Te Reo
the Journal of the Linguistic Society of New Zealand

Volume 69
Issue 1

Research Article

2026
Pages 142–174

October, 2026

The realisation of consonant clusters in te reo Māori loanwords

Craig W. J. S. Stansfield
Te Herenga Waka

This paper is a peer-reviewed contribution from https://www.nzlingsoc.org/journal/current-issue/
©*Te Reo – The Journal of the Linguistic Society of New Zealand*
Editor: Victoria Chen

The realisation of consonant clusters in te reo Māori loanwords

CRAIG W. J. S. STANSFIELD[1]
Te Herenga Waka

Abstract

When foreign words are borrowed into te reo Māori and are made to conform to its phonemic inventory and phonotactics, speakers must often make choices as to how to deal with the loanwords' consonant clusters. Ray Harlow's foundational chapter 'Borrowing and its alternatives in Māori' (Harlow, 2004) sets out the basic choice between anaptyxis and consonant loss, but which strategies are used for which te reo Māori loanwords; why; which vowels are inserted when anaptyxis is involved; and which consonants are lost in which environments when cluster reduction occurs has been little studied. This paper summarises the conclusions in the literature and extends them by a study of the loanword appendix in Herbert William Williams's classic *A Dictionary of the Maori Language* (Williams, 1844/1971) to propose a fuller set of conclusions regarding the phonological adaptation of loanwords in te reo Māori.

Keywords

Māori, borrowing, loanword, consonant cluster, anaptyxis

1 Introduction

In acquiring a loanword, speakers of a language need to make a choice between retaining the source language's phonemic inventory and phonotactics (wholly or in part) or adapting it (wholly or in part) to the phonemic inventory and phonotactics of their own language. A formal distinction can be made between the former type of loanwords, known as *Fremdwörter* (German, 'foreign words'), and the latter type, known as *Lehnwörter* (German, 'loanwords'). The degree of adaptation is not necessarily binary, but rather is often a point on a spectrum, and the adaptation may occur independently of orthography (where this is relevant). For example, the word *garage* has been borrowed from French into many languages, including English, without (in English) any change in spelling, but with variable realisations in English pronunciation. The French pronunciation is /ga.ʁaʒ/; New Zealand English pronunciation ranges from a modified /gəˈɹɐːʒ/ to a much more fully Anglified /ˈgɛ.ɹədʒ/. The former pronunciation can be described as a *Fremwort*, the latter as a *Lehnwort*.

Some of the factors that can influence this choice between Fremdwörter and Lehnwörter are the degrees to which the source and recipient languages have (i) common phonemic inventories; (ii) similar phonotactic constraints such as allowable syllable structures; and (iii) similar word stress patterns. This study investigates te reo Māori Lehnwörter taken from English as well as a small number of tokens from other contact languages. As Māori and

[1] I wish to thank two anonymous reviewers for their constructive and careful feedback on this paper, which has greatly improved it. Any errors of course remain my responsibility.

English possess differing phonemic inventories, allowable syllable structures, and stress patterns, a closer investigation of Māori Lehnwörter is informative for our understanding of how speakers of te reo Māori perceive and choose to modify English source words and how this contrasts with the approach of speakers of other languages in creating their languages' Lehnwörter.

Te reo Māori is a particularly useful case study in this regard because of:

(i) The long history of contact between Māori and English, in a political environment of colonisation and a linguistic environment where English moved from being the language of a small proportion of New Zealand's population to its majority language within 20 years (Bennett, 1861; compare the census of Māori of September 1857–September 1858 and the *Second census of the Colony of New Zealand* (Pākehā/European census) of 24 December 1858) and where most Māori became bilingual.

(ii) The very strong preference of Māori speakers for Lehnwörter over Fremdwörter (Bauer et al., 1993, p. 354).

(iii) The restricted phonemic inventory of te reo Māori, one of the world's smallest (Bauer et al., 1993, pp. 529–537).

(iv) The strict (C)V(V) syllable structure of te reo Māori, with both consonant clusters and codas prohibited (Bauer et al., 1993, pp. 543–544; Warren & Bauer, 2004, pp. 615, Bauer & Warren, 2004, p. 581).

These factors result in a large number of loanwords in the Māori vocabulary (at least prior to the Māori Renaissance of the early 1970s onwards) and a requirement for a very high proportion of loanwords to be modified in the process of being borrowed.

Looking specifically at the interesting approaches taken by Māori speakers to English source word consonant clusters, these considerations give rise to the basic research question of this paper: **How are source word consonant clusters made to conform to the required (C)V(V) syllable structure of te reo Māori?** This in turn gives rise to the detailed research questions of: (a) how is the consonant cluster of the source word modified?; (b) if one or more consonants are deleted, which one(s) are they?; (c) if an anaptyctic vowel is inserted, how is it determined?; and (d) what are the likely phonetic and/or other motivations for all of these choices?

This paper is organised as follows. Section 1 is an introduction to the research area and why it is of interest. Section 2 sets out the background, including an overview of loanwords, an overview of the consonant phonology of New Zealand English and te reo Māori, and the research questions. Section 3 sets out the methodology, including the data source and the excluded tokens. Section 4 sets out the results, firstly for the phonemic correspondences and then the main findings relating to consonant clusters and to anaptyxis (epenthesis of a vowel). Section 5 outlines the conclusions that can be drawn from the results.

2 Background

2.1 Loanwords and systematicity: An overview

There is a substantial body of literature on loanwords. A central question that has been discussed by several recent works concerns the degree of systematicity in loanwords attested in various cases of language contact. Schütz (1985, p. 24), for example, observes that borrowing may be 'a process that consists of ordered rules'. In relation to anaptyxis (epenthesis of a vowel), a key area of research with loanwords, Uffman (2006, p. 1080) comments that:

> Epenthesis … satisfies constraints on native syllable structure, most notably bans on codas and on complex onsets. While such generalizations have been firmly established in the literature at least since Meinhof's (1910:14) observations on vowel epenthesis in loanwords in Bantu languages, there still is comparatively little empirical work regarding the question which vowel exactly is inserted, both within a language, in a given context, and crosslinguistically.

However, as noted by these researchers, these analyses are unlikely to be comprehensive, given the known variability in loanwords deriving from a single source word, which means that the attested variation cannot be fully accounted for. For example, Ka'ai and Moorfield (2009, p. 19) record as deriving from 'Britain' /ˈbɹɘt.ɘn/ the five Māori Lehnwörter *Piritini* /piɾitini/, *Piriteni* /piɾiteni/, *Piritene* /piɾitene/, *Piritani* /piɾitani/, and *Piritana* /piɾitana/: no single rule can account for the /i/~/e/~/a/ variation in the realisation of the unstressed /ɘ/ vowel of the second syllable, nor for the similar /i/~/e/~/a/ variation in the paragogic (final) vowel of these Lehnwörter.

This study assumes that loanwords are amenable to analyses founded on known, 'natural' sound changes, but that the rules resulting from these analyses will not be fully predictive as borrowing occurs lexical item by lexical item, at discrete times, rather than proceeding systematically across the lexicon over a finite period, as happens with other sound changes.

The primary assumption underlying attempts to form ordered (if not exceptionless) rules for borrowing is that the Lehnwort will resemble the source word as much as the recipient language's phonemic inventory and phonotactics permit. The assumption is, as Schütz (2004, p. 285) states, that 'when two phonological systems meet head-on through borrowing, we expect new forms to be as phonetically close to the model as possible'. In commenting on borrowing from English into Fijian Schütz (2004, p. 261) terms this principle one of 'phonetic similarity':

> we can posit a principle for the relationship between the forms: PHONETIC SIMILARITY. That is, we assume that a newly coined Fijian word should sound as much like the English model [that is, the source word] as possible, but still fit within the Fijian system.

2.2 New Zealand English and te reo Māori phonology: an overview

The Māori language is known in Māori as *te reo* (or as *te reo Māori* where there is a need to avoid ambiguity with other languages). Speakers of te reo have been in sustained contact with speakers of other languages since the 1790s, particularly with English speakers from the British Isles, Australia, and North America, but also with speakers of French and other languages. This contact reached a substantial level from the 1840s onwards, following the commencement of systematic and sustained immigration into New Zealand. This relatively massive and sustained contact with speakers of other languages, particularly of English, coupled with the technological, cultural, religious, and political changes and innovations that this contact

entailed, resulted in a large number of loanwords being borrowed into te reo. The massive scale of immigration into New Zealand from the 1840s onwards meant that speakers of te reo became a minority in New Zealand from the mid-1850s (Bennett, 1861; compare the census of Māori of September 1857–September 1858 and the *Second census of the Colony of New Zealand* (Pākehā/European census) of 24 December 1858).

Te Aka Māori Dictionary (Māori Dictionary Project, n.d.) records that it contains a historical loanwords collection, which is

> a collection of over 18,000 loanwords from the Māori-language newspapers and the Paipera Tapu, the Māori Bible, and a few other sources. These were all published from 1842 to 1952. A significant number of these loanwords are still in use in modern Maori….

Many other historically recorded loanwords, however, have fallen out of use, because of further technological changes or because of more recent lexical additions to te reo, particularly through the creation of new terms from existing Māori linguistic resources, and particularly since the creation of Te Taura Whiri i te Reo Māori (the Māori Language Commission) in 1987. For example, the lexical extension of *waka* to include 'car' is now much more common than the loanwords *kā*, *motokā*, and *motukā* (all *Te Aka Māori Dictionary*). Degani (2012, p. 16) notes that 'attitudes towards lexical borrowing changed starting from the 1970s. In this period, some Māori people, and especially those involved in the preservation of "pure" Māori, frowned upon English loans and avoided [using] them as much as possible'. The 'What we do' description of the work of Te Taura Whiri i te reo Māori on its website includes 'Ensure te reo Māori remains a living language by developing new words so that it is used effectively in a changing world' (Te Taura Whiri i te Reo Māori, n.d.).

Te reo Māori and English and most of the other languages from which te reo has acquired lexical items—notably French, Hebrew, and Latin (Williams 1844/1971)—have significantly different (although substantially overlapping) phonemic inventories; very different allowable syllable structures; and frequently different stress patterns. (Williams also lists a loanword from Tahitian, which, being another Central-Eastern Polynesian language, differs much less from te reo Māori.) Māori speakers have invariably preferred to adapt loanwords to the phonology of te reo, that is, Māori has Lehnwörter rather than Fremdwörter; and pronunciations that do not conform to Māori phonology are therefore examples of code switching, not loanwords. Code switching is common in contemporary Māori as virtually all post-adolescent speakers of te reo are fully bilingual in te reo and English or are multilingual.

2.2.1 English and te reo Māori transcription

I have followed Bauer and Warren (2004) and Warren and Bauer (2004) for the transcription of the English source words set out in the Williams appendix. This is a system for the phonemic transcription of New Zealand English known as PhoNZE. The use of a single and contemporary variety of English to transcribe the varieties of English from which the loanwords came is a major but necessary simplification, as the reality is that these loanwords were sourced from a variety of Englishes, including the American English and Australian English of the early whalers and sealers (McNab, 1913; Rickard, 1965) and the Irish, Scottish, Welsh, Australian, and Northern England Englishes of settlers throughout the colonial period and up to the present, as well as the English of Southern England of the stereotypical New Zealand Company immigrant. Nevertheless, the use of New Zealand English as the standard for source word transcription seems to me to be a reasonable simplification of the complexity of the loanword environment, given that:

(i) There is little reliable data about when and from whom any particular loanword was borrowed into te reo.

(ii) The majority of loanwords in the Williams appendix will likely have been borrowed in the latter stages of the approximately 180 years of the source data period (circa 1790 to circa 1970), once New Zealand English had become well established, when speakers of New Zealand English were the English speakers that Māori interacted with most often; and once most speakers of te reo Māori were also fluent in English, that variety of English being New Zealand English.

(iii) The differences in the varieties of English are most relevant to vowels, not to consonants; whereas the present paper is concerned with how consonant clusters are realized in te reo loanwords, making the differences in the varieties of English less important for the present purpose. This difference in dialectal variation between (principally) vowels (for English dialects) and (overwhelmingly) consonants (for Māori mita) is one of the interesting differences between English and Māori.

I have included primary and secondary stress and syllabification markings for the New Zealand English transcriptions, to ensure that it is clear when any particular consonant cluster forms an onset or a coda, which is critical to the analyses; but I have not attempted to do so for the transcription of te reo Māori, as this necessarily involves judgements about the nature of diphthongs and non-diphthongal vowel sequences that are best left to native speakers.

2.2.2 New Zealand English and te reo Māori consonants

New Zealand English has a phonemic inventory of 24 consonants and te reo Māori has a phonemic inventory of 10 consonants.

The 24 New Zealand English consonants consist of 6 plosives /p/, /b/, /t/, /d/, /k/, and /g/; 3 nasals /m/, /n/, and /ŋ/; 2 affricates /tʃ/ and /dʒ/; 9 fricatives /f/, /v/, /θ/, /ð/, /s/, /z/, /ʃ/, /ʒ/, and /h/; 3 approximants /ɹ/, /j/, and /w/; and 1 lateral approximant /l/ (Bauer & Warren, 2004, pp. 593–597). For simplicity, I do not discuss the allophones of the New Zealand English consonants here and only comment on them individually in the results as is needed.

The 10 Māori consonants consist of 3 plosives /p/, /t/, and /k/; 3 nasals /m/, /n/, and /ŋ/; 1 tap /ɾ/; 2 fricatives /ɸ/ and /h/; and 1 approximant /w/. There are no affricates or phonemic lateral approximants in te reo and the small number of fricatives is especially noteworthy; in particular, there are no sibilants (Bauer et al., 1993, pp. 529–533). This is a phonemic inventory of the te reo Māori consonants, and so ignores allophones, which can be extensive. The allophonic range of the two phonemic fricatives /ɸ/ and /h/ is large because of the relative lack of fricatives and the absence of sibilants. The bilabial fricative /ɸ/ can be realised as [f] (labiodental) and as [θ] (dental) as well as as [ɸ] (bilabial); and the glottal fricative /h/ can be realised as [ç] (palatal), as [x] (velar), as [χ] (uvular), and as [ħ] (pharyngeal) as well as as [h] (glottal) (Hohepa, 1965, pp. 9–10). The alveolar tap /ɾ/ can be realised as [ɹ] (that is, as an approximant) and as [l] (that is, as a lateral approximant) as well as as a tap (Harlow, 2007, p. 77).

2.2.3 English and te reo Māori syllable structures

English allows closed and open syllables, that is, syllables with and without codas (respectively). Onsets are permitted and are common but are optional and may consist of one, two, or three consonants. Nuclei may consist of a vowel, either a monophthong or a diphthong, or a syllabic consonant, which may only be /m/, /n/, or /l/ (or /ɹ/, in rhotic varieties). Codas are

permitted and are common but are optional and may consist of one, two, three, four, or five consonants.

The permissible syllable structure can be represented as (ω)ν(κ), where *(ω)* represents an optional onset, *ν* represents the nucleus (a vowel or a syllabic consonant), and *(κ)* represents an optional coda. The permissible syllable structure can also be represented (ignoring, for simplicity, syllabic consonants) as $(C)^3V(V)(C)^5$, where *C* represents a consonant and *V* a vowel, although no actual English syllable realises the theoretical maximal structure with a three-consonant onset and a five-consonant coda. A near-maximal example (being C^3VC^4) is 'strengths' (/stɹeŋkθs/ in New Zealand English).

Te reo Māori allows only open syllables, that is, syllables without codas. Closed syllables (with codas) are prohibited. Onsets are permitted and are common but are optional and may only consist of a single consonant. Therefore, with codas not permitted and onsets only permitted to consist of a single consonant, consonant clusters never occur. Nuclei may only consist of a vowel, either a monophthong or a diphthong, and there are no syllabic consonants.

The permissible syllable structure can be represented as (ω)ν, where *(ω)* represents an optional onset and *ν* represents the nucleus. The permissible syllable structure can also be represented as (C)V(V).

2.3 Research questions

The major differences between Māori and English phonology mean that the realisation of English consonant clusters in Māori Lehnwörter is an interesting aspect of borrowing to study. This paper looks at how consonant clusters are realised in a way that conforms to Māori phonotactics and the relatively small phonemic inventory of te reo. The Māori Lehnwörter used to investigate this are taken principally from English, but include one token of a source word from French, that is, *miel* 'honey'. Te reo Māori (i) prohibits syllabic consonants; (ii) prohibits consonant clusters; (iii) prohibits codas; and (iv) allows only single-consonant onsets. In other words, the permissible Māori syllable structure is (C)V(V). How then are source word consonant clusters made to conform to this required (C)V(V) structure? This overarching research question has four aspects to it. Firstly, how is the consonant cluster of the source word modified to fit? There are several possibilities, that is, that the consonant cluster is deleted in its entirety as part of realising the source word; or that it is simplified through the deletion of one or more of its phonemes; or that an anaptyctic vowel is inserted to split the consonant phonemes into separate syllables; or that amenable consonant phonemes (that is, semivowels) are realised as vowels. Secondly, if one or more consonants are deleted, which one(s) are they? Thirdly, if an anaptyctic vowel is inserted, how is it determined, that is, how are its height, backness, and length selected when there is no equivalent phoneme in the source word to try to match? Fourthly, what are the likely phonetic and/or other motivations for all of these choices?

Research questions. How are source word consonant clusters made to conform to the required (C)V(V) syllable structure of te reo Māori?

(a) How is the consonant cluster of the source word modified?

(b) If one or more consonants are deleted, which one(s) are they?

(c) If an anaptyctic vowel is inserted, how is it determined?

(d) What are the likely phonetic and/or other motivations for all of these choices?

An assumption made in formulating these research questions is that irregularity will occur in how consonant clusters are realised because borrowing is a type of language change that does not necessarily conform to the uniformity of other types of language change, as lexical items are borrowed (and therefore adapted) one-by-one (or occasionally in small groups) and at disparate times, as the lexical need arises, rather than the sound change(s) operating across the lexicon over a distinct and finite period; but that regularities will nevertheless still be present because adaptations of source words to make them conform to the recipient language's phonemic inventory and phonotactics are still sound changes, and are still subject to the constraints on sound change that are captured by the concept of 'naturalness'. In short, the assumption in the research questions is that borrowing is still amenable to analyses that posit natural sound change rules; but that these rules will not be fully predictive in the case of borrowing.

3 Methodology

3.1 Data source

The main data source for this research is Williams (1844/1971). Specifically, the Māori Lehnwörter listed in the appendix on pages 501–507[2] of the seventh edition of *A Dictionary of the Maori Language*, which is entitled 'SOME OF THE MORE IMPORTANT WORDS ADOPTED FROM NON POLYNESIAN SOURCES'. This important lexicographic work was first compiled by William Williams and published in Paihia in 1844. The fifth edition was compiled by his grandson, Herbert William Williams, and published in Wellington in 1917. The seventh edition used for this research was revised and augmented by the Advisory Committee on the Teaching of the Maori Language in the Department of Education and published by A. R. Shearer, the Government Printer, in 1971. It contains a preface to the sixth edition of 1957 by J. M. McEwen and a preface to the seventh edition of 1971 by G. P. Kelly. For reasons of brevity and because of the derived page referencing used for the source I shall refer to this specific resource within Williams (1844/1971) as 'the Williams appendix' rather than systematically as Williams (1844/1971, pp. 501–507).

The Lehnwörter in the Williams appendix are overwhelmingly borrowed from English, but loans from French, Hebrew, Latin, and Tahitian are also listed (the Tahitian loanword is despite the appendix title being 'ADOPTED FROM NON POLYNESIAN SOURCES'). The 585 entries in the Williams appendix constitute just a small subset of the 18 000 loanwords recorded in *Te Aka Māori Dictionary*, but they should comprise some of the most commonly used Lehnwörter in te reo Māori in the period immediately before the Māori Renaissance of the 1970s onwards and the deliberate shift as part of that amazing cultural and linguistic revitalisation to filling lexical needs primarily from existing lexical resources within te reo (which includes replacing loanwords). The choice of a now dated, non-contemporary Māori dictionary from 1971—that is, Williams (1844/1971)—is therefore a deliberate one, in order to gain access to the lexicon of te reo Māori at the likely point of its maximal receptiveness to borrowing.

3.2 Excluded tokens

Six entries from the Williams appendix that contain consonant clusters are excluded from the analyses in this paper. These six entries are *putiputi*, *hiko*, *raiti hiko*, *kāwanatanga*, *taipō*, and

[2] The page numbering of Williams (1844/1971) stops at page 499, after the main dictionary entries, and does not continue through the addenda or the appendix. Page numbers 501–507 are therefore derived rather than printed in the volume.

Tararā, said to come from 'pretty pretty' (but meaning 'flower'), 'electricity', 'electric light', 'government', 'ghost', and 'Yugoslav', respectively. The Williams appendix lists all of these as loanwords, but I consider all of them to be at least questionable, for the reasons set out below, and do not want to complicate the analyses by including Lehnwörter that may not in fact be so, as the resulting possibly erroneous phonemic correspondences could obscure relationships that might otherwise be apparent.

The te reo Māori word ***putiputi***, meaning flower, is said to come from the expression 'pretty pretty', but there are no attestations of that borrowing and there is not full agreement as to 'pretty pretty' as the source. I have therefore excluded *putiputi* from the tokens analysed.

In a comment on a Facebook post on the derivation of *putiputi* from 'pretty pretty' Koro Wheke (2020) observed that:

> Im not sure if putiputi is actually a loan word rather than a archaic term given new meaning. Putik is flower in Malay and in Maori puti refers to a certain type of small bud

Malay and te reo Māori are both Malayo-Polynesian languages. The Malay word ***p**utik* 'flower bud'[3] comes from Proto-Western-Malayo-Polynesian (PWMP) ***p**utik*$_2$ 'to pick, pluck fruit'; as does the Simeulue (Simalur, Simolol) word ***p**uti?* 'newly-formed fruit'. I have not been able to find *puti* referenced as 'bud' in te reo Māori, but te reo has multiple words for 'bud', including *kōpuku*, *matikao*, *nihoniho*, *rāpupuku*, *tipu*, *toroihi*, and *tupu*, so an additional word for 'bud' does not seem unlikely given this rich botanical lexicon.

The usual Malay reflex of *p in Proto-Malayo-Polynesian (PMP) is *p*, but the usual Māori reflex of PMP *p is *ɸ*. However, there are examples of PMP *p being reflected in Māori as *p*, namely:

- *ke**p**it 'press together; pressed together; carry under the arm, press between the legs' → *ko**p**i* '1. to be shut, closed, doubled together (as anything hinged or jointed); 2. to nip between the legs' (*Te Aka Māori Dictionary*); 'doubled together, as of anything hinged or jointed; shut, closed; gorge of a river; gather up into small compass; nip between the legs' (*ACD*); and
- ***p**a(n)tuk*$_2$ 'to knock, strike against' → ***p**atu* 'to strike, hit, beat, assault, kill, subdue, ill-treat' (*Te Aka Māori Dictionary*); 'to strike, beat, thrash; subdue; kill; pound (fern root, etc.); also used of washing clothes' (*ACD*).

On balance therefore, a derivation of ***p**uti**p**uti* 'flower' in te reo from a PMP source for PWMP ***p**utik*$_2$ 'to pick, pluck fruit' seems unlikely, given the change in meaning and the rareness of the *p → *p* phonemic correspondence; but not impossible, given the similar semantic shifts in Malay and in Simeulue, and the *ko**p**i* and ***p**atu* reflexes in te reo that do involve this *p → *p* correspondence.

Hiko is said to come from 'electricity' (or more probably from 'electric'), and the related ***raiti hiko*** from 'electric light'. *Raiti* from 'light' is clear, but *hiko* coming from 'electricity' or 'electric' is very controversial (not to mention difficult to analyse in terms of the relevant phonemic correspondences), and a lexical extension from *hiko* meaning 'to flash' (as of lightning) and 'to stimulate' seems to me (among many others) to be at least as likely. I have therefore excluded *hiko* and *raiti hiko* from the tokens analysed.

[3] In the *ACD* this is *putek* 'rudimentary fruit; the fruit as it appears when the blossom has just been shed' and in the original of *ACD*'s source, i.e., Wilkinson (1901), this is *puteḳ* or *putiḳ* 'The fruit as it appears immediately after the falling of the blossom'.

Kāwanatanga is obviously a mixed term. *Kāwana* is indeed a Lehnwort, from 'governor', but the abstract noun suffix *-tanga* (Harlow, 2007, p. 122) is a highly productive derivational suffix of te reo, meaning that *kāwanatanga* as a whole is not a loan word and I have excluded it (as not deriving from 'government') from the tokens analysed.

Taipō is said to mean 'ghost' and 'typhoid', and while a realisation of 'typhoid' as *taipō* is clear, a realisation of 'ghost' /gɐʉst/ → /taipoː/ *taipō* is bizarre, and I am not convinced that it is a Lehnwort, at least not with 'ghost' as the source word. *Te Aka Māori Dictionary* lists *taipō* as meaning 'goblin, spook, ghost—unwanted supernatural visitors not of human origin that haunt the living', but does not reference it as a loanword; but it does list *taipō* meaning 'typhoid, typhus' as a loanword. I have therefore excluded *taipō* (as not deriving from 'ghost') from the tokens analysed.[4]

Tararā is said to mean 'Yugoslav', but a realisation of 'Yugoslav' /ˈjʉː.gə.slɐːv/ → /tarараː/ Tararā would be bizarre, and I consider that it is much more likely to be a Lehnwort from 'Dalmatian' or 'Dallie'. I have therefore excluded *Tararā* from the tokens analysed.

In addition to these six tokens, *rēhita* and *kawenata* are excluded from my analyses. These two tokens are Lehnwörter, but their analyses are unclear because of the phonemic correspondences involved, and I want to avoid possibly erroneous phonemic correspondences obscuring relationships that might otherwise be apparent. Specifically, in analysing 'register' /ˈɹedʒ.ə.stə/ → /ɾeːhita/ *rēhita*, it is unclear whether the Lehnwort's /h/ phoneme realises the source word's /dʒ/ or /s/ phonemes, and therefore whether the Lehnwort's /i/ phoneme is realising the source word's first /ə/ phoneme or is an anaptyctic vowel to break the /st/ cluster into separate syllables. Similarly, in analysing 'covenant' /ˈkɐv.ən.ənt/ → /kawenata/ *kawenata*, it is unclear whether the Lehnwort's /n/ phoneme realises the first or the second /n/ phoneme of the source word, and therefore whether the Lehnwort's second /a/ phoneme is realising the source word's second /ə/ phoneme or is an anaptyctic vowel to break the /nt/ cluster into separate syllables. I have therefore also excluded *rēhita* and *kawenata* from the tokens analysed.[5]

4 Results

This section sets out the results of this study. Before entering into the core discussion, I begin by presenting a brief overview of the phonemic correspondences within the consonant clusters in the Williams appendix (section 4.1). The main results relating to consonant clusters are then presented in section 4.2 and those relating to anaptyxis (epenthesis of a vowel) are presented in section 4.3.

4.1 Phonemic correspondences for retained phonemes in consonant clusters

Unattested phonemes. Within the Williams appendix data /v/, /ð/, /ʒ/, and /h/ never form part of a consonant cluster. This is unsurprising, as /v/, /ð/, and /h/ are rare in English consonant clusters and /ʒ/ never occurs in a cluster.

[4] An anonymous reviewer of this paper has pointed out that phono-semantic matching to *taipō* (meaning 'ghost') could have eased the acceptance of the loanword for 'typhoid'.

[5] An anonymous reviewer of this paper has advised that the derivations from /dʒ/ and from the first /n/ are clear to them, and I agree that these are the most likely correspondences in both cases, particularly given the greater salience crosslinguistically of earlier syllables and phonemes. However, as there is at least some possibility of the alternative analyses involving the final syllables, I have excluded these two tokens from the tokens analysed.

Retained and **deleted phonemes.** Of the remaining 20 English consonants, /θ/, /ʃ/, and /w/ are always retained and /z/ is always deleted when they occur in consonant clusters in the Williams appendix data. However, this is probably a small numbers problem, rather than any quality of these consonants that requires them always to be or never to be realised in te reo Māori Lehnwörter, as they are rare (one token of /θ/, one token of /ʃ/, four tokens of /w/, and one token of /z/).

Unchanged phonemes. If they are not eliminated through a cluster reduction process, of the 20 consonants in the consonant clusters attested in the Williams appendix, the three voiceless plosives /p/, /t/, and /k/ and two of the three nasals, that is, /m/ and /n/, which are all also phonemes in te reo, are always unchanged (29 tokens of /p/, 56 tokens of /t/, 29 tokens of /k/, two tokens of /m/, and 30 tokens of /n/).

Non-coronal voiced plosives. If they are not eliminated through a cluster reduction process, of the three voiced plosives in the attested consonant clusters, two, that is, /b/ and /g/, are always realised in te reo as /p/ (17 tokens) and /k/ (nine tokens), respectively. These phonemes are therefore always devoiced. This is unsurprising as there are no voiced obstruents in te reo and this is as reported in Ryan (1972, p. 144), Harlow (2004, p. 160), and Kaʻai and Moorfield (2009, pp. 19–20 & 23–24).

Voiceless affricate. If it is not eliminated through a cluster reduction process, when /tʃ/ occurs in the attested consonant clusters it always realised in te reo as /h/, that is, it is debuccalised. There is, however, only one token of /tʃ/ that is not eliminated, namely ‘inch’ /ənʧ/ → /iːnihi/ *īnihi*. This is as reported in Ryan (1972, p. 144), Kearns (1990, p. 68), Harlow (2004, p. 160), and Kaʻai and Moorfield (2009, p. 22), although Kearns (1990, p. 68) and Harlow (2004, p. 160) note that /tʃ/ can also be realised as /ti/ and Kaʻai and Moorfield (2009, p. 22) note that /tʃ/ can also be realised as /t/ and /ti/. None of these alternative realisations occur in the Williams appendix consonant clusters.

Voiced affricate. If it is not eliminated through a cluster reduction process, when /dʒ/ occurs in the attested consonant clusters it is always realised in te reo as /h/, that is, it is debuccalised. There is, however, only one token of /dʒ/ that is not eliminated, namely ‘hinge’ /həndʒ/ → /iːnihi/ *īnihi* (identical to the Lehnwort for ‘inch’, discussed immediately above under ‘Voiceless affricate’). This is as reported in Kearns (1990, p. 68), Harlow (2004, p. 160), and Kaʻai and Moorfield (2009, pp. 23–24), although Kearns (1990, p. 68) notes that /dʒ/ can also be realised as /ti/ or /ɾi/, Harlow (2004, p. 160) notes that /dʒ/ can also be realised as /ti/, /ɾi/, or /ɾ/, and Kaʻai and Moorfield (2009, pp. 23–24) note that /dʒ/ can also be realised as /t/, /ti/, or /ɾi/. None of these alternative realisations occur in the Williams appendix consonant clusters.

Voiceless dental fricative. When /θ/ occurs in the attested consonant clusters it is always realised in te reo as /t/, that is, it is strengthened to the corresponding coronal plosive. There is, however, only one token of /θ/, namely ‘thread’ /θɹed/ → /taɾete/ *tarete*. This is as reported in Ryan (1972, p. 144) and Kaʻai and Moorfield (2009, p. 24) and contrasts with the prediction in Lin (2009, p. 9) that manner features for consonants in loanwords are more faithfully retained than place features (which would predict a realisation as /ɸ/ or /h/).

Voiceless postalveolar fricative. When /ʃ/ occurs in the attested consonant clusters it is always realised in te reo as /h/, that is, it is debuccalised. There is, however, only one token of /ʃ/, namely ‘launch’ /loːnʃ/ → /ɾoːnihi/ *rōnihi*. This is as reported in Ryan (1972, p. 144), Kearns (1990, pp. 67–68), Harlow (2004, p. 160), and Kaʻai and Moorfield (2009, pp. 20, 24), although Kearns (1990, pp. 67–68) and Kaʻai and Moorfield (2009, p. 24) note that /ʃ/ can also be

realised as /t/, that is, strengthened to the corresponding coronal plosive. This alternative realisation does not occur in the Williams appendix consonant clusters.

Postalveolar approximant. If it is not eliminated through a cluster reduction process, when /ɹ/ occurs in the attested consonant clusters it is always realised in te reo as /ɾ/ (58 tokens), that is, it is tapped. This is unsurprising as there are, phonemically, no coronal approximants in te reo and the existence of an approximant [ɹ] allophone of the alveolar tap phoneme /ɾ/ in te reo (Harlow, 2007, p. 77) means that there may have been little or no phonetic difference between the source word and the Lehnwort in respect of this phoneme, depending on the phonetic context and the rapidity of speech.

Palatal approximant. If it is not eliminated through a cluster reduction process, when /j/ occurs in the attested consonant clusters it is always realised in te reo as /i/ (six tokens), that is, it is realised as a close front vowel. This is as reported in Kaʻai and Moorfield (2009, p. 24) and is noted in Harlow (2004, p. 160), where he observes the realisation of /i/ as [j] before a stressed vowel in te reo and notes the utility of this synchronic process for realising /j/ in te reo Lehnwörter as /i/.

Labial-velar approximant. When /w/ occurs in the attested consonant clusters it is always realised in te reo as /u/ (four tokens), that is, it is realised as a close back vowel.

Alveolar lateral approximant. If it is not eliminated through a cluster reduction process, when /l/ occurs in the attested consonant clusters it is always realised in te reo as /ɾ/ (36 tokens), that is, it is tapped. This is unsurprising as there are, phonemically, no coronal approximants or lateral approximants in te reo and is as reported in Ryan (1972, p. 144), Harlow (2004, p. 160), and Kaʻai and Moorfield (2009, pp. 21, 24). In addition, the existence of a lateral approximant [l] allophone of the alveolar tap phoneme /ɾ/ in te reo (Harlow, 2007, p. 77; Warren & Bauer, 2004, p. 618) means that there may have been little or no phonetic difference between the source word and the Lehnwort in respect of this phoneme, depending on the dialects of the speakers and the phonetic context.

This leaves the phonemes /d/, /ŋ/, /f/, and /s/ from the attested consonant clusters. These have variable realisations.

Voiced alveolar plosive. If it is not eliminated through a cluster reduction process, when /d/ occurs in the attested consonant clusters it is sometimes realised as /t/ (five tokens), that is, it is devoiced, and sometimes as /ɾ/ (one token), that is, it is tapped. The realisation as /t/ is as reported in Ryan (1972, p. 144) and the variable realisations as /t/ and as /ɾ/ are as reported in Harlow (2004, p. 160) and Kaʻai and Moorfield (2009, pp. 20, 22). The preference for realisation as /t/ over /ɾ/ conforms to the prediction in Lin (2009, p. 9) that manner features for consonants in loanwords are more faithfully retained than place and voicing features. In contrast to Lin (2009, p. 9), Hohepa (1965, p. 10) states that '[in te reo t]he essential contrast between members of the oral fricative continuants is the presence or absence of voicing—and a secondary opposition is the area of friction. …… [/ɾ/] may be interpreted as the voiced counterpart of /t/'. If Hohepa is correct in this regard a realisation of /d/ as /ɾ/ is unsurprising, as maintaining the voicing distinction would remain important despite the absence of voiced plosives in te reo.

The one token realising /d/ as /ɾ/ is part of a cluster reduction from /dz/, that is, '[playing] cards' /kɐːdz/ → /kaːɾi/ *kāri*. It seems possible that Williams (1844/1971) is incorrect in viewing the plural 'cards' as the source word, with the singular 'card' as the actual source word, in which case there is no cluster reduction from /dz/ involved. Te reo speakers exposed to the English word 'cards' for one or more of the many games involving playing cards would also have been exposed to the singular 'card', and probably at a similar or higher frequency (as the

game is played). In addition, nouns for non-human entities are not marked for number on the noun itself in te reo Māori (but rather the relevant noun phrase is marked for number via the definite article), so *kāri*, when used in isolation, can mean both 'card' and 'cards'. The realisation of /d/ as /ɾ/ can therefore possibly be discounted, in which case /d/ conforms to the uniform devoicing of the voiced plosives discussed above for /b/ and /g/ under 'Non-coronal voiced plosives'.

Velar nasal. If it is not eliminated through a cluster reduction process, when /ŋ/ occurs in the attested consonant clusters it is sometimes unchanged (one token, namely 'ink' /ɘŋk/ → /iːŋiki/ *īngiki*) and is sometimes realised as /n/ (one token, namely 'leggings' /ˈleg.ɘŋz/ → /ɾeːkena/ *rēkena*). As far as I am aware, this variation with dorsal and coronal realisations of /ŋ/ has not been addressed in the literature. However, it seems possible that the phonetic realisation of 'leggings' was [ˈleg.ɘnz], given the frequency of an alveolar realisation of the nasal in the *-ing* suffix in English (Hazen, 2006), particularly when followed by an alveolar fricative, as here[6], in which case there is no variation in the realisation of any of the nasals (/m/, /n/, and /ŋ/) when retained.

Voiceless labiodental fricative. If it is not eliminated through a cluster reduction process, when /f/ occurs in the attested consonant clusters it is sometimes realised as /p/ (six tokens), that is, it is strengthened to the corresponding labial plosive; sometimes as /ɸ/ (one token), that is, it is labialised; and sometimes as /h/ (one token), that is, it is debuccalised. The single token as /ɸ/ is 'golf' /gɒlf/ → /koɾoɸa/ *korowha*. The single token as /h/ is 'wolf' /wʊlf/ → /wuɾuhi/ *wuruhi*. However, *Te Aka Māori Dictionary* records /wuɾuɸi/ *wuruwhi* as well as /wuɾuhi/ *wuruhi* as alternative Lehnwörter for 'wolf'. The far greater realisation as /p/ contrasts with the prediction in Lin (2009, p. 9) that manner features for consonants in loanwords are more faithfully retained than place features (which the less common /ɸ/ and /h/ realisations conform to), but this three-way variation in the realisation of /f/ does not seem to have been discussed in the literature (except for the realisation as /p/ being reported in Ryan, 1972, p. 144). However, Kaʻai and Moorfield (2009, p. 23) do note 'golf' as part of the /g/ → /k/ realisation, and they note the varying realisations of 'wolf' as *wuruwhi* and *wuruhi* as part of a discussion on the four syllables /ɸu/, /ɸo/, /wu/, and /wo/ only being used in te reo Māori in Lehnwörter (Kaʻai & Moorfield, 2009, p. 27).

Voiceless alveolar fricative. If it is not eliminated through a cluster reduction process, when /s/ occurs in the attested consonant clusters it is sometimes realised as /h/ (14 tokens), that is, it is debuccalised, and sometimes as /t/ (one token), that is, it is strengthened to the corresponding coronal plosive. The single token as /t/ is 'slate' /slæet/ → /teɾeti/ *tereti*. The preference for realisation as /h/ over /t/ conforms to the prediction in Lin (2009, p. 9) that manner features for consonants in loanwords are more faithfully retained than place features. The realisation as /h/ is as reported in Ryan (1972, p. 144) and Harlow (2004, p. 160) and the variable realisations as /h/ and as /t/ are as reported in Kearns (1990, pp. 67–68) and Kaʻai and Moorfield (2009, pp. 20–21 & 24).

Having discussed the phonemic correspondences that occur in the consonant clusters of the Williams appendix data, I now set out the principal findings of this study in relation to consonant clusters (section 4.2) and to anaptyxis (section 4.3).

[6] I am grateful to an anonymous reviewer of this paper for reminding me of the possibility of an alveolar realisation of the nasal in the *-ing* suffix of English being the cause of the /n/ token in *rēkena*.

4.2 Consonant clusters

This section of the paper sets out the results relating to consonant clusters, related to research questions (a), (b), and (d). Section 4.2.1 discusses the possible strategies for resolving loanword consonant clusters in te reo Māori. Section 4.2.2 sets out the principal results of the paper relating to consonant clusters. Section 4.2.3 sets out the hierarchy of retention that is manifested by the principal results. Section 4.2.4 briefly discusses some phonetic motivations for the choices made between consonant cluster strategies, and relates specifically to research question (d). Section 4.2.5 discusses some variant realisations of source words that reinforce that conclusions in respect of loanwords cannot be definitive. Section 4.2.6 looks briefly at the specific circumstances of the resolution of the relatively rare (among the source words in the data analysed) three-consonant onsets. Finally, sections 4.2.7 and 4.2.8 look specifically at word-initial and word-final consonant clusters, respectively, given the significance that some of the literature on loanwords has placed on these phonetic environments for how source word consonant clusters are resolved.

4.2.1 Consonant cluster strategies

As discussed in section 2, a major difference between English and te reo Māori phonology is the prohibitions on consonant clusters and on codas in te reo. With Lehnwörter in languages such as te reo Māori that require an open (C)V(V) syllable structure, there are three different strategies for dealing with two-phoneme consonant clusters (that is, CC) in source words, two of which have two variants. These are retaining both consonants; deleting one consonant; and deleting the whole consonant cluster.

Retain both consonants. Firstly, both consonants can be retained. However, a CC structure is not permitted in a language that requires a (C)V(V) syllable structure. There are two sub-strategies for resolving this tension between phonetic similarity and the required syllable structure.

The first sub-strategy is to utilise anaptyxis (that is, epenthesis of a vowel), so that an anaptyctic vowel is inserted between the two consonants, that is, CC → CV.C. This sub-strategy with anaptyxis necessarily produces an additional syllable in the Lehnwort because of the phonotactic requirement to conform to a (C)V(V) syllable structure.

The other sub-strategy when retaining both consonants, which is only an option when the second consonant of the cluster is a semivowel (that is, /j/ or /w/), is to realise the semivowel as a vowel, retaining the relevant front or back feature of the semivowel (that is, realisation as /i/ or /u/, respectively). A Cj consonant cluster is therefore realised as a Ci full syllable and a Cw consonant cluster is therefore realised as a Cu full syllable. This sub-strategy is only relevant to onsets as no English coda contains a semivowel.

Delete one consonant. Secondly, one of the consonants can be deleted and the other retained, that is, CC → C. This also involves two sub-strategies, one where the first consonant is retained and the second is deleted, that is, C_1C_2 → C_1; the other where the second consonant is retained and the first is deleted, that is, C_1C_2 → C_2. This departs from the principle of phonetic similarity to the source word compared to the first strategy of retaining both consonants, but it can facilitate adherence to the principle in other ways, such as maintaining the number of syllables between the source word and the Lehnwort (which anaptyxis can disrupt) and maintaining the primary and other lexical stresses on the same syllables (which both anaptyxis and semivowel conversion can disrupt, as stress in te reo Māori is predictable based on the positions of the long vowels and the diphthongs).

Delete the consonant cluster. Thirdly, the entire consonant cluster can be deleted, that is, CC → ∅. This departs most radically from the principle of phonetic similarity to the source word, but again can facilitate adherence to the principle in relation to the number of syllables and the positions of the lexical stresses.

The required (C)V(V) syllable structure also means that the results are usually different for codas (κ) than they are for onsets (ω). With codas, as they follow the syllabic nucleus (ν), the nucleus is not available to provide the vowel of the required (C)V(V) syllable structure, and paragogic anaptyxis must therefore occur whenever *any* coda consonant from the source word is retained (that is, in both the first and the second strategies). This means that in the first strategy, where both consonants are retained and anaptyxis occurs, anaptyxis must occur twice in codas, as inserted vowels are needed after both retained consonants. The results in codas are therefore usually longer than those in onsets. In the third strategy, as with onsets, the result in codas is still CC → ∅, but in the second strategy it is CC → CV (not CC → C as it is for onsets), and in the first strategy it is CC → CV.CV (not CC → CV.C as it is for onsets).

Harlow (2004, p. 160) observes that in te reo Lehnwörter 'Consonant clusters are either split up by means of epenthetic vowels or simplified by the loss of all but one of the consonants'. Degani (2012, p. 20) similarly notes that 'Since consonant clusters are not allowed in Māori, their integration into Māori involves either the reduction of the consonant cluster to just one consonant or the insertion of epenthetic vowels'. The second strategy as applied to the semivowel-final onsets retains both phonemes of the cluster, without anaptyxis, but instead with conversion of the C_2 semivowel to a vowel (/j/ → /i/ and /w/ → /u/), and is used by 10 tokens; and the third strategy deletes both consonants of the cluster and is used by two tokens. Harlow's and Degani's statements are therefore accurate in respect of 183 of 195 tokens (94% of the data), but not of the remaining 12 tokens. Anaptyxis and cluster reduction to a single consonant are by far the most commonly used strategies, as Harlow (2004, p. 160) and Degani (2012, p. 20) observe, but complete loss of the cluster and retention of the cluster without anaptyxis (but instead with conversion of the C_2 semivowel to a vowel) do also occur.

4.2.2 Principal consonant cluster results

I now turn to the presentation of the principal consonant cluster results. In the analysis of the Williams appendix consonant clusters nine groups of consonant clusters have been identified. These nine groups involve differing rates of utilisation of the three strategies and sub-strategies for conforming source word consonant clusters to the required (C)V(V) syllable structure of te reo Māori.

In the identification of four of these nine groups, those involving sibilants, it is necessary to refer to the sonority sequencing principle (SSP). 'Sonority' refers to the loudness of speech sounds relative to other speech sounds of the same stress, length, and pitch. The sonority sequence (from most to least sonorous phonemes) is vowels > semivowels > liquids > nasals > voiced fricatives > voiceless fricatives > voiced affricates > voiceless affricates > voiced plosives > voiceless plosives.

The SSP is the principle that the sonority of phonemes is at a peak in the syllabic nucleus (ν) and that sonority falls away, progressively, from the nucleus through the phonemes of both the onset (ω) and the coda (κ), with the least sonorous phonemes therefore being those furthest away from the nucleus, that is, the first phoneme of the onset and the last phoneme of the coda (Carlisle, 2001, pp. 4–5; François, 2010, pp. 405–406). Coda consonant clusters that conform to the SSP therefore have their phonemes arranged in the sonority sequence set out above (becoming less sonorous as they increase in distance from the nucleus), whereas onset

consonant clusters that conform to the SSP have their phonemes arranged in the opposite order (so as to similarly become less sonorous as they increase in distance from the nucleus).

The sonority sequence of phonemes from most to least sonorous is, however, somewhat language dependent (Carlisle, 2001, pp. 4–5; François, 2010, pp. 406, 413). English consonant clusters always conform to the SSP, except for (i) in some clusters formed from the addition of a derivational or inflectional morpheme, such as the plural marker, for example, CARD + {plural} → /kɐː**dz**/, where the voiced fricative /z/ is more distant from the nucleus /ɐː/ than the voiced plosive /d/ is, contrary to the SSP; and (ii) clusters of /s/ and a plosive, for example, 'school' /**sk**ʉːl/ and 'tax' /tɛ**ks**/, where the /s/ voiceless fricatives are more distant from the vowel nuclei than the /k/ voiceless plosives are, again contrary to the SSP. This compliance with the SSP except for /s/ + plosive clusters is a feature of English's inheritance from Proto-Indo-European, and is common to many Indo-European languages (François, 2010, p. 413; Carlisle, 2001, pp. 4–5).

The nine identified groups of consonant clusters in the Williams appendix data are:

- liquid-final onsets, for example, 'priest' /**pɹ**iːst/ (table 1);
- semivowel-final onsets, for example, 'quilt' /**kw**ɘlt/ (table 2);
- lateral approximant-led codas, for example, 'milk' /mɘ**lk**/ (table 3);
- nasal-led codas, for example, 'bank' /bɛ**ŋk**/ (table 4);
- two voiceless plosive codas, for example, 'contract' /ˈkɒn.tɪɛ**kt**/ (table 5);
- SSP-ordered sibilant onsets, for example, 'slate' /**sl**æet/ (table 6);
- SSP-ordered sibilant codas, for example, 'cask' /kɐː**sk**/ (table 7);
- non-SSP-ordered sibilant onsets, for example, 'school' /**sk**ʉːl/ (table 8); and
- non-SSP-ordered sibilant codas, for example, 'tax' /tɛ**ks**/ (table 9).

The results of the nine consonant cluster groups are then compared and contrasted in table 10.

The nine groups are not necessarily mutually exclusive, and there are 14 consonant clusters that are common to liquid-final onsets and to SSP-ordered sibilant onsets, that is, /sl/ (one token); to liquid-final onsets and to non-SSP-ordered sibilant onsets, that is, /stɹ/ (two tokens); to semivowel-final onsets and to non-SSP-ordered sibilant onsets, that is, /stj/ (one token); and to nasal-led codas and to SSP-ordered sibilant codas, that is, /ntʃ/ (one token), /nʃ/ (one token), /ns/ (three tokens), /ndʒ/ (three tokens), /nz/ (one token), and /ŋz/ (one token).

The final row in each of tables 1–9 presents (i) a generalisation of the realisations according to the relevant consonant cluster strategy utilised by the tokens in the relevant column; (ii) the totals of the tokens in the relevant column (the sums of the '№' (number) columns); and (iii) the proportions (as percentages) of the total tokens that utilise the relevant strategy of their column compared to all of the tokens of the relevant consonant cluster group.

Table 1. Liquid-final onsets

English	te reo	№	%	te reo	№	%
pɹ	pVɾ	14	100%			
bl	pVɾ	7	100%			
kl	kVɾ	7	100%			
pl	pVɾ	6	100%			
gɹ	kVɾ	5	100%			
kɹ	kVɾ	3	100%			
fl	pVɾ	2	100%			
θɹ	tVɾ	1	100%			
sl	tVɾ	1	100%			
bɹ	pVɾ	9	90%	ɾ	1	10%
tɹ	tVɾ	12	86%	ɾ	2	14%
fɹ	pVɾ	4	80%	ɾ	1	20%
dɹ	tVɾ	3	75%	ɾ	1	25%
gl	kVɾ	4	67%	ɾ	2	33%
stɹ	tVɾ	1	50%	ɾ	1	50%
	P̥Vɾ	79	91%	ɾ	8	9%
V = vowel; P̥ = voiceless plosive						

The liquid-final onset cluster /sl/ (one token) is also represented in table 6 (SSP-ordered sibilant onsets) and /stɹ/ (two tokens) is also represented in table 8 (non-SSP-ordered sibilant onsets).

The liquid-final onsets demonstrate a strong preference for phonetic similarity using anaptyxis (79 of 87 tokens, that is, 91%). This is as noted by Kang (2011, p. 15).

The loss of the obstruent in obstruent + liquid clusters is noted by Ryan (1972, p. 143), who specifically references ‘frying pan’ → *raupani* and ‘trousers’ → *rautete*.

The loss of the obstruent in obstruent + liquid clusters is congruent with the frequent loss of the plosive in plosive-continuant clusters (that is, plosive + continuant clusters and continuant + plosive clusters) set out by Schütz (2004, p. 268) for Fijian Lehnwörter. Fijian is a Central Oceanic language, as is te reo Māori.

Table 2. Semivowel-final onsets

English	te reo	№	%	te reo	№	%	te reo	№	%
dj	ti	2	100%						
mj	mi:	1	100%						
tw	tu	1	100%						
kw	ku	3	75%	k	1	25%			
nj	ni	2	50%	n	2	50%			
bj	pi	1	50%				Ø	1	50%
tj				t	2	100%			
stj				t	1	100%			
	CG → CV	10	59%	occlusive	6	35%	Ø	1	6%
V = vowel; C = consonant; G = semivowel									

The semivowel-final onset cluster /stj/ (one token) is also represented in table 8 (non-SSP-ordered sibilant onsets).

The semivowel-final onsets demonstrate a moderate preference for phonetic similarity via semivowels becoming syllabic (10 of 17 tokens, that is, 59%).

Table 3. Lateral approximant-led codas

English	te reo	№	%	te reo	№	%	te reo	№	%
lk	ɾVkV	2	100%						
lf	ɾVFV	2	100%						
lt				ɾV	1	50%	tV	1	50%
ld				ɾV	1	100%			
	ɾVℂV	4	57%	ɾV	2	29%	tV	1	14%
V = vowel; ℂ = obstruent; F = fricative									

The lateral approximant-led codas demonstrate a moderate preference for phonetic similarity using anaptyxis (four of seven tokens, that is, 57%).

Bauer and Warren (2004, p. 595) note in respect of the lateral approximant in New Zealand English that 'In coda position there is variation between a "darker" lateral, perhap [l^{ʔ}], and a vowel of variable quality'. As a result, it is possible that /soːlt/ is not the pronunciation of 'salt' that was realised as the Lehnwort *tote* /tote/, but rather a pronunciation without an /l/, meaning there was no /lt/ coda consonant cluster for the /l/ to be lost from. If this is correct, all lateral approximant-led codas are realised either through anaptyxis or through cluster reduction to the liquid, as expected.

The loss of the obstruent in lateral approximant-led clusters is noted by Ryan (1972, p. 142), who specifically references 'gold' → *koura* and 'quilt' → *kuira*.

The loss of the obstruent in obstruent + liquid clusters is congruent with the frequent loss of the plosive in plosive-continuant clusters (that is, plosive + continuant clusters and continuant + plosive clusters) set out by Schütz (2004, p. 268) for Fijian Lehnwörter. Fijian is a Central Oceanic language, as is te reo Māori.

Table 4. Nasal-led codas

English	te reo	№	%	te reo	№	%	te reo	№	%
ntʃ	nVhV	1	100%						
nʃ	nVhV	1	100%						
ns	nVhV	2	67%	nV	1	33%			
ndʒ	nVhV	1	33%	nV	2	67%			
ŋk	ŋVkV	1	33%				kV	2	67%
nt	nVtV	5	29%	nV	5	29%	tV	7	41%
nd				nV	6	100%			
nz				nV	1	100%			
ŋz				nV	1	100%			
mp							Ø	1	100%
	NVℂV	11	30%	nV	16	43%	PV *or* Ø	10	27%
V = vowel; N = nasal; ℂ = obstruent; P = plosive									

The nasal-led coda clusters /ntʃ/ (one token), /nʃ/ (one token), /ns/ (three tokens), /ndʒ/ (three tokens), /nz/ (one token), and /ŋz/ (one token) are also represented in table 7 (SSP-ordered sibilant codas).

The nasal-led codas demonstrate a moderate preference for cluster reduction (26 of 37 tokens, that is, 70%), usually via deletion of the obstruent (16 of 37 tokens, that is, 43%).

Table 5. Two voiceless plosive codas

English	te reo	№	%
kt	kV	1	100%
pt	pV	1	100%
	P_1V	2	100%
V = vowel; P = plosive			

The two voiceless plosive codas always involve cluster reduction via deletion of the second plosive.

Table 6. SSP-ordered sibilant onsets

English	te reo	№	%	te reo	№	%
sl	tVɾ	1	100%			
sm				m	1	100%
sn				n	1	100%
	tVɾ	1	33%	N	2	67%
V = vowel; N = nasal						

The SSP-ordered sibilant onset cluster /sl/ (one token) is also represented in table 1 (liquid-final onsets).

The SSP-ordered sibilant onsets demonstrate a moderate preference for cluster reduction via deletion of the sibilant (two of three tokens, that is, 67%). The variation between cluster reduction via deletion of the sibilant and anaptyxis, with the former the preferred strategy, is as noted by Kang (2011, p. 15).

Table 7. SSP-ordered sibilant codas

English	te reo	№	%	te reo	№	%	te reo	№	%
ntʃ	nVhV	1	100%						
nʃ	nVhV	1	100%						
ns	nVhV	2	67%	nV	1	33%			
ndʒ	nVhV	1	33%	nV	2	67%			
st	hVtV	1	13%	tV	4	50%	hV	3	38%
nz				nV	1	100%			
ŋz				nV	1	100%			
sk							hV	1	100%
	CV∁V	6	32%	occlusiveV	9	47%	hV	4	21%
V = vowel; C = consonant; ∁ = obstruent									

The SSP-ordered sibilant coda clusters /ntʃ/ (one token), /nʃ/ (one token), /ns/ (three tokens), /ndʒ/ (three tokens), /nz/ (one token), and /ŋz/ (one token) are also represented in table 4 (nasal-led codas).

The SSP-ordered sibilant codas also demonstrate a moderate preference for cluster reduction (13 of 19 tokens, that is, 68%), usually via deletion of the sibilant (nine of 19 tokens, that is, 47%).

Commenting on the loss of phonemes in the realisation of consonant clusters in Fijian, Schütz (2004, p. 284) observes that:

> *Some of the examples of the simplification of English final consonant clusters given earlier might be instead instances of interpreting* -s *as a marker of plural. If so, the borrower analogised from his/her knowledge of English morphology. Examples are* kasi.tama/kasi.taba *'customs',* lai.seni *'license'.*

If this had occurred among Māori speakers with a good enough knowledge of English morphology to understand the English plural marker, which could have been a significant number given the early development of bilingualism among Māori speakers, then the deletion of the sibilant from coda clusters ending in /s/ and /z/ is perhaps explained. The Williams appendix contains 11 such tokens, including (as with Schütz, 2004, p. 284) 'license', and three actual plurals, namely 'cards', 'rates', and 'leggings', as set out at (1) below.

(1) Coda clusters ending in /s/ and /z/
'box' /bɒ**ks**/ → /poua**k**a/ *pouaka*
'box' /bɒ**ks**/ → /pa:**k**i/ *pāki*
'box' /bɒ**ks**/ → /pa:**k**a/ *pāka*
'tax' /tɛ**ks**/ → /ta:**k**e/ *tāke*
'cards' /kɐ:**dz**/ (incorporates {plural}) → /ka:**ɾ**i/ *kāri*
'Scotsman' /ˈskɒ**ts**.mən/ → /ka:**t**imana/ *Kātimana*
'rates' /ɹæe**ts**/ (incorporates {plural}) → /ɾei**t**i/ *reiti*
'Wednesday' /ˈwe**nz**.dæe/ → /we**n**eɾei/ *Wenerei*
'wax' /wɛ**ks**/ → /wa:**k**i/ *wāki*
'leggings' /ˈleg.ə**ŋz**/ (incorporates {plural}) → /ɾe:ke**n**a/ *rēkena*
'license' /ˈlɑe.sə**ns**/ → /ɾaiha**n**a/ *raihana*

Table 8. Non-SSP-ordered sibilant onsets

English	te reo	№	%	te reo	№	%
stɹ	tVɾ	1	50%	ɾ	1	50%
sk	hVk	1	25%	k	3	75%
st	hVt	2	12%	t	15	88%
sp				p	4	100%
stj				t	1	100%
	ℂVC	4	14%	C	24	86%
V = vowel; C = consonant; ℂ = obstruent						

The non-SSP-ordered onset cluster /stɹ/ (two tokens) is also represented in table 1 (liquid-final onsets) and /stj/ (one token) is also represented in table 2 (semivowel-final onsets).

The non-SSP-ordered sibilant onsets demonstrate a strong preference for cluster reduction, losing the sibilant (24 of 28 tokens, that is, 86%). The variation between cluster reduction via deletion of the sibilant and anaptyxis, with the former the preferred strategy, is as noted by Kang (2011, p. 15).

Table 9. Non-SSP-ordered sibilant codas

English	te reo	№	%	te reo	№	%
ks	kVhV	3	50%	kV	3	50%
ts				tV	2	100%
dz				ɾV	1	100%
	kVhV	3	33%	CV	6	67%
V = vowel; C = consonant						

The non-SSP-ordered sibilant codas demonstrate a moderate preference for cluster reduction via deletion of the sibilant (six of nine tokens, that is, 67%).

Table 10 below summarises and compares the results of the nine groupings of consonant clusters.

Table 10. Cluster comparisons

	Anaptyxis/semivowel conversion		Preferred consonant		Dispreferred consonant/total deletion	
Cluster type		№: %		№: %		№: %
liquid-final onsets	P̥Vɾ	79: 91%	ɾ	8: 9%		
semivowel-final onsets	CG → CV	10: 59%	occlusive	6: 35%	Ø	1: 6%
lateral approximant-led codas	ɾVȻV	4: 57%	ɾV	2: 29%	tV	1: 14%
nasal-led codas	NVȻV	11: 30%	nV	16: 43%	PV *or* Ø	10: 27%
two voiceless plosive codas			P_1V	2: 100%		
SSP-ordered sibilant onsets	tVɾ	1: 33%	N	2: 67%		
SSP-ordered sibilant codas	CVȻV	6: 32%	occlusiveV	9: 47%	hV	4: 21%
non-SSP-ordered sibilant onsets	ȻVC	4: 14%	C	24: 86%		
non-SSP-ordered sibilant codas	kVhV	3: 33%	CV	6: 67%		
total		118		75		16

V = vowel; C = consonant; G = semivowel; N = nasal; Ȼ = obstruent; P = plosive; P̥ = voiceless plosive

I now look to summarise these results. The figures given here are not directly comparable to those set out in the summary Table 10 above because of the 14 tokens that are double counted in Table 10 as they are represented in more than one of the nine identified groups of consonant clusters. There are, in other words, 195 tokens in total, but Table 10 sums to 209 as a result of the cross-group double representations.

Phonetic similarity, retaining both of the phonemes of the consonant cluster, is the most common strategy, representing 57% of the data (111 of 195 tokens). This is, however, much more pronounced for the onsets, where 92 of 131 tokens (70%) retain both phonemes of the

source word, compared to the codas, where only 19 of 64 tokens (30%) do. This much stronger preference for phonetic similarity in onsets is probably to be expected given both the known greater saliency of the beginnings of words for lexical recognition and the prohibition of codas in te reo Māori.

Three groups of consonant clusters exhibit a preference for phonetic similarity, retaining both of the phonemes of the consonant cluster: liquid-final onsets, semivowel-final onsets, and lateral approximant-led codas. Although this is only three of the nine groups of clusters, these three groups represent 57% of the data (111 of 195 tokens).

The other six groups of consonant clusters exhibit a preference for cluster reduction, to different degrees. The two voiceless plosive codas always delete the second plosive in the cluster, although this result is unlikely to be meaningful given that there are only two tokens in this very small group.

The nasal-led codas are the most varied of the groups. They exhibit a moderate preference for cluster reduction (26 of 37 tokens, that is, 70%) but vary in how they achieve this, with 16 tokens retaining the nasal, nine tokens retaining the plosive, and one token deleting the whole cluster. Overall, however, the nasal is usually retained (27 of 37 tokens, that is, 73%).

The sibilant clusters generally delete the sibilant: 41 out of 59 tokens (69%). This preference to delete the sibilant perhaps reflects the complete lack of sibilants in the phonemic inventory of te reo. However, where the position of the English sibilant unusually does not break the sonority sequencing principle, the preference to delete the sibilant is eliminated, with 11 out of 22 tokens (50%) deleting the sibilant. Where the position of the English sibilant does break the sonority sequencing principle, as is more common, it is much more commonly deleted: 30 out of 37 tokens (81%). Deletion of the sibilant is more common in sibilant onsets (26 of 31 tokens, that is, 84%) than it is in sibilant codas (15 of 28 tokens, that is, 54%). Where one consonant is deleted in sibilant-led clusters, it is always the sibilant that is deleted.

This loss of the sibilant in sibilant-led clusters is noted by Ka‘ai and Moorfield (2009, p. 27) and is a wider realisation of the /sp/ → /p/ correspondence noted by Ryan (1972, p. 144). However, it contrasts with the frequent loss of the plosive in plosive-continuant clusters (that is, plosive + continuant clusters and continuant + plosive clusters) set out by Schütz (2004, p. 268) for Fijian Lehnwörter. Fijian is a Central Oceanic language, as is te reo Māori. Source word sibilant + plosive clusters are realised as sibilants in Fijian but as plosives in te reo Māori.

I now examine the relative sonority of the cluster types, and have grouped some of the nine groups together in order to be able to do so more coherently, as follows:

- the liquid-final onsets and the lateral approximant-led codas are grouped together as “liquid clusters”;
- the SSP-ordered sibilant onsets and the SSP-ordered sibilant codas are grouped together as “sibilant clusters (SSP-ordered)”; and
- the non-SSP-ordered sibilant onsets and the non-SSP-ordered sibilant codas are grouped together as “sibilant clusters (non-SSP-ordered)”.

Table 11 below then sets out the cluster types by sonority. It indicates that:

(i) Overall, the consonant clusters with the more sonorous phonemes (that is, semivowels and liquids) are likely to adhere to the principle of phonetic similarity and to retain both phonemes, whereas the consonant clusters with the less sonorous phonemes (occlusives and obstruents) are likely to be resolved through cluster reduction or deletion.

(ii) With the exception of the semivowel clusters, where the consonant clusters with the more sonorous phonemes exhibit cluster reduction, they are much more likely to keep the more sonorous phoneme (that is, the liquid or the nasal).

(iii) With the clusters with the less sonorous sibilant phonemes, where these clusters exhibit cluster reduction, it is very much more likely that the more sonorous phoneme is the one retained when the source word cluster adheres to the sonority sequencing principle (SSP) than when it breaks it.

Table 11. Cluster types by sonority

		Retain both phonemes		Retain more sonorous phoneme in cluster reduction	
Sonority	Cluster type	№	%	№	%
most	semivowel clusters	10 of 17	59%	0 of 7	0%
↑	liquid clusters	83 of 94	88%	10 of 11	91%
	nasal clusters	11 of 37	30%	16 of 26	62%
	sibilant clusters (SSP-ordered)	7 of 22	32%	11 of 15	73%
↓	sibilant clusters (non-SSP-ordered)	7 of 37	19%	1 of 30	3%
least	two plosive clusters	0 of 2	0%	—	—

The exception to these general principles is the semivowels, that is, the group of the most sonorous consonant phonemes. This group is distinct from the others in using the conversion of semivowels to full vowels (that is, syllabification) to achieve retention of both phonemes rather than using anaptyxis. Semivowel clusters do exhibit the sonorous cluster preference for retaining both phonemes (10 of 17 tokens, that is, 59%), but they do so at a lower rate than the less sonorous liquid clusters (88%), and, where they do exhibit cluster reduction, the semivowel is the phoneme that is always deleted, with the much less sonorous occlusive retained.

Having presented the main results relating to consonant clusters, the following four sections draw a key implication relating to a hierarchy of retention (section 4.2.3); set out the likely phonetic motivations of some of the choices between anaptyxis and cluster reduction (section 4.2.4); note some variant realisations that utilise multiple strategies (section 4.2.5); and look at the conformity of the three-consonant consonant clusters (that is, CCC) to the general results (section 4.2.6).

4.2.3 Hierarchy of retention

The results set out in section 4.2.2 above indicate that there is a hierarchy of retention of consonants. Preferably both source word phonemes of the consonant cluster are retained, to ensure phonetic similarity to the source word. However, if only one consonant is retained, this is generally determined according to the hierarchy of retention liquid > occlusive > semivowel and fricative. So liquids are preferably retained over occlusives (plosives, nasals, and affricates); and liquids and occlusives are preferably retained over semivowels and fricatives.

This hierarchy also fully accounts for the three three-consonant consonant cluster tokens in the Williams appendix, that is, ‘steward’ /ˈ**stj**ʉː.əd/ → /**t**uaɾi/ *tuari* (occlusive realised from fricative + occlusive + semivowel); ‘strawberry’ /ˈ**st.ɹ**oː.bəɹ.i/ → /**ɾ**oːpeɾe/ *rōpere* (liquid realised from fricative + occlusive + liquid); and ‘street’ /**st.ɹ**iːt/ → /**tiɾ**iti/ *tiriti* (occlusive and liquid realised from fricative + occlusive + liquid).

This hierarchy of retention perhaps also explains why the liquid clusters and (to a lesser extent) the semivowel clusters predominantly adhere to the principle of phonetic similarity, retaining both source word phonemes, while the sibilant clusters predominantly exhibit cluster reduction, retaining the non-sibilant consonant only: because fricatives are the least preferred class of phonemes for retention and realisation. This phenomenon in turn perhaps reflects the paucity of fricatives in te reo Māori (only /ɸ/ and /h/) and the complete absence of sibilants.

4.2.4 Phonetic motivations

The choice between anaptyxis and cluster reduction can sometimes be seen to be phonetically motivated. Realisations of the /bɹ/ cluster are usually (nine tokens) /pVɾ/, that is, anaptyxis has occurred, in line with the requirement for phonetic similarity, but cluster reduction to /ɾ/ is used for 'umbrella' /ɐmˈ**bɹ**el.ə/ → /ma**ɾ**aɾa/ *marara*, where the immediately preceeding /m/, retained in the Lehnwort, provides the labial feature that realisation of the /b/ as /p/ would usually have provided, while allowing realisation of the three-syllable source word as a three-syllable Lehnwort.

Similarly, realisations of the /dɹ/ cluster are usually (three tokens) /tVɾ/, that is, anaptyxis has occurred, but cluster reduction to /ɾ/ is used for 'hundredweight' /ˈhɐn.**dɹ**əd.wæet/ → /hana**ɾ**ete/ *hanarete*, where the immediately preceeding /n/ provides the coronal feature that realisation of the /d/ as /t/ would usually have provided, but without introducing an additional syllable.

Again, realisations of the /gl/ cluster are usually (four tokens) /kVɾ/, that is, anaptyxis has occurred, but cluster reduction to /ɾ/ is used for 'singlet' /ˈsəŋ.**gl**ət/ → /hiŋa**ɾ**eti/ *hingareti* and for 'English' /ˈəŋ.**gl**əʃ/ → /iŋa**ɾ**ihi/ *Ingarihi*, where the immediately preceeding /ŋ/ provides the dorsal feature that realisation of the /g/ as /k/ would usually have provided, but without introducing additional syllables.

4.2.5 Variant realisations

As noted, strategies with Lehnwörter are not fully predictable (compare the five realisations of 'Britain' as *Piritini*, *Piriteni*, *Piritene*, *Piritani*, and *Piritana* noted in section 1), and there are variant Lehnwörter in the source data that utilise multiple strategies.

For '**trousers**' /ˈ**tɹ**æʉ.zəz/ the Williams appendix has both /**ɾ**autete/ *rautete* (cluster reduction to the liquid) and /**tar**au.te.te/ *tarautete* (anaptyxis).

For '**quarter**' /ˈ**kw**oː.tə/ the Williams appendix has /**k**oata/ *koata* (cluster reduction to the plosive), but *Te Aka Māori Dictionary* has this form and also /**ku**ata/ *kuata* (semivowel-to-vowel conversion).

For '**newspaper**' /ˈ**nj**ʉːsˌpæe.pə/ the Williams appendix has /**n**uːpepa/ *nūpepa* (cluster reduction to the nasal), but *Te Aka Māori Dictionary* has this form and also /**ni**upepa/ *niupepa* (semivowel-to-vowel conversion).

For '**New Year**' /ˌ**nj**ʉːˈjiə/ the Williams appendix has /**n**uia/ *Nu Ia* (cluster reduction to the nasal), but *Te Aka Māori Dictionary* has /**n**uːia/ *Nū Ia* (cluster reduction to the nasal, with a long vowel) and /**ni**uia/ *Niu Ia* (semivowel-to-vowel conversion).

4.2.6 Three-consonant clusters (onsets)

With three-consonant clusters there are too few data to reach any significant conclusions specifically in relation to these clusters. There are only three tokens in the Williams appendix,

from the two clusters /stɹ/ and /stj/, both onsets, namely ‘strawberry’ /ˈ**stɹ**oː.bəɹ.i/ → /**ɾ**oːpeɾe/ *rōpere*; ‘street’ /**stɹ**iːt/ → /**tiɾ**iti/ *tiriti*; and ‘steward’ /ˈ**stj**ʉː.əd/ → /**t**uaɾi/ *tuari*. These each use different strategies to reduce their clusters to /ɾ/, /tiɾ/, and /t/. It can, however, probably be concluded that three-consonant clusters are too long to involve full retention and anaptyxis between all three consonants, as this would involve the creation of three syllables out of a single onset.

The possibility of the loss of the initial consonant or consonant cluster is noted by Ryan (1972, p. 142), who specifically references ‘strawberry’ → *rōpere*.

Having presented the main and minor results in relation to the consonant clusters, the remaining two sections in section 4 address a number of minor findings that reinforce or contrast with the observations in Kang (2011, pp. 14–15) in relation to word-initial consonant clusters (section 4.2.7) and word-final consonant clusters (section 4.2.8).

4.2.7 Word-initial consonant clusters

Kang (2011, p. 14) observes that:

> Several languages use a deletion strategy in combination with epenthesis word-initially. Polynesian languages that lack /s/ in their inventory delete /s/ in /s/-initial clusters while all other clusters are repaired by epenthesis. Examples include Maori (frock → [poraka], Scotsman → [koːtimana]; Harlow 2004; Herd 2005)….

This observation would indicate that, word initially, only the third strategy is used, unless the consonant cluster starts with /s/. Ignoring the onset consonant clusters where there is neither deletion nor anaptyxis, that is, those where the semivowel is realised as a vowel, this rule of Kang’s in relation to Polynesian languages is generally correct with regard to te reo, although the Williams appendix data does have eight counter examples, one of a word-initial cluster where the intial /s/ is not deleted and seven of other word-initial clusters that feature consonant deletion rather anaptyxis. Of these seven, two involve the loss of an obstruent and the retention of a liquid, and five involve the loss of a semivowel and the retention of an occlusive, as set out at (2) below.

(2) Counter-examples to Kang (2011, p. 14)
‘slate’ /**sl**æet/ → /**t**eɾeti/ *tereti*
‘trousers’ /ˈ**tɹ**æʉ.zəz/ → /**ɾ**autete/ *rautete*
‘frying pan’ /ˈ**fɹ**ɑe.əŋˌpɛn/ → /**ɾ**aupani/ *raupani*
‘quarter’ /ˈ**kw**oː.tə/ → /**k**oata/ *koata*
‘Tunisia’ /**tj**ʉːˈnəz.i.ə/ → /**t**unihia/ *Tunihia*
‘Tuesday’ /ˈ**tj**ʉːz.dæe/ → /**t**uːɾei/ *Tūrei*
‘newspaper’ /ˈ**nj**ʉːsˌpæe.pə/ → /**n**uːpepa/ *nūpepa*
‘New Year’ /ˌ**nj**ʉːˈjiə/ → /**n**uia/ *Nu Ia*

Kang (2011, p. 14) is therefore generally correct in respect of te reo Māori, but word-initial /s/ deletion from /s/-led onsets and anaptyxis of all other word-initial onsets are strong tendencies, not a rule.

4.2.8 Word-final consonant clusters

In relation to word-final coda consonant clusters Kang (2011, p. 15) observes that:

> Yet another group of languages show both epenthesis and deletion depending on the segmental composition of the offending structure. These languages include … Maori (Harlow 2004)….

The Williams appendix data support this, but as Kang's statement includes all possibilities relevant to codas and do not exclude any possibilities (as semivowel-to-vowel conversion is only possible in onsets, as semivowels do not occur in English codas) this is inevitable. Kang's point, however, is to compare the group of languages (including te reo Māori) that do allow both anaptyxis and cluster reduction word finally to those that do not, not to analyse how word-final consonant clusters are resolved in Māori Lehnwörter.

4.3 Anaptyxis

This section of the paper sets out the results relating to anaptyxis (that is, the insertion of an epenthetic vowel), related to research questions (c) and (d). Section 4.3.1 discusses three possible anaptyxis strategies for resolving loanword consonant clusters in te reo Māori. Section 4.3.2 discusses the first strategy, default anaptyxis; section 4.3.3 discusses the second strategy, copy anaptyxis; and section 4.3.4 discusses the third strategy, feature spreading. Section 4.3.5 compares how well the different anaptyxis strategies fit the Williams appendix data.

4.3.1 Anaptyxis strategies

As shown in the consonant cluster results, anaptyxis is a common strategy to separate the consonants in source word consonant clusters and to append to coda consonants to achieve open syllables, thereby ensuring that the Lehnwort conforms to the required (C)V(V) syllable structure of te reo Māori. Given that the anaptyctic vowel is not present in the source word, this raises the question how the *particular* vowel to be inserted is determined. In other words, how are the height, backness, and length of the anaptytic vowel to be determined if there is no starting point for this phoneme of the Lehnwort in the source word (as there are for the consonants and for the non-anaptyctic vowels)?

> Schütz (2004, p. 269) poses the question in relation to his research on Fijian thus:
>
> vowels have been inserted or added to insure that all Fijian syllables are open. But how are these vowels chosen? …… The important question with respect to an added vowel is this: is the choice of vowel random or patterned?

The assumption that Schütz goes on to make is that anaptyxis is patterned, that is, that it is rule-governed. I make the same assumption and this question regarding how the choice of the specific anaptyctic vowel to be inserted is made is one of this paper's research questions.

Posing this question with specific reference to te reo Māori, how is it determined whether it is /i/, /u/, /e/, /o/, or /a/; or even /iː/, /uː/, /eː/, /oː/, or /aː/; or indeed even a diphthong that realises the need for an anaptyctic vowel, when there is nothing in the source word for the Lehnwort to try to match? How is the principle of phonetic similarity to be observed, when there is no starting point for the recipient language's phonemic choice, and nothing to be faithful to?

The literature on anaptyxis (Kearns, 1990; Kitto & de Lacy, 1999; Schütz, 2004; Uffmann, 2006) indicates that there are three basic strategies for this.

(i) Default anaptyxis, where the same anaptyctic vowel is used each time (or the same small group of anaptyctic vowels, chosen for fit relevant to the receiving phonetic environment, are used each time).
(ii) Copy anaptyxis, where the anaptyctic vowel is copied from a vowel in the source word (modified as needed to conform to the recipient language's phonemic inventory). This is usually the immediately following one (regressive or anticipatory anaptyxis), most commonly used for onsets, or the immediately preceding one (progressive or persevatory anaptyxis), most commonly used for codas.
(iii) Feature spreading, where certain features of a consonant, or a consonant cluster, usually the one immediately preceding the anaptyctic vowel, determine that vowel's features.

These three anaptyxis strategies are summarised in Table 12 below. The feature spreading is assumed in the schema to be from the immediately preceding consonant. The anaptyctic vowels in the Lehnwort are **bolded**. C_1 and C_2 represent the first and second consonants of the consonant cluster; V_α represents an unspecified vowel of the nucleus (ν) of the source word and its anaptyctic copy or copies; V_δ represents a default anapyctic vowel; and V_1 and V_2 represent anaptyctic vowels derived from copying features of the similarly numbered consonants.

Table 12. Anaptyxis strategies

Anaptyxis strategy	Source word	Lehnwort
Onsets		
default vowel	$C_1C_2V_\alpha$	$C_1\mathbf{V_\delta}.C_2V_\alpha$
copy vowel (regressive anaptyxis)	$C_1C_2V_\alpha$	$C_1\mathbf{V_\alpha}.C_2V_\alpha$
feature spreading	$C_1C_2V_\alpha$	$C_1\mathbf{V_1}.C_2V_\alpha$
Codas		
default vowel	$V_\alpha C_1C_2$	$V_\alpha.C_1\mathbf{V_\delta}.C_2\mathbf{V_\delta}$
copy vowel (progressive anaptyxis)	$V_\alpha C_1C_2$	$V_\alpha.C_1\mathbf{V_\alpha}.C_2\mathbf{V_\alpha}$
feature spreading	$V_\alpha C_1C_2$	$V_\alpha.C_1\mathbf{V_1}.C_2\mathbf{V_2}$

As can be seen from this schema, using anaptyxis increases the number of syllables, significantly so in the case of codas, which is why anaptyxis is not always the preferred strategy for addressing consonant clusters, with cluster reduction also being used (see section 3 for the results relating to retention and deletion of consonant cluster phonemes).

4.3.2 Default anaptyxis

Default anaptyxis, if used in te reo Māori, would seem to involve /a/ being the default vowel, as it is by far the most common anaptyctic vowel used, being 69 out of 169 tokens (41%). This compares to 50 tokens of /i/, 29 tokens of /e/, 14 tokens of /o/, five tokens of /u/, and two tokens of /a:/ as an anaptyctic vowel (with no tokens of the other long vowels, that is, /i:/, /e:/, /o:/, and /u:/, as anaptyctic vowels), as set out in table 13 below. Table 13 also compares these data to the frequencies of the vowels in te reo Māori in general, as reported by Harlow (2007, p. 68).

Table 13. Frequencies of vowels in te reo Māori

source	i	i:	u	u:	e	e:	o	o:	a	a:	total
te reo in general	20%	1%	11%	1%	15%	1%	10%	3%	31%	8%	100%
Williams appendix	30%	—	3%	—	17%	—	8%	—	41%	1%	100%

The more common the vowel is in te reo in general, the more likely it is to be used as an anaptyctic vowel: /a/: 31% → 41%; /i/: 20% → 30%; /e/: 15% → 17%. Conversely, the less common the vowel is in te reo in general, the less likely it is to be used as an anaptyctic vowel: /u/: 11% → 3%; /o/: 10% → 8%; /aː/: 8% → 1%; /oː/: 3% → not utilised; /iː/, /uː/, and /eː/: 1% each → not utilised.

This is an example of the Matthew effect, a common phenomenon found across very many sciences. It was named by Merton (1968) after Matthew 25:29: "For unto every one that hath shall be given, and he shall have abundance: but from him that hath not shall be taken away even that which he hath".

However, although /a/ is by far the most common choice as an anaptyctic vowel, default anaptyxis leaves some 59% of the data not accounted for. It therefore seems unlikely that default anaptyxis is the sole or primary strategy used to determine the anaptyctic vowel of te reo Māori Lehnwörter, even if a bias towards use of the most common vowels, and particularly /a/, can be seen in the data, comforming to the Matthew effect.

4.3.3 Copy anaptyxis

Copy anaptyxis is the strategy to copy the following vowel (regressive or anticipatory anaptyxis) into onset consonant clusters and to copy the preceding vowel (progressive or persevatory anaptyxis) into coda consonant clusters. Kitto and de Lacy (1999, p. 14) note that in Māori Kūki 'Āirani (Cook Islands Māori) 'epenthetic vowels are usually copies'. If copy anaptyxis is similarly used in te reo Māori, it would account for 110 of 169 tokens (65%). This is much more satisfactory than for default anaptyxis, accounting for the majority rather than a minority of the data; but it still leaves 35% of the data not accounted for. However, the difference between onsets and codas with this strategy is noteworthy. Of the onsets, copy anaptyxis accounts for 69 of the 86 tokens (80%). This compares to the codas, where copy anaptyxis only accounts for 41 of the 83 tokens (49%).

This difference in the results between onsets and codas may invite the enquiry as to whether it is correct that copy anaptyxis is in fact progressive in codas, and whether regressive anaptyxis in both onsets and codas is more explanatory. However, this is not the case. Regressive anaptyxis in codas would explain only 1 token out of 83 (1%). The anaptyctic vowel in this token is /e/, which provides 14 of the 83 coda tokens (17%) and 29 of the 169 total tokens (also 17%), so this 1% result for regressive anaptyxis in codas is very much worse than chance. Similarly, looking (for completeness) at progressive anaptyxis in onsets, this would explain only 3 tokens out of 85 (4%), which is again very much worse than chance. The logical order of regressive anaptyxis in onsets and progressive anaptyxis in codas is therefore correct; and this accounts for 110 of 169 tokens (65%).

In this regard Kitto and de Lacy (1999, p. 11) observe that:

> … the most common reason why the usual direction of copying is prevented is that in some environments there is simply nothing to copy. For example, epenthetic vowels in Hawaiian (Polynesian) loanwords are copied from the left, but when epenthesis is initial there is nothing appropriate to copy, so direction reverses:

I consider that this observation is only partially correct (unless Hawaiian works substantially differently in this regard from te reo Māori). It seems to me that the location of the anaptyctic vowel within an onset or coda source word cluster, or after a coda consonant, is what determines whether the copying is regressive or progressive, not whether or not it is word initial. Anaptyctic copying in medial onsets is still regressive. For example, in 'supreme court' /sʉːˌpɨːmˈkoːt/ → /huːpɨrɨmikoːti/ *hūpirimi kooti*, the anaptyctic vowel /i/, used to divide the

/pɹ/ consonant cluster into separate syllables, is copied regressively from the nucleus of the syllable /pɹ**i**ːm/, not progressively from that of the preceding syllable /s**u**ː/. It is not the case that anaptyctic copying in initial onsets is regressive but that in medial onsets it is progressive, as Kitto and de Lacy (1999, p. 11) imply, and the syllable is the relevant unit for the operation of anaptyxis, not the word.

4.3.4 Feature spreading

Feature spreading is a term that can cover a wide range of phonological processes. Its precise nature as applied to anaptyxis depends upon which features are spread, and from which consonant(s). This research has looked at several possibilities for the operation of feature spreading to explain anaptyxis in te reo Māori Lehnwörter, all of which have proven to be unsatisfactory. All of the explored possibilities assumed that features are spread progressively to the anaptyctic vowel from the immediately preceding (single) consonant. However, an analysis that matched the differential analysis that I have stated to be logical for copy anaptyxis, with regressive anaptyxis in onsets and progressive anaptyxis in codas, could possibly have produced better results; and other feature spreading possibilities, such as from consonant clusters (as opposed to single consonants), could at least in theory also be possible. These possibilities are left to future research.

In terms of feature spreading from the immediately preceding single consonant, four possibilities were explored:

(i) Spreading of features from front consonants to front vowels and from back consonants to back vowels.
(ii) Spreading of features from acute consonants to acute vowels and from grave consonants to grave vowels.
(iii) Spreading of features from compact consonants to compact vowels and from diffuse consonants to diffuse vowels.
(iv) Spreading of features from labial, coronal, dorsal, and laryngeal consonants to front, central, and back vowels.

All of the possibilities have proved problematic for analyses of te reo Māori. Use of acute and grave features falters on the established lack of either feature in specifying /a/ (that is, the central vowel of the vowel system) in te reo (Hohepa, 1967, pp. 7–8; Harlow, 2007, pp. 63–64). This contrasts with analyses of /a/ in other languages, such as Turkish, where /a/, despite being phonetically central (Zimmer & Orgun, 1999, p. 155) is usually regarded as a back vowel because of how it acts in vowel harmony processes, forming the unrounded equivalent of /o/ (Göksel & Kerslake, 2005, p. 10). Use of front and back features is problematic for the same reason, as /a/ is analysed as central (so neither front nor back). Conversely, use of compact and diffuse features falters on the established presence of both features in specifying /e/ and /o/ (that is, the mid vowels of the vowel system) in te reo (Hohepa, 1967, pp. 7–8; Harlow, 2007, pp. 63–64).

Use of labial, coronal, dorsal, and laryngeal place features is problematic given that all vowels are dorsal, and even though there is a possible association of coronal consonants and front vowels (Hume, 1996), this does not resolve how labial, dorsal, and laryngeal consonants associate with central and back vowels. All possible associations of labial, coronal, dorsal, and laryngeal consonants with front, central, and back vowels have been examined, without any clear association being established. This results from the fairly balanced distribution of the vowels with regarding to the preceding consonants, as set out in Table 14 below. For simplicity

these associations do not take into account allophonic variation in te reo (the most significant aspects of which are outlined in section 2.2.2). This allophonic variation could see the bilabial fricative /ɸ/ (when realised as [θ], i.e., as a dental) included in the coronal group and the glottal fricative /h/ (when realised as [ç], i.e, as a palatal, [x], i.e., as a velar, or as [χ], i.e., as a uvular) included in the dorsal group (Hohepa, 1965, pp. 9–10). However, neither of these adjustments result in any great improvement in the associations involved.

This lack of any association can be most clearly seen, and is most attributable to, the very balanced distribution of /a/, which (as set out above in section 4.3.2) is 41% of the data (or 42% if /a:/ is also included). This is 29% labial, 41% coronal, 25% dorsal, and 6% laryngeal.

Table 14. Associations between consonant and following vowel

		Vowel						
		Front		Central		Back		
Consonant		i	e	a	a:	u	o	Total
Labial	p	6	10	19	0	4	8	47
	m	0	0	0	0	0	0	0
	ɸ	0	0	1	0	0	0	1
Coronal	t	12	9	15	2	0	0	38
	n	10	7	9	0	0	0	26
	ɾ	1	0	4	0	1	1	7
Dorsal	k	8	3	17	0	0	4	32
	ŋ	1	0	0	0	0	0	1
	w	0	0	0	0	0	0	0
Laryngeal	h	12	0	4	0	0	1	17
	Total	50	29	69	2	5	14	169

4.3.5 Comparison of anaptyxis strategies

Default anaptyxis, if used in te reo Māori, would seem to involve /a/ being the default vowel, as it is by far the most common anaptyctic vowel used, being 69 out of 169 tokens (41%). Copy anaptyxis, if used in te reo Māori, would account for 110 of 169 tokens (65%), a much more satisfactory result as it explains a majority, not a minority, of the data. The difference between onsets and codas with a strategy of copy anaptyxis is nevertheless important. Copy anaptyxis accounts for 69 of the 86 onset tokens (80%) but only for 41 of the 83 coda tokens (49%). A third possibility, that of feature spreading from the immediately preceding consonant, was not found to be satisfactory as an analysis tool for te reo Māori given the established analyses of the distinctive features of te reo (Hohepa, 1967, pp. 7–8; Harlow, 2007, pp. 63–64) and the lack of any consistent association between consonants and following vowels. However, the possibility of other types of feature spreading operating, such as vowel harmony or feature spreading from following consonants or from consonant clusters or from non-adjacent consonants cannot be completely discounted.

These results suggest that copy anaptyxis is the preferred strategy, being the best fit with the data. This is reinforced by the finding that copy anaptyxis explains 69 of the 86 onset tokens (80%). Nevertheless, even this best-fit strategy still leaves 35% of the data not accounted for, and 51% of the coda data unaccounted for, which is a lot to have unexplained.

The literature on anaptyxis in other languages is perhaps instructive in this regard. Kitto and de Lacy (1999, pp. 5–6) and Uffmann (2006, pp. 1091–1093 & 1108) both report multiple

languages using *multiple* anaptyxis strategies, in preference orders. In this regard Uffmann (2006, pp. 1091–1093 & 1108) reports that Samoan (a Nuclear Polynesian language, as is te reo Māori), Shona (an Atlantic-Congo language), and Sranan Tongo (an English-based creole) all use all three strategies of default anaptyxis, copy anaptyxis, and feature spreading, although in different orders. Kitto and de Lacy (1999, pp. 5–6) report that Māori Kūki ʻĀirani (Cook Islands Māori, an Eastern Polynesian language, as is te reo Māori) uses copy anaptyxis after /ɾ/ and two default vowels in other phonetic environments, namely /a/ after /iɾ/ and /i/ everywhere else. Given these reported multiple strategies in use in Māori Kūki ʻĀirani, Samoan, Shona, and Sranan Tongo, it seems possible that this could also be the case in te reo Māori.

The evidence that speakers of te reo Māori use copy anaptyxis, among the anaptyctic strategies available to them, is strongest, as Kitto and de Lacy (1999, p. 14) report for speakers of Māori Kūki ʻĀirani and Kang (2011) reports for Polynesian languages word initially (Kang, 2011, p. 14) and for te reo Māori word finally (Kang, 2011, p. 15). It is possible that speakers of te reo Māori also use additional strategies as well, which may include default anaptyxis. If default anaptyxis is used, the default anaptyctic vowel is likely to be /a/.

5 Conclusions

This paper has explored how source word consonant clusters are made to conform to the required (C)V(V) syllable structure of te reo Māori. It has addressed how consonant clusters of source words are modified; which consonants are deleted when consonant clusters are reduced; how the specific properties of anaptyctic vowels are determined when they are inserted; and what the likely phonetic and/or other motivations are for these choices.

The principal conclusions of this paper are that:

(i) In the creation of te reo Māori Lehnwörter, the consonant clusters with the more sonorous phonemes (that is, semivowels and liquids) are likely to maintain the principle of phonetic similarity and to retain both phonemes, whereas the clusters with the less sonorous phonemes (occlusives and obstruents) are likely to be resolved through cluster reduction or cluster deletion.
(ii) With the exception of the semivowel clusters, in the creation of te reo Māori Lehnwörter, where the consonant clusters with the more sonorous phonemes do exhibit cluster reduction, they are much more likely to keep the more sonorous phoneme (that is, the liquid or the nasal).
(iii) With the less sonorous sibilant phonemes, where these clusters exhibit cluster reduction, it is very much more likely that the more sonorous phoneme is the one retained when the source word cluster adheres to the sonority sequencing principle (SSP) than when it breaks it.
(iv) Among the strategies for determining the features of the anaptyctic vowels of te reo Māori Lehnwörter, the evidence that speakers of te reo use copy anaptyxis is the strongest. This operates regressively for onsets and progressively for codas.
(v) If speakers of te reo Māori also use a strategy of default anaptyxis for determining the features of the anaptyctic vowels of te reo Māori Lehnwörter, this default vowel is highly likely to be /a/.

References

Bauer, L., & Warren, P. (2004). New Zealand English: Phonology. In E. W. Schneider, K. Burridge, B. Kortmann, R. Mesthrie, & C. Upton (Eds.), *A Handbook of varieties of English* (Vol. 1: Phonology) (pp. 580–602). Mouton de Gruyter.

Bauer, W., Parker, W., & Evans, T. K.. (1993). *Maori*. Routledge.

Bennett, J. B. (1861). *Report to the Honorable E. W. Stafford, Colonial Secretary, &c., &c., &c.* Registrar-General's Office.

Carlisle, R. S. (2001). Syllable structure universals and second language acquisition. *International Journal of English Studies, 1*(1), 1–19. https://revistas.um.es/ijes/article/view/47581/45601

Degani, M. (2012). Language contact in New Zealand: A focus on English lexical borrowings in Māori. *Academic Journal of Modern Philology, 1*, 13–24.

François, A. (2010). Phonotactics and the prestopped velar lateral of Hiw: Resolving the ambiguity of a complex segment. *Phonology, 27*(3), 393–434. https://www.cambridge.org/core/journals/phonology/article/phonotactics-and-the-prestopped-velar-lateral-of-hiw-resolving-the-ambiguity-of-a-complex-segment/C16AD657A5F47490BC087ED97993212F

Göksel, A., & Kerslake, C. (2005). *Turkish: A comprehensive grammar*. Routledge.

Harlow, R. (2004). Borrowing and its alternatives in Māori. In J. Tent & P. Geraghty (Eds.), *Borrowing: A Pacific perspective* (pp. 145–169). Pacific Linguistics, Research School of Pacific and Asian Studies, The Australian National University. https://openresearch-repository.anu.edu.au/bitstream/1885/254149/1/PL-548.145.pdf

Harlow, R. (2007). *Māori: A linguistic introduction*. Cambridge University Press.

Hazen, K. (2006). IN/ING Variable. In K. Brown (Ed.), *Encyclopedia of language and linguistics* (2nd ed.) (pp. 581–584). Elsevier.

Hohepa, P. W. (1965). *A profile-generative grammar of Maori* [Doctoral dissertation, Indiana University]. ProQuest Dissertations & Theses. https://go.openathens.net/redirector/wgtn.ac.nz?url=https://www.proquest.com/dissertations-theses/profile-generative-grammar-maori/docview/302133597/se-2?accountid=14782

Hohepa, P. W. (1967). *A profile-generative grammar of Maori*. Indiana University Publications in Anthropology and Linguistics, Memoir 20. University of Indiana.

Hume, E. V. (1992). *Front vowels, coronal consonants and their interaction in nonlinear phonology* [Doctoral dissertation, Cornell University]. ProQuest Dissertations & Theses. https://go.openathens.net/redirector/wgtn.ac.nz?url=https://www.proquest.com/dissertations-theses/front-vowels-coronal-consonants-their-interaction/docview/303990431/se-2?accountid=14782

Ka'ai, T M., & Moorfield, J. C. (2009). Ngā kupu Arotau—eweri tāima: Loanwords in Māori 1842–1952: Some interesting features of loanwords collected from the Māori language newspapers, the Paipera Tapu and some early published writings. *Te Kaharoa: The e-journal on indigenous Pacific issues, 2*(1), 1–38. https://ndhadeliver.natlib.govt.nz/delivery/DeliveryManagerServlet?dps_pid=IE13653296 [48-138-1-PB.pdf]

Kang, Y. (2011). Loanword phonology. In M. van Oostendrop, C. J. Ewen, E. Hume, & K. Rice (Eds.), *The Blackwell companion to phonology: Phonological interface* (pp. 1–25). John Wiley & Sons, Ltd. https://doi.org/10.1002/9781444335262.wbctp0095

Kitto, C., & de Lacy, P. (1999). Correspondence and epenthetic quality. In C. Kitto & C. Smallwood (Eds.), *Proceedings of Austronesian Formal Linguistics Association VI* (in press). Holland Academic Graphics.

Koro Wheke [now Ngaiterangi Smallman]. (2020). *Im not sure if putiputi is actually a loan word rather than a archaic term given new meaning. Putik is* [Comment]. Facebook. https://www.facebook.com/Whakatauki/posts/kia-orai-remember-this-friend-that-i-had-back-in-nz-who-in-every-second-word-wou/569806076385195/

Lin, Y.-H. (2009). Loanword adaptation and phonological theory. *Proceedings of the 21st North American Conference on Chinese Linguistics (NACCL-21), 1*, 1–12.
Māori Dictionary Project. (n.d.). *Te Aka Māori Dictionary*. https://maoridictionary.co.nz/dictionary-info
Moorfield, J. C. (n.d.). Kā. In *Te Aka Māori dictionary*. Retrieved October 28, 2026, from https://maoridictionary.co.nz/search?idiom=&phrase=&proverb=&loan=&histLoanWords=&keywords=k%C4%81
Moorfield, J. C. (n.d.). Koata. In *Te Aka Māori dictionary.* Retrieved October 28, 2026, from https://maoridictionary.co.nz/search?idiom=&phrase=&proverb=&loan=&histLoanWords=&keywords=koata
Moorfield, J. C. (n.d.). Kopi. In *Te Aka Māori dictionary.* Retrieved October 28, 2026 from https://maoridictionary.co.nz/search?idiom=&phrase=&proverb=&loan=&histLoanWords=&keywords=kopi
Moorfield, J. C. (n.d.). Kuata. In *Te Aka Māori dictionary.* Retrieved October 28, 2026, from https://maoridictionary.co.nz/search?idiom=&phrase=&proverb=&loan=&histLoanWords=&keywords=kuata
Moorfield, J. C. (n.d.). Motokā. In *Te Aka Māori dictionary.* Retrieved October 28, 2026, from https://maoridictionary.co.nz/search?idiom=&phrase=&proverb=&loan=&histLoanWords=&keywords=motok%C4%81
Moorfield, J. C. (n.d.). Motukā. In *Te Aka Māori dictionary.* Retrieved October 28, 2026, from https://maoridictionary.co.nz/search?idiom=&phrase=&proverb=&loan=&histLoanWords=&keywords=motuk%C4%81
Moorfield, J. C. (n.d.). Niu Ia. In *Te Aka Māori dictionary.* Retrieved October 28, 2026, from https://maoridictionary.co.nz/search?idiom=&phrase=&proverb=&loan=&histLoanWords=&keywords=Niu+Ia
Moorfield, J. C. (n.d.). Niupepa. In *Te Aka Māori dictionary.* Retrieved October 28, 2026, from https://maoridictionary.co.nz/search?idiom=&phrase=&proverb=&loan=&histLoanWords=&keywords=niupepa
Moorfield, J. C. (n.d.). Nū Ia. In *Te Aka Māori dictionary.* Retrieved October 28, 2026, from https://maoridictionary.co.nz/search?idiom=&phrase=&proverb=&loan=&histLoanWords=&keywords=N%C5%AB+Ia
Moorfield, J. C. (n.d.). Nūpepa. In *Te Aka Māori dictionary.* Retrieved October 28, 2026, from https://maoridictionary.co.nz/search?idiom=&phrase=&proverb=&loan=&histLoanWords=&keywords=n%C5%ABpepa
Moorfield, J. C. (n.d.). Patu. In *Te Aka Māori dictionary.* Retrieved October 28, 2026, from https://maoridictionary.co.nz/search?idiom=&phrase=&proverb=&loan=&histLoanWords=&keywords=patu
Moorfield, J. C. (n.d.). Taipō. In *Te Aka Māori dictionary.* Retrieved October 28, 2026, from https://maoridictionary.co.nz/search?idiom=&phrase=&proverb=&loan=&histLoanWords=&keywords=taip%C5%8D
Moorfield, J. C. (n.d.). Waka. In *Te Aka Māori dictionary.* Retrieved October 28, 2026, from https://maoridictionary.co.nz/search?idiom=&phrase=&proverb=&loan=&histLoanWords=&keywords=waka
Moorfield, J. C. (n.d.). Wuruhi. In *Te Aka Māori dictionary.* Retrieved October 28, 2026, from

https://maoridictionary.co.nz/search?idiom=&phrase=&proverb=&loan=&histLoanWords=&keywords=wuruhi

Moorfield, J. C. (n.d.). Wuruwhi. In *Te Aka Māori dictionary.* Retrieved October 28, 2026, from https://maoridictionary.co.nz/search?idiom=&phrase=&proverb=&loan=&histLoanWords=&keywords=wuruwhi

McNab, R. (1913). *The old whaling days: A history of Southern New Zealand from 1830 to 1840*. Whitcombe and Tombs.

Merton, R. K. (1968). The Matthew Effect in science: The reward and communication systems of science are considered. *Science, 159*(3810), 56–63. https://garfield.library.upenn.edu/merton/matthewii.pdf

Rickard, L. S. (1965). *The whaling trade in old New Zealand*. Minerva.

Ryan, J. S. (1972). The form and range of borrowings from English in Modern Maori. *Orbis: Bulletin international de Documentation linguistique, 21*(1), 136–166.

Schütz, A. J. (1985). Accent and accent units in Māori: The evidence from English borrowings. *The Journal of the Polynesian Society, 94*(1), 5–26. https://www.jstor.org/stable/20705907

Schütz, A. J. (2004). English loan words in Fijian. In J. Tent & P. Geraghty (Eds.), *Borrowing: a Pacific perspective* (pp. 253–294). Pacific Linguistics, Research School of Pacific and Asian Studies, The Australian National University. https://doi.org/10.15144/PL-548.253

Te Taura Whiri i te Reo Māori. (n.d.). *Our mahi: What we do and how we do it*. https://en.tetaurawhiri.govt.nz/mahi

Uffmann, C. (2006). Epenthetic vowel quality in loanwords: Empirical and formal issues. *Lingua, 116*(7), 1079–1111. https://www.sciencedirect.com/science/article/pii/S0024384105001075

Warren, P.l, & Bauer, L. (2004). Maori English: Phonology. In E.W. Schneider, K. Burridge, B. Kortmann, R. Mesthrie, & C. Upton (Eds.), *A handbook of varieties of English* (Vol. 1: Phonology) (pp. 614–624). Mouton de Gruyter.

Wilkinson, R. J. (1901). ڤوتيق puteḳ or putiḳ. In *A Malay-English Dictionary* (Part I, p. 477). Kelly & Walsh.

Williams, H. W. (with Kelly, G. P., & McEwen, J. M.). (1971). *A dictionary of the Maori language* (7th ed.). A. R. Shearer, Government Printer. (First edition edited by William Williams published 1844, fifth edition edited by Herbert William Williams published 1917)

Zimmer, K., & Orgun, O. (1999). Turkish. In International Phonetic Association (Ed.), *Handbook of the International Phonetic Association: A guide to the use of the International Phonetic Alphabet* (pp. 154–158). Cambridge University Press.