss, and the reference; in a few cases less transparent etyma are also indicated. The dates of the attestations are those mentioned in the primary sources or in Crowley (1998) and Tryon and Charpentier (2004). When an exact year could not be established, a hyphen preceding it is used, which reads "in or before".

2.2 Onsets

Consider first the fate of reflexes of etymological obstruent + liquid clusters. The evidence from early records of Bislama shows that these were disallowed. Three subtypes of vowel epenthesis are employed as repair strategies for their resolution. The examples under (1)–(3) illustrate the insertion of [e], as a default epenthetic vowel:

(1) /pr-/
    copperah 'copra' -1919 (Fletcher 1923, 326)
(2) /dr-/
    a. *derrown* 'to drown' 1899 (Pionnier 1913, 193)
    b. *derronk* 'drunk' -1919 (Fletcher 1923, 330)
    c. *derrink* 'to drink' -1919 (Fletcher 1924, 121)

(3) /gl-/
    Ingerlish 'English' -1919 (Fletcher 1923, 329)

A second repair strategy is vowel copying. As seen in the examples below, the epenthetic vowel is a copy of the etymological vowel in the following syllable:

(4) /gr-/
    a. *guirisse* [giris] 'fat' 1899 (Pionnier 1913, 116)
    b. *angérè* 'hungry' 1899 (Pionnier 1913, 119)

Labial attraction is also attested, as in (5), whereby the [LABIAL] consonant /f/ enforces the insertion of the [LABIAL] vowel [o] as the epenthetic vowel:

(5) /fr-/
    Forailleray 'Friday' 1899 (Pionnier 1913, 193)

    Obstruent + /w/ clusters appear to have been permitted, as shown by the following examples:

(6) /tw-/
    touaneté 'twenty' 1899 (Pionnier 1913, 187)
(7) /kw-/
    a. *quick* 'quick' 1914 (Jacomb 1914, 94)
    b. *quinine* 'quinine' -1919 (Fletcher 1924, 154)

Consider next reflexes of etymological /s/-initial clusters. Two of these are illicit and undergo reduction via vowel epenthesis. In /s/ + oral stop clusters the epenthetic vowel is occasionally default [i], as in (8):

(8) /st-/
    sitone 'stone, rock' 1931 (Titaÿna and Lugeon 1931, 31)

In other cases, the epenthetic vowel is enforced via vowel copying, as in (9)–(11):

(9) /sp-/
    soupoune [spun] 'spoon' 1899 (Pionnier 1913, 115)
(10) /st-/
    a. *sitil* 'to steal' 1899 (Pionnier 1913, 185)
    b. *sitima* 'steamer' 1899 (Pionnier 1913, 115)

(11) /sk-/
    a. *sikine* 'skin; body' 1899 (Pionnier 1913, 193)
    b. *sikinime* [sikinim] 'to peel' 1899 (Pionnier 1913, 191)

Another /s-/-initial cluster which undergoes reduction is /s/ + nasal stop. The only repair strategy attested is vowel copying:

(12) /sm-/
    sémèle 1899 (Pionnier 1913, 191)

The available evidence from early records of the language suggests that the /sl-/ is permitted (but see also section 3.2). As shown below, this cluster appears to be preserved as such:

(13) /sl-/
    a. *slipe* 'to sleep' 1899 (Pionnier 1913, 191)
    b. *sleep* 'to sleep' 1914 (Jacomb 1914, 94)
    c. *sleep* 'to sleep' -1919 (Fletcher 1924, 121)
    b. *slack im* 'to strike matches' 1914 (Jacomb 1929, 30)

Also permitted is the cluster /sw-/. Consider the following examples:

(14) /sw-/
    a. *swim* 'to bathe' 1899 (Pionnier 1913, 191)
    b. *sweet* 'sweet' 1914 (Jacomb 1914, 98)
    c. *sweat* 'to sweat' -1919 (Fletcher 1924, 154)
    d. *swim* 'to swim' 1934 (Marshall 1937, 71)

In light of the data examined in this section, it may be concluded that not all types of complex onset undergo reduction in Early Bislama. The only repair strategy employed for the resolution of illicit onset clusters is vowel epenthesis. Note that vowel epenthesis reduces both clusters that do not violate the Sonority Sequencing Generalization, i.e. obstruent + liquid and /s/ + nasal stop clusters, and those that do, i.e. /s/ + oral stop clusters.

2.3 Codas

Several Early Bislama forms etymologically derived from words with simplex codas appear, at first sight, to exhibit paragogic vowels. As shown below, however, there is evidence which suggests that these forms should not be taken at face value.

Consider first the following reflexes of etyma ending in the voiceless stop /p/:

(15) /-p/
    a. *supa* 'soup' 1899 (Pionnier 1913, 116)
    b. *seepy* 'sheep' 1914 (Jacomb 1914, 101)

The form *supa* is not attested either in Melanesian Pidgin English or in any other source for Early Bislama. Also, one of the Modern Bislama forms for 'soup' is *supsup* (Guy 1974, 203, Crowley 2003, 266)*,* without an intrusive vowel. An identical form is also attested in Pijin (Jourdan and Maebiru 2002, 233). This suggests that *supsup* would have existed in Early Bislama as well. If so, the non-reduplicated forms *sup* < English *soup* (Crowley 2003, 266) and *lasup* < French *la soupe* (Guy 1974, 203) attested in Modern Bislama are later developments. Moreover, the non-etymological word-final vowel in *supa* is [a], i.e. neither a vowel copy, nor a [LABIAL], vowel, nor default [i]/[e]. It may therefore be assumed that the word-final <a> in the form *supa* 'soup' in Pionnier (1913) is an error of transcription and does not represent a paragogic vowel. A similar case can be made against *seepy*. The fact that the Modern Bislama, Tok Pisin and Pijin forms are *sipsip* (see Guy 1974, 153 and Crowley 2003, 383 for Modern Bislama, Volker *et al*. 2008, 90 for Tok Pisin, and Jourdan and Maebiru 2002, 211 for Pijin) shows that, as noted by Crowley (1998, 96), "the reduplicated form had early currency" and that the form *sipi*, with an intrusive vowel copy, recorded by Jacomb (1914) "represented a temporary development".

Consider next the forms under (16), which are all derived from etyma ending in /-ŋ/:

(16) a. *alonga* 'in' 1878 (Wawn 1893, 144)
    b. *alonga* 'in' 1880s (Thomas 1886, 359)
    c. *along-a* 'in' 1890 (Tryon and Charpentier 2004, 234)
    d. *longa* 'in' 1877 (Giles 1968 [1877], 37)
    e. *longa* 'to' -1919 (Fletcher 1923, 327), 'with' (Fletcher 1923, 325), 'because' (Fletcher 1923, 329)
    f. *belong-a* 'of' 1869 (Watt 1896, 369)
    g. *belong-a* 'of' 1870 (Kay 1872, 79)
    h. *belonga* 'of' 1878 (Wawn 1893, 143)
    i. *belong a* 'of' 1880s (Thomas 1886, 246)
    j. *belong-a* 'of' 1890 (Tryon and Charpentier 2004, 234)

All these forms are widely attested in late 19th-century varieties of Melanesian Pidgin English (see Tryon and Charpentier 2004, and Avram 2005, 200). Moreover, the intrusive word-final vowel is once again [a], and not a vowel copy or default [i]/[e]. To conclude, these forms appear to be illustrative of the pre-stabilization stage of Bislama.

One last example, derived from an etymon ending in /-l/, is reproduced below:

(17) /-l/
　　killa 'to hit' -1919 (Fletcher 1923, 326)

This form should also be dismissed as doubtful. Firstly, this is the only occurrence of such a reflex of English *kill*. Secondly, this form does not contain the expected transitive suffix *-im*. Thirdly, the expected intrusive vowel would have been either a vowel copy or default [e].

Summing up, word-final simplex codas are permitted in Early Bislama.

Consider next the fate of etymological complex codas. While these are disallowed in Early Bislama, their composition determines the way in which they are adjusted.

As shown below, reflexes of /k/ + /s/ (18) and /g/ + /z/ (19) coda clusters are resolved via the epenthesis of the default vowel [i] or [e]:

(18) /-ks/
　　a. *bokis* 'box' 1867 (Paton 1894, 77)
　　b. *sikis* 'six' 1899 (Pionnier 1913, 187)
　　c. *ackis* 'axe' 1912 (Alexander 1927, 214)
　　d. *ackus* 'axe' 1917 (Johnson 1921, 48)
　　e. *bokkus* 'box' (Johnson 1921, 170)
(19) /-gz/
　　en'guis 'egg' 1899 (Pionnier 1913, 116) < E *eggs*

Two remarks are in order here. The spelling <-us>, as in (18d) and (18e) presumably represents [-es]. As for <n'g> in (19), this presumably stands for the pre-nasalized stop [ᵑg], as a reflex of etymological /g/. Note that intervocalic pre-nasalized stops still occur in Modern Bislama (Tryon 1987, 6).

Early Bislama reflexes of etymological nasal stop + /θ/ (20) or nasal stop + /s/ (21) coda clusters exhibit the default epenthetic vowel [i]:

(20) /-nθ/
　　maniche 'month' 1899 (Pionnier 1913, 112) < E *month*

(21) /-ns/
    a. *baniche* 'fence' 1899 (Pionnier 1913, 117) < E *fence*
    b. *banis* 'fence' 1912 (Alexander 1927, 214)
    c. *fenys* 'fence' 1914 (Jacomb 1914, 99)
    d. *danis* 'dance' 1934 (Marshall 1937, 88)

Note that Pionnier (1913) has [-ʃ] as the Early Bislama reflex of both /-θ/ and /-s/, as in (20) and (21a), respectively. The latter, however, surfaces as [-s] in other sources, as in (21b-d). This most probably reflects inter-speaker variation, given that the same holds for Modern Bislama, in which "/s/ is normally realized as [s] in the Bislama of most speakers" and "only occasionally as sh" (Tryon 1987, 6).

    Etymological nasal stop + affricate coda clusters are also resolved via epenthesis of the default vowel [i]:

(22) /-ntʃ/
    a. *lanish* 'speedboat' 1914 (Jacomb 1914, 96) < E *launch*
    b. *Frennich* 'French' -1919 (Fletcher 1923, 328)
    c. *lannitch* 'speedboat' -1919 (Fletcher 1923, 325)
    d. *lanich* 'speedboat' 1931 (Titaÿna and Lugeon 1931, 31)
(23) /-ndʒ/
    oranige 'orange' 1899 (Pionnier 1913, 116)

Here again a few comments are in order. As can be seen, the Early Bislama reflexes of English /-tʃ/ are either [ʃ], as in (22a) and (22d), or [tʃ], as in (22b)–(22c). Note, however, that the latter forms are produced by a "white man" and are therefore not necessarily representative. As for English /-dʒ/, Pionnier's (1913) transcription, based on French orthography, suggests that its Early Bislama reflex is apparently [ʒ], as in (23). There are two reasons, however, which suggest that this is rather unlikely. In Modern Bislama, as shown by Crowley (2004, 11) "the contrast between voiced and voiceless segments is lost word-finally […] with only voiceless segments being found". In all likelihood, the same must have been true of the earlier stages of the language. Also, Pionnier (1913, 117) himself has *ca-biche* 'cabbage', where etymological /dʒ/ corresponds to Early Bislama [ʃ], as the spelling with <che> indicates.

    A relatively large number of forms are relevant to the adjustment of etymological word-final codas of the type consonant + /t/ or /d/, in which the consonant is the fricative /s/, the nasal /n/ or the liquid /l/. The forms at issue consist of reflexes of /-st/ (24), /-nt/ (25), /-nd/ (26), /-lt/ (27), and

/-ld/ (28). As can be seen in the examples below, all these illicit coda clus-
ters are resolved via deletion of /t/ or /d/:

(24) /-st/
      a. *firs'* 'first' 1907 (Grimshaw 1907, 240)
      b. *fas'* 'stuck' 1914 (Jacomb 1914, 97)
(25) /-nt/
      a. *wane* 'to want' 1899 (Pionnier 1913, 192)
      b. *kaun'* 'debt' 1914 (Jacomb 1914, 100) < E *account*
      c. *pine* 'point' 1914 (Jacomb 1914, 103) < E *point*
(26) a. /-nd/
      b. *graoun* 'ground' 1899 (Pionnier 1913, 193)
      c. *sane* 'sand' 1899 (Pionnier 1913, 110)
      d. *ouine* 'wind' 1899 (Pionnier 1913, 110)
      e. *behin'* 'behind' 1907 (Grimshaw 1907, 240)
(27) /-lt/
      saul 'salt' 1899 (Pionnier 1913: 116) < E *salt*
(28) /-ld/
      a. *col* 'cold' 1899 (Pionnier 1913, 112)
      b. *hole* 'hold (in ship)' 1912 (Alexander 1927, 214)

A single form in the corpus illustrates the adjustment of the etymological
/l/ + /k/ coda cluster:

(29) /-lk/
      milik 'milk' 1914 (Jacomb 1914, 103)

To conclude, complex word-final codas are prohibited in Early Bislama.
Even consonant clusters in which sonority decreases away from the nucle-
us, i.e. which respect sonority requirements for codas, undergo reduction.
Two repair strategies are employed for reduction, depending on the type of
consonant cluster: vowel epenthesis and consonant deletion.
    Consider finally word-medial codas. Only four forms relevant to their
treatment are found in the available records of Early Bislama. Three of
these are reproduced below:

(30) /-t/
      taketer 'doctor' 1914 (Jacomb 1914, 103)
      dokkitor 'doctor' -1919 (Fletcher 1923, 333)
(31) /-l/
      allersame 'as, like' 1912 (Alexander 1927, 214)

Since Early Bislama seems to have always permitted simple codas, as already shown, the forms in (30)-(31) are isolated cases. In fact, they appear to have been inherited from Melanesian Pidgin English. According to Crowley (1998, 72), *dokita* "is attested as an occasional archaic variant in modern Bislama". Crowley (1998, 72) further writes that this "is also the shape of the word that was borrowed in the Loyalties", where Bislama was still spoken at the beginning of the 20th century (Tryon and Charpentier 2004, 205). As for Early Bislama *allersame*, a similar form is found in Early Pijin: *alle same* (Avram 2007, 228). It is only the fourth form which is truly relevant to the adjustment of etymological word-medial codas:

(32) *saulouara* 'sea' 1899 (Pionnier 1913, 193) < E *salt water*

As expected, the illicit word-medial cluster /-lt/ undergoes reduction via deletion of /t/.

The data examined in this section show that word-medial simplex codas are permitted in Early Bislama, whereas word-medial complex codas are prohibited and are reduced just as their word-final counterparts.

3 Modern Bislama

3.1 Corpus

The sources for Modern Bislama consists of: two corpora of spoken Bislama (Bislama n.d., Bislama_transcripts); two corpora of written Bislama (Charpentier 1979, Jarraud-Leblanc 2013); dictionaries (Camden 1977, Crowley 2003, Moon 2007); a grammar (Crowley 2004); textbooks and workbooks of Bislama (Guy 1974, Tryon 1987, Anon. a 2014); textbooks in Bislama (Anon. b 2017, Anon. c 2017); religious texts (*Gud Nyus bilong Jisas Krais...* 1972, *Nyutestamen long Bislama ...* 1980, *Ol sam long Bislama* 1980).

All examples in this section are reproduced in the standardized orthography of written Bislama (see e.g. Moon 2007).

3.2 Onsets

The adjustment of onset clusters in Modern Bislama has been only briefly discussed in the literature. Tryon (1987, 9), for instance, writes that "speakers […] normally select as the vowel to be inserted between consonant sequences the same vowel as follows the sequence". According to Crowley (2003, 12), "a vowel that is the same as the vowel of an adjacent

syllable (but sometimes just /i/) can be inserted between two consonants that come together when one is /l/, /r/ or /s/'". In other words, the epenthetic vowel is either a vowel copy or the default vowel [i]. Crowley (2008a, 158) states that "initial […] clusters are much less likely to undergo reduction" and that "initial two-member clusters may be simplified by the optional insertion of an epenthetic vowel".

    In what follows, as in section 2.2, the treatment of etymological complex onsets is analyzed with reference to the particular type of onset cluster. Moreover, the variation mentioned by Tryon (1987, 9) and Crowley (2008a, 158) is illustrated with forms from the two corpora of spoken Modern Bislama (Bislama n.d., Bislama_transcripts).

    Consider first reflexes of obstruent + liquid clusters. Examples (32)–(35) from Charpentier (1979, 186, 276), Tryon (1987, 9), Crowley (2003, 12), Crowley (2004, 23), Jarraud-Leblanc (2013, 72, 84, 96, 208) illustrate the co-existence of variants exhibiting vowel copy epenthesis and non-epenthesized ones:

(33) /bl-/
    blu ~ bulu 'blue'
(34) /br-/
    brum ~ burum 'broom'
(35) /fr-/
    from ~ forom 'from'

The insertion of [i], as a default epenthetic vowel is only attested in religious texts (*Gud Nyus bilong Jisas Krais...* 1972, *Nyutestamen long Bislama ...* 1980, *Ol sam long Bislama* 1980).

(36) /kl-/
    klok ~ kilok 'clock'

Finally, labial attraction is not found.

    While epenthesis into this type of onset cluster is certainly attested, it is important to note that there is considerable variation. Note that in all tables below the non-epenthesized vs. epenthesized forms are set out per number of occurrences and per number of documents, the latter indicated between brackets. As shown in Table 1, some forms exhibit variation of non-epenthesized and epenthesized variants:

Table 1-1 Frequency of variable reflexes of obstruent + liquid onset clusters in spoken Bislama

Form	Bislama (n.d.)		Bislama transcripts (n.d.)	
	Non-epenthesized	Epenthesized	Non-epenthesized	Epenthesized
ples ~ peles 'place'	*ples* 6935 (1642)	*peles* 1 (1)	*ples* 1196 (67)	*peles* 15 (2)
blue ~ bulu 'blue'	*blue* 154 (67)	*bulu* 2 (2)	*blu* 22 (9)	*bulu* 0
fri ~ firi 'free'	*fri* 271 (149)	*firi* 1 (1)	—	—
from ~ forom 'from'	*from* 7541 (2055)	*forom* 365 (26)	*from* 868 (68)	*forom* 0
turu ~ turu 'true'	*tru* 2083 (898)	*turu* 1 (1)	*tru* 247 (53)	*turu* 2 (2)
klinim ~ kilinim 'to clean'	*klinim* 2 (2)	*kilinim* 1 (1)	*klinim* 228 (109)	*kilinim* 1 (1)
klosap ~ kolosap 'near, close'	*klosap* 1089 (557)	*kolosap* 223 (172)	*klosap* 113 (32)	*kolosap* 96 (33)

As can be seen, the epenthesized variants are in the minority. This already suggests that epenthesis into obstruent + liquid onset clusters is becoming less frequent in Modern Bislama. This is confirmed by the fact that other words only occur in a non-epenthesized form, as illustrated in Table 2 below:

Table 1-2 Frequency of non-variable reflexes of obstruent + liquid onset clusters in spoken Bislama

Form	Bislama (n.d.)		Bislama_transcripts (n.d.)	
	Non-epenthesized	Epenthesized	Non-epenthesized	Epenthesized
brum 'broom'	*brum* 27 (11)	0	*brum* 7 (1)	0
fraede 'Friday'	*fraede* 44 (29)	0	*fraede* 5 (5)	0
klia 'clear'	*klia* 520 (339)	0	*klia* 44 (19)	0
Inglis 'English'	*Inglis* 361 (170)	0	*Inglis* 75 (13)	0
gris 'grease'	*gris* 30 (18)	0	—	0
kopra 'copra'	*kopra* 54 (33)	0	*kopra* 10 (8)	0
hangri 'hungry'	*hangri* 10 (8)	0	*hangri* 3 (2)	0

As shown below, obstruent + /w/ onset clusters are permitted:

(37) /tw-/
    twin 'twin'
(38) /kw-/
    kwaet 'quiet'

Consider next reflexes of /s/-initial clusters. According to Guy (1974, 7), "in words starting with an s followed by one or two consonants, a vowel may be inserted between that s and the consonant which follows it", i.e. there is variation between non-epenthesized and epenthesized forms. As for the quality of the epenthetic vowel, Guy (1974, 7) writes that it "is a u whenever the next vowel is u; otherwise it is i". That is to say, [i] functions as the default epenthetic vowel. Insertion of [i] is illustrated by the following examples of reflexes of /s/ + oral stop clusters (examples from Guy 1974, 7, Tryon 1987, 9, Jarraud-Leblanc 2013, 96):

(39) /sp-/
    spolem ~ sipolem 'to damage'
(40) /str-/
    strap ~ sitrap 'strap'
(41) /sk(r)-/
    a. *skel ~ sikel* 'scales'
    b. *skul ~ sikul* 'school',
    c. *skrasem ~ sikrasem* 'to scratch'

Note, however, that (41b) exhibits epenthetic [i], even though the next vowel is /u/, *contra* Guy (1974, 7). As for vowel copying, it is illustrated by forms such as those reproduced below:

(42) /st-/
    stima ~ sitima 'vessel, ship' < E *steamer*
(43) /sk-/
    a. *skin ~ sikin* 'skin, bark, hide'
    b. *skul ~ sukul* 'school'

The data set out in Table 3 allow a more precise assessment of the extent of variation in reflexes of /s/ + oral stop onset clusters:

Table 1-3 Frequency of variable reflexes of /s/ + oral stop onset clusters in spoken Bislama

Form	Bislama (n.d.)		Bislama transcripts (n.d.)	
	Non-epenthesized	Epenthesized	Non-epenthesized	Epenthesized
spia ~ sipia 'spear'	*spia* 91 (44)	*sipia* 0	*spia* 69 (7)	*sipia* 2 (1)
ston ~ siton 'stone'	*ston* 855 (251)	*siton* 0	*ston* 216 (35)	*siton* 2 (1)
stoa ~ sitoa 'store'	*stoa* 123 (70)	*sitoa* 3 (2)	*stoa* 7 (3)	*sitoa* 0
stap ~ sitap 'to stop'	*stap* 22865 (3250)	*sitap* 1 (1)	*stap* 3481 (71)	*sitap* 0
skin ~ sikin 'skin, bark, hide'	*skin* 231 (114)	*sikin* 0	*skin* 20 (10)	*sikin* 24 (9)
skul ~ sikul 'school'	*skul* 1254 (397)	*sikul* 0 *sukul* 0	*skul* 187 (30)	*sikul* 11 (2) *sukul* 0
skrasem ~ sikrasem 'to scratch'	*skrasem* 14 (7)	*sikrasem* 3 (2)	*skrasem* 3 (3)	*sikrasem* 15 (5)

Additional evidence of variation comes from the process of morphological partial reduplication. As shown by Crowley (1990, 310), Meyerhoff (2003, 232-234), Crowley (2004, 26, 2008b, 451), two types of morphological partial reduplication are attested in Modern Bislama. In CV reduplication (44a), the first consonant and the following vowel of the base are reduplicated. In (C)VC reduplication (44b), the first consonant, the next vowel and the following consonant of the base are reduplicated. A base can undergo both types of reduplication:

(44) a. CV reduplication
       si-sikinim 'to peel'
     b. (C)VC reduplication
       sik-sikinim 'to peel'

As can be seen, both types of reduplication presuppose a base *sikin*, i.e. with epenthetic [i].

Once again, while variation certainly occurs, data collected more recently include non-epenthesized reflexes of /s/ + oral stop onset clusters, some of which–*spolem* 'to spoil' and *stima* 'vessel ship' previously had epenthesized variants as well–see examples (39) and (42). Consider the forms in Table 4 below:

Table 1-4 Frequency of non-variable reflexes of /s/ + oral stop onset clusters in spoken Bislama

Form	Bislama (n.d.)		Bislama transcripts (n.d.)	
	Non-epenthesized	Epenthesized	Non-epenthesized	Epenthesized
spolem 'to spoil'	*spolem* 478 (268)	0	*spolem* 43 (18)	0
spun 'spoon'	*spun* 22 (7)	0	*spun* 4 (1)	0
stima 'vessel, ship'	*stima* 17 (13)	0	*stima* 13 (5)	0
stilim 'to steal'	*stilim* 96 (58)	0	*stilim* 29 (9)	0

If epenthesized, reflexes of /s/ + nasal stop clusters only exhibit insertion of [i] or [e] as default epenthetic vowels:

(45) /sm-/
    small ~ simol / semol 'small'
(46) /sn-/
    snek ~ sinek 'snake'

The frequency of occurrence of such epenthesized variants is extremely low, as shown in Table 5:

Table 1-5 Frequency of variable reflexes of /s/ + nasal stop onset clusters in spoken Bislama

Form	Bislama (n.d.)		Bislama transcripts (n.d.)	
	Non-epenthesized	Epenthesized	Non-epenthesized	Epenthesized
smol ~ simol / semol 'small'	*smol* 2716 (1066)	*simol* 1 (1) *semol* 1 (1)	*smol* 610 (64)	*simol* 5 (2) *semol* 0
snek ~ sinek 'snake'	*snek* 141 (32)	*sinek* 0	*snek* 63 (8)	*sinek* 1 (1)

However, here again there is additional evidence of variation from morphological partial reduplication. In the forms below (from Crowley 1990, 310), illustrative of CV reduplication and (C)VC reduplication, respectively, both types of reduplication presuppose a base *simol*, i.e. with epenthetic [i]:

(47) a. CV reduplication
      si-simol 'very little'
   b. (C)VC reduplication
      sim-simol 'very little'

Reflexes of /sl-/ exhibit the occasional insertion of [i] as a vowel copy:

(48) /sl-/
      slip ~ silip 'sleep'

However, as shown in Table 6, the frequency of occurrence of epenthe-sized variants is low:

Table 1-6 Frequency of reflexes of /sl-/ onset clusters in spoken Bislama

Form	Bislama (n.d.)		Bislama transcripts (n.d.)	
	Non-epenthesized	Epenthesized	Non-epenthesized	Epenthesized
slip 'sleep' *slakem* 'to strike match-es'	*slip* 383 (201) *slakem* 15 (10)	*silip* 86 (42) 0	*slip* 83 (24) *slakem* 3 (2)	*silip* 69 (24) 0

That variation does exist is also shown by morphological partial reduplica-tion. Consider the following forms (from Crowley 2004, 26):

(49) a. CV reduplication
      si-silak 'loose'
   b. (C)VC reduplication
      sil-silak 'loose'

Once again, both types of reduplication presuppose a base *silak*, i.e. with epenthetic [i]. Finally, a remark is in order here with respect to the treat-ment of etymological /sl-/ clusters. According to Rickford (1986, 162), "studies of current usage can be used to shed light on the interpretation of documentary texts". On this view, it follows that Modern Bislama variants such as *silip* 'sleep' suggest that /sl-/ clusters must have undergone [i] epenthesis in Early Bislama as well, even though no such form is attested (see section 2.2). The situation in the modern variety therefore raises doubts about the accuracy of the transcriptions of Early Bislama which

suggested that the onset cluster at issue survives as such. As for the /sw-/ cluster, it is permitted, as illustrated by the forms below:

(50)   *swit* 'sweet'

The findings of this section can be summarized as follows. Fossilized forms such as *kilok* 'clock' occur only in religious texts.

In epenthesized forms only two repair strategies are employed: insertion of default [i]/[e] or vowel copying.

There is an ongoing tendency towards the relaxation of phonotactic constraints on onsets. This accounts both for variation in reflexes of etymological complex onsets and for the occurrence of exclusively non-epenthesized forms.

Anglicization of standard Modern Bislama is reflected in the fact that all the types of obstruent + liquid, obstruent + glide, /s/-initial clusters are allowed.

3.3 Codas

As with complex onsets, the adjustment of etymological word-final complex codas has only been briefly touched upon in the literature. With respect to word-final complex codas in reflexes of monosyllabic etyma, Crowley (2003, 23) writes that "sometimes, a word may be pronounced by different speakers–or by the same speaker on different occasions–as either a single syllable or as two syllables". Crowley (2008a, 158) rightly notes that "final clusters which […] are regularly reduced in Bislama".

As shown in what follows, the repair strategy used for the resolution of illicit word-final complex codas depends on their make-up. For instance, reflexes of consonant + /t/ or /d/ consistently exhibit deletion of /t/ (51) or /d/ (52)–(53), respectively:

(51) /-kt/
     distrik 'district'
(52) /-nd/
     han 'hand'
(53) /-ld/
     gol 'gold'

Word-final consonant deletion also occurs in reflexes of nasal stop + [LABIAL], [DORSAL] clusters (examples from Crowley 2008a, 158):

(54) /-mp/
     stam 'stamp
(55) /-ŋk/
     stiŋ 'to stink'

Another repair strategy employed is epenthesis. Charpentier (1979, 276), for instance, writes that "la où il y avait groupement consonantique en anglais, une voyelle épenthétique a été intercalée", but fails to observe that this only holds for specific types of etymological coda clusters. More recently, Crowley (2008a, 158) states that "some word-final clusters involving a consonant followed by a sibilant are optionally separated by an epenthetic front vowel". As shown below, the epenthetic vowel at issue is [i] or, more rarely, [e].

   A first type of word-final complex codas which may undergo epenthesis is represented by consonant + fricative clusters. Insertion of [i] as a default epenthetic vowel is amply documented in e.g. Charpentier (1979, 276), Crowley (2003, 12, 2008: 158), Bislama (n.d.), Bislama_transcripts (n.d.). Below are some illustrative examples:

(56) /-ks(t)/
     a.   *akis* 'axe'
     b.   *boks ~ bokis* 'box'
     c.   *neks ~ nekis* 'next'
(57) /-nθ/
     mans ~ manis 'month'
(58) /-ns(t)/
     a.   *jans ~ janis* 'chance'
     b.   *akenis* 'against'
(59) /-nʧ/
     lanis 'speedboat'
(60) /-lv/
      twelef 'twelve'

Contra Crowley (2008a, 158), epenthesis not limited to sibilants, as in (56)–(59), but also extends to non-sibilant fricatives, as in (60). The correct generalization, therefore, is that default [i]/[e] may be epenthesized into consonant + [+strident] coda clusters.

   Some of the examples under (56)–(60) attest to the occurrence of variation between non-epenthesized and epenthesized variants. However, the data set out in Table 7 clearly point to a preference for epenthesized variants of reflexes of consonant + sibilant clusters:

Table 1-7 Frequency of variable reflexes of consonant + sibilant clusters in spoken Bislama

Form	Bislama (n.d.)		Bislama transcripts (n.d.)	
	Non-epenthesized	Epenthesized	Non-epenthesized	Epenthesized
aks ~ akis 'axe'	*aks* 0	*akis* 42 (26)	*aks* 2 (1)	*akis* 8 (2)
boks ~ bokis 'box'	*boks* 14 (6)	*bokis* 271 (80)	boks 2 (1)	*bokis* 9 (5)
siks ~ sikis 'six'	*siks* 22 (12)	*sikis* 194 (120)	*siks* 58 (18)	*sikis* 47 (11)
neks ~ nekis 'next'	*seks* 22 (14)	*nekis* 378 (238)	*neks* 16 (7)	*nekis* 62 (20)
taks ~ takis 'tax'	—	—	*taks* 18 (9)	*takis* 91 (53)
faks ~ fakis 'fax'	*faks* 2 (2)	*fakis* 0	—	—
foks ~ fokis 'fox'	*foks* 17 (6)	*fokis* 2 (2)	*foks* 8 (5)	*fokis* 2 (2)
dans ~ danis 'dance'	*dans* 0	*danis* 225 (51)	*dans* 4 (3)	*danis* 115 (17)
fens ~ fenis 'fence'	*fens* 5 (4)	*fenis* 44 (19)	*fens* 6 (3)	*fenis* 9 (3)
jans ~ janis 'chance'	*jans* 15 (13)	*janis* 284 (208)	*jans* 8 (5)	*janis* 29 (13)
mans ~ ma-nis 'month'	*mans* 4 (4)	*manis* 779 (418)	*mans* 12 (4)	*manis* 60 (19)
jens ~ jenis 'change'	*jens* 69 (41)	*jenis* 502 (283)	*jens* 26 (8)	*jenis* 20 (9)

Epenthesized variants overwhelmingly outnumber non-epenthesized ones. The only two exceptions in Table 7 are *faks* 'fax', which does not exhibit variation and *foks ~ fokis* 'fox', with the non-epenthesized variant prevailing. However, these exceptions can be accounted for: the term for 'fax' is obviously a neologism and the fox is an animal not indigenous to Vanuatu. Another rare exception is given below:

(61) /-lθ/
     helt 'health'

This form is exceptional on two counts: the reflex of etymological /θ/ is [t], not [s]; it contains a word-final complex coda.

Liquid + consonant clusters are resolved via two repair strategies. Epenthesis of default [i] or [e] is illustrated by forms or variants such as the following:

(62) /-lm/
       filem 'film'
(63) /-rn/
       lasitern ~ lasiterin 'ground water reservoir' < F *la citerne*

The other repair strategy is vowel copying. Consider the examples below:

(64) /-lp/
       help ~ helep 'help'
(65) /-lk/
       a.   *melek* 'milk'
       b.   *silik* 'silk'

Variation between non-epenthesized and epenthesized variants is captured by the data set out in Table 8:

Table 1-8 Frequency of liquid + consonant clusters in spoken Bislama

Form	Bislama (n.d.)		Bislama_transcripts (n.d.)	
	Non-epenthesized	Epenthesized	Non-epenthesized	Epenthesized
milk ~ melek 'milk'	*milk* 14 (11)	*melek* 86 (44)	*milk* 0	0
silk ~ silik 'silk'	*silk* 0	*silik* 4 (3)	*silk* 0	0
film ~ filem 'film'	*film* 33 (11)	*filem* 29 (15)	—	—

It should be noted that *milk* for 'milk' is clearly an anglicized variant.

In conclusion, complex codas are generally disallowed in Modern Bislama and are resolved via consonant deletion or vowel epenthesis. Complex codas do occasionally occur, but only in urban Bislama or anglicized forms.

Word-medial codas appear to be rarely subject to adjustment. The only repair strategy attested is insertion of the default epenthetic vowel [i]. Consider the following examples (from Crowley 2004, 12, Bislama n.d., Bislama_transcripts n.d.):

(66) /-t/
     dokta ~ dokita 'doctor'
(67) /-l/
     a. *alkol ~ alikol* 'alcohol'
     b. *olsem ~ olisem* 'like'

The frequency of occurrence of non-epenthesized and epenthesized variants in reflexes of etymological word-medial codas is given in Table 9:

Table 1-9 Frequency of reflexes word-medial codas in spoken Bislama

Form	Bislama (n.d.)		Bislama transcripts (n.d.)	
	Non-epenthesized	Epenthesized	Non-epenthesized	Epenthesized
dokta ~ dokita 'doctor'	*dokta* 181 (101)	*dokita* 0	*dokta* 16 (8)	*dokita* 0
alkol ~ alikol 'alcohol'	*alkol* 17 (9)	*alikol* 0	*alkol* 2 (2)	*alikol* 0
olsem ~ olisem 'like'	*olsem* 12516 (2908)	*olisem* 0	*olsem* 258 (68)	*olisem* 2 (2)

As can be seen, epenthesized variants occur extremely rarely. Moreover, two of these are archaic variants: *dokita* for 'doctor' and *olisem* for 'like', see also the discussion of examples (30) and (31) in section 2.3.

In light of the above, it can safely be concluded that word-medial codas are generally permitted in Modern Bislama;

4 Conclusions

Early Bislama tends to disallow complex syllable margins. The repair strategies employed for the adjustment of illicit complex onsets and complex codas are epenthesis and consonant deletion. Forms attested in spoken and written Modern Bislama, however, reflect the interplay of several factors: fossilization, relaxation of phonotactic constraints on onsets, and anglicization (especially of urban Bislama). The repair strategies–epenthesis and consonant deletion–are mostly restricted to the resolution of illicit coda clusters.

Consider next the quality of the epenthetic vowels. In most languages epenthetic vowels tend to be minimally marked or underspecified segments. Cross-linguistically it is featurally unmarked vowels such as [i] or [e] that are selected in epenthesis. This is the case of both Early Bislama

and Modern Bislama. Note, incidentally, that Bislama is not the only Pacific English-lexifier contact language with two default epenthetic vowels: the same is true of e.g. Early Tok Pisin (Avram 2005, 88) and Early Pijin (Avram 2007, 2009, 2014). Also, epenthetic may be contextually coloured, i.e. subject to various sorts of assimilation (Avram 2005, 206), e.g. vowel copying, labial attraction. As shown, vowel copying is found throughout the documented history of Bislama and labial attraction is attested in Early Bislama.

Early Bislama and Modern Bislama also differ with respect to the status of the epenthetic vowels. In Early Bislama, the epenthetic vowels are part of the underlying representations, i.e. they are lexical. As for Modern Bislama, however, the picture that emerges is a more complex one. In most reflexes of etymological complex onsets, the originally epenthetic vowels do not occur anymore, i.e. they are no longer lexical; in a small number of reflexes of etymological complex onsets and in some reflexes of etymological complex codas these are still part of lexical entries; other forms exhibit variation between phonetic realizations with or without epenthetic vowels.

Finally, the findings of the present paper are also relevant to the fate of some features of English syllable structure in language-contact situations, such as those in which pidgins emerge and develop. Blevins (2006, 14) identifies several what she calls "stable features of English word/syllable structure". These include the following (Blevins 2006, 14): onsets may be simple or complex; codas may be simple or complex; CR onsets are possible (R = liquid); sC onsets are possible. Blevins (2006, 16) rightly observes that, for instance, onsets may be simple or complex "in all varieties of Modern English that have enjoyed relatively natural histories with little change induced by external factors". Blevins (2006, 16) further notes that CR onsets and sC onsets are also "stable features of Modern English" in this sense. However, these stable features of English syllable structure may become, to quote Blevins (2006, 16), "unstable features under contact". One of the examples given by Blevins (2006, 16) is a consequence of the fact that "new varieties of English have arisen in close contact with languages that do not allow complex onsets". In such cases "changes eliminating complex onsets are in evidence, including C-deletion and V-epenthesis". As shown in 2.2, consonant deletion is not attested as a strategy for the reduction of complex onsets in Early Bislama, which only resorts to vowel epenthesis. Also, some complex onsets are allowed. Nevertheless, the Early Bislama data examined confirm Blevins's (2006: 17) general conclusion that "the stability of [English] complex onsets […] is rendered unstable by intense contact or influence from languages with

distinct phonotactics". Blevins (2006) does not discuss the fate under contact of the stable feature of English "codas may be simple or complex". *Mutatis mutandis*, however, her conclusions with respect to the instability of English complex onsets also apply to English complex codas, as demonstrated by both Early Bislama and Modern Bislama data.

References

Alexander, Gilchrist. 1927. *From the Middle Temple to the South Seas*. London: John Murray.

Anon a. 2014. *Bislama Workbook*, 3rd revised edition. Port Vila: Peace Corps Vanuatu.

Anon. b. 2017. *Laef long komyuniti. Tija gaed yia 2*. Port Vila: Ministri blong Edukesen mo Trenin.

Anon. c. 2017. *Lanwis mo komyunikesen. Tija gaed yia 2*. Port Vila: Ministri blong Edukesen mo Trenin.

Avram, Andrei A. 2005. *On the Syllable Structure of English Pidgins and Creoles*. Bucharest: Editura Universităţii din Bucureşti.

Avram, Andrei A. 2007. "Syllable Restructuring in Early Solomon Islands Pidgin English." *Bucharest Working Papers in Linguistics* IX (1): 219–30.

Avram, Andrei A. 2009. "The Epenthetic and Paragogic Vowels of Pijin". *Revue roumaine de linguistique* LIV (3–4): 365–82.

Avram, Andrei A. 2014. "Syllable Restructuring in Early Solomon Islands Pidgin English: An Optimality-Theoretic Analysis". *Analele Universităţii din Bucureşti. Limbi şi literaturi străine* LXIII (2): 15–36.

Avram, Andrei A. 2022. "An Optimality-Theoretic Account of Syllable Restructuring in Early Bislama." *Bucharest Working Papers in Linguistics* XXIV (2): 59–79.

Baker, John R. 1929. *Man and Animals in the New Hebrides*. London: Routledge.

Bislama. n.d. https://www.corpus.dynamicsoflanguage.edu.au.

Bislama_transcripts. n.d. https://go.coedl.net/bislama_corpus.

Blevins, Juliette. 2006. "New Perspectives on English Sound Patterns. "Natural" and "Unnatural" in Evolutionary Phonology". *Journal of English Linguistics* 34 (1): 6–25.

Camden, William G. 1977. *A Descriptive Dictionary Bislama to English*. Villa: Maropoa Bookshop.

Charpentier, Jean-Michel. 1979. *Le pidgin Bislama(n) et le multilinguisme aux Nouvelles-Hébrides*. Paris: Société d'Études Linguistiques et Anthropologiques de France (SELAF).

Crowley, Terry. 1990. *Beach-la-Mar to Bislama. The Emergence of a National Language in Vanuatu*. Oxford: Clarendon Press.

Crowley, Terry. 1998. "The Bislama Lexicon before the First World War: Written Attestations". In *Papers in Pidgin and Creole Linguistics 5*, edited by Peter Mühlhäusler, 61–106. Canberra: The Australian National University.

Crowley, Terry. 2003. *A New Bislama Dictionary*, 2nd edition. Vila: Institute of Pacific Studies, Pacific Languages Unit.

Crowley, Terry. 2004. *Bislama Reference Grammar*. Honolulu: University of Hawai'i Press.

Crowley, Terry. 2008a. "Bislama: Phonetics and Phonology". In *Varieties of English*, vol. 3, *The Pacific and Australasia*, edited by Kate Burridge and Bernd Kortmann,143–163. Berlin New York: de Gruyter Mouton.

Crowley, Terry. 2008b. "Bislama: Morphology and Syntax". In *Varieties of English*, vol. 3, *The Pacific and Australasia*, edited by Kate Burridge and Bernd Kortmann, 444–66. Berlin New York: de Gruyter Mouton.

Fletcher, Robert J. 1923. *Isles of Illusion: Letters from the South Seas*. London: Constable and Co.

Fletcher, Robert J. 1924. *Gone Native*. Boston: Small, Maynard and Company.

Giles, William E. 1968 [1877]. *A Cruize in a Queensland Labour Vessel to the South Seas*. Canberra: Australian National University Press.

Grimshaw, Beatrice. 1907. *From Fiji to the Cannibal Islands*. London: Eveleigh Nash.

Gud Nyus bilong Jisas Krais. We Matyu, mo Mak, mo Luk, mo Jon oli raitim. 1972. Hong Kong: The Bible Society in New Zealand.

Guy, Jean B. M. 1974. *Handbook of Bichelamar. Manuel de bichelamar*. Canberra: The Australian National University.

Harrisson, Tom. 1937. *Savage Civilization*. New York: Alfred A. Knopf.

Imhaus, Edouard N. 1890. *Les Nouvelles-Hébrides*. Paris: Berger-Levrault et Compagnie.

Jacomb, Edward. 1914. *France and England in the New Hebrides: The Anglo-French Condominium*. Melbourne: Robertson.

Jacomb, Edward. 1929. *The Joy Court (comédie rosse)*. London: Braybrook and Dobson.

Jarraud-Leblanc, Cendrine. 2013. "The Evolution of Written Bislama". PhD diss., University of Auckland and Université de la Nouvelle-Calédonie.