TE REO

Journal of the Linguistic Society of New Zealand

Volume 55   *2012*

Contents

THE INFLECTION-DERIVATION DIVIDE IN MĀORI AND ITS IMPLICATIONS[1]

Laurie Bauer: *School of Linguistics and Applied Language Studies, Victoria University of Wellington, PO Box 600, Wellington 6140 <laurie.bauer@vuw.ac.nz>*

Winifred Bauer: *Te Kawa a Māui, Victoria University of Wellington , PO Box 600, Wellington 6140 <winifred.bauer@vuw.ac.nz>*

Abstract

This paper considers the application of the categories of inflection and derivation to the morphology of New Zealand Māori. It is shown that the categories as they are usually defined do not fit well with the facts about Māori. Nevertheless, a case can be made for an inflection-derivation divide in Māori. However, if this division is made as proposed in the paper, it has wider typological implications which need to be considered. In particular it re-opens the debate on defining heads by their position in the word.

1. Introduction

The distinction between inflection and derivation was first developed for the Indo-European languages, and then subsequently applied to languages of other families. In many instances, the types of morphology we find in these other languages seem to fit more or less easily into the framework for the distinction which is constructed on the basis of Indo-European, but there is no necessary reason why this should be true in all instances. The Māori language has a relatively limited range of affixal morphological types, and thus looks like an interesting case to consider from a typological point of view. This problem has

not, to our knowledge, previously been specifically addressed for Māori, and so the investigation may also be taken to add something to our understanding of the Māori language. Of these two aims, however, it is the former which is the main focus of this paper.

2. Some background

2.1 Māori

The Māori language is the language of the indigenous people of New Zealand. It is an Eastern Polynesian language within the Austronesian language family. It is closely related to Tahitian, and more distantly to Hawai'ian and Rapanui of Easter Island.

Structurally, Māori is superficially a VSO language, with modifying words following the head in Noun Phrases (and also with left-headed compounds such as *roro-hiko* 'brain-electricity = computer').

Māori is a threatened language. Good figures are hard to come by, but at a maximum, 22% of the ethnically Māori population of New Zealand (approximately 15% of the 4.1m inhabitants) show any real fluency in the language (Harlow 2007: 195). Even such figures seem rather optimistic. Because of the efforts at language revitalisation since the 1970s, a large proportion of the people who speak Māori are L2 learners. It would be expected that the manipulation of morphological structure – particularly derivational morphology, if the term is relevant – among this group would be relatively constrained or uncertain since morphological productivity is rarely taught in L2 classes and cannot be deduced in any straightforward manner from the frequency of existing forms. There is some evidence of an increasing use of some types of Māori words among the non-Māori population (Macalister 1999), but this is simply loan vocabulary, and does not carry with it any necessary knowledge of morphological structure.

2.2 Inflection and derivation

Although the distinction between inflection and derivation is often introduced as a clear-cut distinction between incompatible categories, many linguists have suggested that there is a cline between the two, with the distinction between the two categories being canonical rather than determined by necessary and sufficient conditions (Scalise 1988, Dressler 1989, Plank 1994, Haspelmath 2002, Bauer 2003). Plank (1994) presents some 25 criteria which may be used

to distinguish between inflection and derivation. If the division really was a straightforward one, it would be redundant to have so many criteria; the fact that so many criteria may be needed indicates that the question is not a simple one.

Of course, it is always possible to make the distinction simple by accepting a single criterion as providing the crucial evidence in a given context. This is probably what is often done, albeit covertly. The only discussion of Māori morphology of which we are aware which specifically makes this distinction (Harlow 2007: 114–21) simply states that some of the morphology is inflect–ional and other processes are derivational without motivating that division; the division that Harlow proposes could be determined by a single criterion, that of the semantics of the categories, but we cannot be sure how Harlow actually decides on the allocation of processes. A more interesting question, though, is whether there is evidence from the criteria available to allow a relatively clear decision to be taken. Plank (1994), for example, shows that the criteria align rather well in motivating a distinction between inflection and derivation in English, a language with a notoriously impoverished inflectional system. To this end, we shall consider the Māori morphological system, and attempt to apply the criteria from Plank and other sources to the Māori data to see what if any distinction can be drawn in Māori. The criteria that are used in this paper are standard ones referred to in the works cited above, and their application can be discovered from these sources. It is expected that not all of the criteria available in the literature will necessarily apply in any individual case, and even that there may be criteria which go against the general run of the evidence. This problem will be tackled in section 3.3.

One of the points which should be borne in mind in judging the outcome of this exercise is that individual languages do not necessarily have both inflectional and derivational morphology. Greenberg (1966: 93) reports that "if a language has inflection, it always has derivation", with the implication that there are languages with only one of these types of morphology (or, indeed, with compounding, but neither inflection nor derivation). A perfectly reasonable conclusion might therefore be that there is no distinction to be made in Māori between inflection and derivation. So the fundamental question is not merely what is inflectional and what is derivational in Māori, but whether there is a distinction, and if so how that distinction is to be drawn.

Another point to consider is that the distinction between inflection and derivation has been elaborated in recent years by the introduction of a distinction between contextual and inherent inflection (Booij 1996).

Contextual inflection is the kind of inflection that marks agreement between categories, while inherent inflection is independent of other parts of the sentence (and often contrastively meaningful, as for instance, the difference between past and non-past tense in English) but still meets other criteria for inflectional morphology.

The main criteria to which appeal will be made in this paper, from the 25 or so offered by Plank and others, are set out below. The ordering of the presentation does not indicate importance.

a. Derivational affixes tend to be phonologically more word-like than inflectional affixes. This point is specifically made by Plank (1994). Applied to Māori, this distinction would be realised by a difference in moraic structure. No lexical word of Māori can be less than two moras long (Bauer 1993: 536), and monomoraic forms thus look grammatical and, in Plank's terms, more inflectional.

b. Inflectional affixes tend to be semantically and formally regular; derivational affixes may not be. Where there is lexicalisation of an affix, so that either form or meaning has become unpredictable in some complex forms, this may thus indicate derivational status.

c. Inflectional affixes tend to be fully productive both in the sense that they apply to all bases in a class and also in the sense that they are automatically available to any new base. Derivational affixes are typically less productive, showing more unmotivated gaps in the paradigm.

d. Complex words containing derivational affixes, but not complex words containing inflectional affixes, can normally be replaced in context by a morphologically simpler word which is not overtly marked for the category under consideration. In the clearest cases, derivatives but not inflected forms can be replaced in the sentence by monomorphemic words. Compare (1) and (2) below, which use English examples, with the relevant morphology in (1) being derivational and in (2) being inflectional.

(1)   Any poacher can find a pheasant.
      Any boy   can find a pheasant.

(2)   She prefers to paint her house herself.
      She *want   to paint her house herself.

This criterion is, in effect, a way of operationalising the notion that inflection is obligatory, derivation is optional.

e. Some categories are more frequently inflectional or derivational than others, because they reflect semantic categories which have the potential to be very widespread across lexemes (inflection) or categories which are not likely to be applicable to as many bases (derivation). Thus tense is often assumed to be likely to be inflectional (it can apply to virtually any verb), while instrumental nouns are less likely to apply to large numbers of bases and are likely to be derivational. Linked to this, a category which recurs in very few languages is likely to be derivational rather than inflectional.

f. Inflection is syntactic, derivation is lexical. While this criterion is virtually a slogan, its interpretation is not necessarily easy. Anderson (1982: 587) says that 'inflectional morphology is what is relevant to the syntax', but leaves open the question of how to recognise something that is 'relevant to the syntax'. Clearly, contextual morphology is relevant to the syntax, and if this is what Anderson means it is relatively helpful, but restrictive. In a wider sense, the difference between a nominal and a verbal clause might be considered syntactic, and thus a nominalisation marker might be considered as inflectional, or passivisation might be relevant because the redistribution of arguments which accompanies passivisation is syntactic (see Bauer 2003: 104–5). Comments on the application of this criterion thus have to be made suitably tentatively. In many cases, change of argument-structure (e.g. the difference between transitive and intransitive verbs) is seen as typically falling within the domain of derivational morphology, so there is some tension here in interpreting the evidence. Contextual inflection is more clearly inflectional under this criterion than is inherent inflection; morphology which appears on grammatical words (such as articles) as opposed to lexical words (such as nouns) is also likely to be inflectional under this criterion.

g. Other things being equal, derivational morphology occurs closer to the root than inflectional morphology, and inflectional morphology is peripheral in the word-form. This does not prevent sequences of derivational or inflectional markers. However, relative ordering may

be an important guide, since, for instance, anything that occurs closer to a root than a derivational affix is unlikely to be inflectional.

One set of criteria, usually considered very important in drawing a distinction between inflection and derivation, does not apply to Māori. In most languages, it is taken as important evidence if a particular morphological process causes a change in word-class. Thus the fact that *manage* is a verb but *management* is a noun is taken as evidence that the affixation of *ment* in English is derivational (but see Haspelmath 1996, 2002 for a dissenting view). Such arguments are not available in Māori because Māori does not clearly have word-classes (Bauer 1997: 65). Most, perhaps all, morphologically simple lexemes in Māori are multifunctional: the same form may head a verb phrase or a noun phrase. Words which look like adjectives to western Europeans pattern as verbs in Māori. Even words which can take plural marking (and thus are about as nominal as you get in Māori) can be found with nominalisation marking or passive marking. This means that criteria that depend upon a unique class being identifiable as the base or on causing a change in word-class are ruled out by the nature of Māori. This conclusion might appear at odds with the fact that we find morphological categories in Māori called 'nominalisation' and 'passive'. We deal with these in more detail below.

To see the multifunctionality of forms in Māori, consider the examples in (3) which illustrate different syntactic usages of the form *mōhio* usually glossed as 'know'.[2]

(3)   a.   Ka mōhio ia   ki taku ingoa (verb)
            TNS know 3SG DO my name
            '(S)he knows my name.'

      b.   Ki tōku mōhio hei   te   marae   te   hui (noun)
            To my   know at(FUT) DET marae   DET meeting
            'As far as I know, the meeting is at the marae.'

      c.   He ngākau mōhio ia (noun modifier)
            A   heart   know 3SG
            '(S)he shows great understanding.'

      d.   Hongi   mōhio ana rāua (verb modifier)
            press-nose know TNS 3DU
            'They hongied knowing [who they greeted].'

Even a word like *rākau* 'tree', which is typically used nominally, may be found used verbally, as illustrated in (4).

(4)   Ka whakatoki-a tōna purapura ki te   oneone, ka   tupu-a
        TNS bleach-PASS his   seed   to DET beach   TNS grow-PASS

        kia roa   anō, ka rākau-tia
        TNS long again TNS tree-PASS
        'His seed was bleached on the beach, and grew, and after a period
        of time it became a tree.'

We even find examples like that in (5), where a place name is used verbally.

(5)   I Puketapu ai   te ingoa, he tū-ranga   nō
        TNS name   POST-VERBAL_PARTICLE DET name a stand-NOM belong

        taua puhi raka
        that head there
        'The hill was called Puketapu [tapu-hill] because the head stood
         there.'

Such examples illustrate the impossibility of assigning forms to classes in Māori.

2.3 Nominalisation and passive

If there are no word-classes in Māori, there might seem to be some contradiction in having a suffix called 'nominalisation' (which seems to imply a word-class of noun) and 'passive' (which seems to imply a word-class of verb).

In a language like English, the term 'nominalisation' is read as making a noun from a form belonging to some other word-class. When we apply it to Māori, it has to be read as marking a form which is preferentially used in noun phrases, without specifying the word-class of the base.

The first point to be made here is that the syntax of Māori does construct nominal phrases differently from verbal phrases: they take different particles, for instance. In standard European-based grammar these things are called noun phrases and verb phrases, thus confusing the nomenclature for the word-class and the function of the phrase. In many languages, including English,

this makes sense. A noun form like *animal* is found in the head of a nominal phrase and not in the head of a verbal phrase, while a verb form like *deny* is found in the head of a verbal phrase but not in a nominal one. Even in English, there are many forms which are ambiguous: *man* and *mother* can be found as the heads of noun phrases or the heads of verb phrases. It is not generally taken that the existence of these ambiguous forms invalidates the fundamental distinction between nouns and verbs in English.

Māori is just the opposite. The vast majority of forms, and particularly the morphologically simple ones, occur freely as the heads of nominal or verbal phrases. There are occasional forms which strongly prefer one or the other function, but these should not detract from the fundamental observation on the lack of word-classes. The fact that the forms termed nominalisations occur almost exclusively as heads of nominal phrases should not mean that Māori has word-classes any more than the existence of *man* and *mother* should prove that English does not have them.

It should also be noted that Māori nominalisations do not always occur as the heads of noun phrases, nor do passives necessarily occur in verb phrases. Waite (1989: 78) notes specifically that 'the passive verb can be inserted into the true NP' and gives examples such as (6).

(6)   He   uaua   te   mōhio-tia   o   ngā whakaaro o   ngā kaihaina
       TNS difficult DET know-PASS GEN DET thought   GEN DET signatory
       'It is difficult to know what the signatories had in mind.'

The example in (7) illustrates a word with a nominalisation marker being used in a verbal position.

(7)   I   te   wā anō e   paka-nga   nei rāua ko   Tiapani…
       at DET time also TNS quarrel-NOM here 2DU SPEC Japan
       'At the time when it and Japan were at war'

3. Māori morphology

3.1 Some irrelevant morphology

Māori is largely isolating in morphological structure. Tense/aspect, for instance, is marked by particles rather than by morphological means. Possession classes are marked on prepositions and determiners, and not on the content words.

Compounding is productive in Māori. As well as the left-headed word-word compounds mentioned above and illustrated in (8) (from Bauer 1993: 519–21), there are forms with an *ā* link illustrated in (9). These are also left-headed. Bauer (1993: 522) says there is no reason to equate this linking element with the possessive preposition *a* (sometimes lengthened prosodically to *ā*), but compound-like constructions based on possessive constructions are widespread cross-linguistically. This *ā* could thus be seen as a preposition, or as a specific linking element. Since we are not concerned with the nature of compounding in Māori here, it makes very little difference which solution is adopted. In a wider sense, if the *ā* is a preposition, this might be used as an argument for saying that the items in (9) are complex lexical items but not compounds (see Bauer 2001: 704–5 for such argumentation with reference to other languages), though they are usually referred to as compounds in the literature on Māori.

(8)	kopa-mārō	wallet-hard	'briefcase'
	ipu-para	container-waste	'rubbish tin'
	pānui-whakamārama	notice-explain	'pamphlet'
(9)	waiata-ā-ringa	song-LINK-hand	'action song'
	utu-ā-hāora	price-LINK-hour	'hourly pay rate'

Reduplication is also widespread in Māori, and structurally and semantically complex (see Bauer 1993: 525–8 and Harlow 2007: 127–9 and references there). We shall largely ignore reduplication in what follows, since it does not add significantly to the discussion. With a single possible exception, which we shall discuss below, reduplication in Māori behaves like derivational morphology, and nothing would be added to the discussion here by including it, while its complications would make the discussion much harder to follow.

In what follows we consider the remaining synchronic morphological processes of Māori, and their place on the inflection-derivation cline. Krupa (1966) lists some affixes which will not be treated here on the grounds that they are no longer synchronically analysable.

3.2 Relevant morphological processes in Māori

Those processes which are deemed relevant here are those which might be classified as inflectional or derivational. Most of them are affixal, the plural marker being the only exception. The processes listed here, along with

the compounding and reduplication mentioned in section 3.1, provide an exhaustive list of analysable morphological processes in Māori.

a.   Seven or eight nouns (all denoting people and family relations) mark the plural by vowel lengthening. Only one other noun has plural marking (Bauer 1993: 354). The nouns concerned are set out in (10) (see Bauer 1997: 160; Harlow 2007: 115).

(10) | SINGULAR | PLURAL | GLOSS |
|---|---|---|
| matua | mātua | 'parent' |
| tangata | tāngata | 'person' |
| teina | tēina | 'same sex younger sibling' |
| tipuna/tupuna | tīpuna/tūpuna | 'grandparent' |
| tuahine | tuāhine | 'sister of a man' |
| tuakana | tuākana | 'same sex elder sibling' |
| wahine | wāhine | 'woman, wife' |

and for some speakers:

whaea	whāea	'mother'

Note that a word like *tungāne* 'brother of a female' and other relationship words do not mark plurality in this way.

The vowel lengthening could be viewed as a matter of apophony or as a matter of reduplication. Again, the analysis does not seem to be particularly relevant for our purposes. We can illustrate that there are features which are canonically inflectional and canonically derivational, even for this set of forms (see Table 1).

Table 1: Number marking on some nouns

INFLECTION	DERIVATION
• Meaning	• Irregular (not all relationship nouns)
• Marker one mora long	• Unproductive
	• Plural not usually marked on nouns

Table 1 is to be interpreted as follows. The relevant number marking looks inflectional because number-marking is frequently inflectional across languages and because the marker is monomoraic. The number marking looks derivational for the reasons given in the second

column of Table 1: the pattern of marking does not cover a natural class of nouns, it is unproductive, and the category is not usually marked on nouns in Māori at all. It might be possible to add to this list from the criteria given by Plank, but these seem to be the criteria which apply clearly (albeit not consistently) to these forms.

b. Singularity is marked on determiners such as *tēnei* 'this', *ēnei* 'these' (Bauer 1993: 386). This example can be used to represent the various forms with an initial *t* in the singular (see Bauer 1997: 151–6 for some of the others). The major reasons for attributing this marking to either inflection or derivation are as given in Table 2, but these reasons are not particularly convincing.

Table 2: Marking on deictics

INFLECTION	DERIVATION
• Grammatical	• Not productive
	• Not used on all determiners

Historically, the initial *t-* in such forms derives from the form *te*, usually glossed as the singular definite article (see Bauer 1997: 144 for some discussion), and while the morphophonemics of the mergers is to some extent unpredictable, it seems clear that this is univerbation of a syntactic process. It is thus not particularly surprising that the inflection-derivation distinction does not fit neatly onto what was originally a syntactic sequence.

c. The prefix *taki-* added to numerals 1–9 means 'in groups of' (Bauer 1993: 498).

d. The prefix *hoko-* added to numbers 1–9 (possibly only 2–7) means '20 times', thus *hoko·whitu* (20 x 7 = 140) means a group of 140 or a large number, esp. an army (Bauer 1993: 499).

(11)   Ka   whao·na   te   wharau nei e   te   hoko·whitu rā
         TNS enter·PASS DET shed   this by DET 20 x 7   there
        'The party entered the construction' (Bauer 1993: 499)

Bauer notes that there may or may not have been precisely 140 people involved.

e. The prefix *tua*- is added to numbers 1–9 to mark ordinals (Bauer 1993: 497)

(12)   Ko te tohu tua·rua tēnei i   tae   mai   ki a   ia
       EQ the sign ORD·two this   TNS arrive hither to PERS 3SG
      'This was the second sign that came to him' (Bauer 1993: 247)

There is a syntactic periphrasis that can be used in place of this, and which must be used with numbers greater than nine.

f. The prefix *toko*- is added to numbers from 1–9 when counting people (Bauer 1993: 496–7), and also on the question words for numbers: *toko-hia* 'how many people?' and some quantifiers.

(13)   Toko·rima ōna   tuākana
      COUNT·five 3PL.GEN same-sex_sibling
      'He had five older brothers' (Bauer 1993: 496)

g. The prefix *pū*- is used on a few bases and probably not productively to mark attenuation: *pū·whero* 'reddish' (Bauer 1993: 511).

All of these markers in (c)–(g) look as if they are fairly well-behaved instances of derivation. They are not productive; their semantics seems to belong to categories which might be expected to be derivational; the meaning 'twenty times' is not common cross-linguistically; they are all two-mora markers, and thus like words in phonological structure; they can all be replaced by monomorphemic words in context; most of them are not required by the syntax. However, the prefix in (f) has some inflectional characteristics, being used on grammatical words, being obligatory in its narrowly-defined construction and showing agreement for person-hood. The presence of a syntactic periphrasis for the item in (e) might also be seen as suggesting inflectional characteristics.

h. The suffix -*(C)anga* (where C represents a variable consonant³) marks a nominalisation. The same affix applies to some modifiers of the nominalisation: *puta·nga ohorere·tanga* appear·NOM sudden·NOM = 'sudden appearance' (Bauer 1993: 512). There is a certain amount of lexicalisation of these nominalisations, so that *kā·inga* 'home' derives from *kā* 'to set on fire, burn', and *moe·nga* 'bed' derives from *moe* 'to sleep'. The contrasting nominalisations *poro·nga* 'end' and

poro·hanga 'fragment' both come from *poro* 'end, broken off" (Bauer 1997: 516).

Table 3 shows the features of this affix which might be considered inflectional or derivational.

Table 3: Nominalisation marking

INFLECTION	DERIVATION
• Used to mark agreement on modifiers • Productive – possibly fully • Observed outside passive: *kite·a·tanga* 'see·PASS NOM being seen', though this is rare • Possibly grammatical usage	• Formally unpredictable[4] • Some lexicalisation (semantic); some different allomorphs contrast • Mainly bimoraic structure • Can be replaced by monomorphemic form in some uses

i. The prefix *kai-* creates human agentives where the base would be transitive if it headed a verbal phrase, as illustrated in (14) (Bauer 1993: 514; Harlow 2007: 124).

(14) | BASE | GLOSS | COMPLEX FORM | GLOSS |
|---|---|---|---|
| ako | 'teach' | kaiako | 'teacher' |
| kōrero | 'speak, speech' | kaikōrero | 'speaker, story-teller' |
| mahi | 'work, do' | kaimahi | 'worker' |
| titiro | 'watch' | kaititiro | 'watchman' |
| waiata | 'sing, song' | kaiwaiata | 'singer' |
| whakahaere | 'administer' | kaiwhakahaere | 'administrator' |

This prefix seems fairly solidly derivational, as is shown in Table 4.

Table 4: Agentive marking

INFLECTION	DERIVATION
• Regular form and meaning	• Not productive on all base types: **kaihaere* • Meaning is of the type expected for derivation • Bimoraic structure • Can be replaced by monomorphemic words

j. The prefix *whaka-* is a causative marker as illustrated in (15) (Bauer
    1993: 515–6; Harlow 2007: 124–6).

(15)

BASE	GLOSS	COMPLEX FORM	GLOSS
atu	'away from speaker'	whakaatu	'point out'
haere	'go'	whakahaere	'administer'
kā	'burn'	whakakā	'turn on (a machine etc)'
kāhore	NEG	whakakāore	'deny'
kite	'see'	whakakite	'show'
roa	'long'	whakaroa	'lengthen'
rongo	'hear'	whakarongo	'listen'
tangata	'man'	whakatangata	'change into a man'

We see in (15) some unexpected classes of base (and *whaka-* may be
added to phrases, though this is not illustrated in (15)), some cases of
lexicalisation and some instances where argument structure has been
affected. The characteristics of inflection and derivation shown by
this prefix are listed in Table 5. In a few lexicalised forms there is an
alternative form *whā*, as in *whāngote* 'breastfeed' from *ngote* 'suck'
or *whāinu ~ whakainu* 'cause to drink'.

Table 5: Causative marking

INFLECTION	DERIVATION
• Form and meaning regular • High productivity • Possibly grammatical usage in that affects argument structure	• Occurs inside *kai-* marking • Can be replaced by monomorphemic forms • Bimoraic structure • Lexicalisation of *whaka-* forms

k. The suffix *-(C)ia* (where C represents a variable consonant) is used
    to mark the passive, which in verb phrases with canonical transitive
    verbs may be functionally an imperative (Bauer 1997: 447). Manner
    particles, the quantifier *katoa* 'all' and lexical modifiers take the
    -(C)ia suffix in agreement with a passive verb when they follow
    the verb: *kite·a rawa·tia* see-PASS INTENS-PASS 'finally seen' (Bauer

1997: 487). There are a few verbs which do not traditionally take passive endings, even in the imperative, such as *hōmai, hōatu* 'give' or *waiho* 'to leave (behind)'; the result is that sometimes it is possible to replace forms with the suffix *-(C)ia* with a form with no passive ending in context. The passive suffix may be added to names and other words which might be expected to be nouns, and even occasionally to nominal phrases (Harlow 2007: 120).

The morphology of the Māori passive has been well covered by linguists since Hale (1968). The particular consonant that is used in the suffix is largely lexically determined, though there is a default used on unfamiliar words, in agreement contexts, or if memory fails. What that default consonant is varies from dialect to dialect: /t/ in some, /ŋ/ (written <ng>) or /h/ in others (Harlow 2007: 116). The result is that the same verb may be heard with different passive suffixes, though usually the complex forms thus created are synonymous. Just occasionally, different passive suffixes may be semantically distinguished (see Williams 1971 sv *aroha* 'love'[5]). The extent to which the various forms of the passive can be treated as allomorphs of the same morpheme is thus marginally in doubt in the current state of the language.

Table 6: Passive marking

INFLECTION	DERIVATION
• Meanings • Used to mark agreement • Occurs outside causative marking • Does not usually co-occur with *-Canga* • Grammatical usage in that it affects argument structure (but the syntax is controversial)	• Unpredictable form, albeit with default • In some uses can be replaced by forms with no passive ending • Largely bimoraic structure • Some lexicalisation: *kawea* (< 'carry') = 'situated'

3.3 Summarising the findings

The discussions that have been provided in section 3.2, and which can be taken to cover all the potentially relevant morphology of Māori, can be summarised as in Table 7, where some of the relevant criteria are listed and 'I'

indicates that the morphological markers for that category appear inflectional by the relevant criterion and 'D' indicates that they appear derivational. Cases where there are problems have been discussed.

Table 7: Summary of findings

	PROD-UCTIVE	AGREE-MENT	REGULAR FORM	MEANING	SYNTAX	LEXICAL-ISED	MORAIC STRUC-TURE	REPLACE WITH SIMPLER
Pl	D	D	I	I	I	I	I	D
Numbers (& attenuation)	D	D	I	D	D	D	D	D
Noml	I	I	D	?	I	D	D	D
Agt	I	D	I	D	D	D	D	D
Caus	I	D	I	?	I	D	D	D
Pass	I	I	D	I	I	D	D	(D)

The rather irregular placement of 'I's and 'D's in Table 7 is, of itself, a rather unsatisfactory result for a theory which expects to see a fairly clear distinction between the two categories (as Plank 1994 finds for English, for instance). However, there are two things to be said about a table of this kind.

The first is that not all the criteria are of equal value: regular form in derivation is far from unusual and is thus a relatively weak argument for inflectional status; correspondingly, irregular inflection is well-known, but usually in specific word-forms rather than in the general form of the affix concerned; the problems with the criterion of syntactic relevance were discussed earlier; a certain amount of lexicalisation is not uncommon in inflection (consider forms like *brethren* in English which is not synonymous with *brothers*) and so some may easily be tolerated. Moreover, as we have seen, the application of some of the criteria is in doubt, and these criteria are presumably less relevant in their application to Māori than they might be in relation to other languages.

The second is that there appears to be an oversupply of 'D's in the table. This could indicate, in terms of the discussion above, that there is no clear distinction between inflection and derivation in Māori and that the distinction simply does not apply. Alternatively, it could be a signal that if there is a

difference between inflection and derivation in Māori the distinction is not precisely the same as the one that is found in English (or, a fortiori, in other languages, whether in the Indo-European family or not). In other words, if we have a distinction between inflection and derivation in Māori, it may be a distinction which is defined for Māori, and where the criteria do not neatly match the criteria that are found in other languages.

Acknowledging that possibility, we should like to suggest that there may well be a useful distinction to be drawn in Māori. That distinction would be based on deciding which of the criteria in Table 7 are the most important and prioritising those criteria. If we focus on the first two criteria in Table 7 (as illustrated in Table 8), then nominalisation and passive start to look as though they might be candidates for the inflectional category in Māori, while all the other morphology is derivational.

Table 8: Selecting criteria from the summary

	PROD-UCTIVE	AGREE-MENT	REGULAR FORM	MEANING	SYNTAX	LEXICAL-ISED	MORAIC STRUC-TURE	REPLACE WITH SIMPLER
Pl	D	D	I	I	I	I	I	D
Numbers (& attenuation)	D	D	I	D	D	D	D	D
Noml	**I**	**I**	D	?	I	D	D	D
Agt	I	D	I	D	D	D	D	D
Caus	I	D	I	?	I	D	D	D
Pass	**I**	**I**	D	I	I	D	D	(D)

In the discussion section the reasons for this choice will be made clearer, and the implications of such a choice for linguistics and in particular linguistic typology will be examined.

It should be noted, however, that if the conclusion sketched in Table 8 is adopted, it is a very selective conclusion: over twenty of Plank's criteria are being – if not ignored – downgraded in strength for this language. This may be a problem. Where the evidence is as far from clear-cut as it is here, it is obvious that alternative analyses of the data would be possible; all that would be required would be a different prioritising of the criteria. In particular, we have argued that we cannot use word-class to argue for derivation in Māori.

If nominalisation and passivisation were taken to be derivational on some reinterpretation of the word-class arguments, then a strong argument could be made for there not being any inflection in Māori at all. However, as we have indicated, we do not believe that such a strategy would be justified. Furthermore, if were used, it would leave some other factors to be explained, as we show below.

4. Discussion

The proposal put forward in the last section can be argued for on the basis of the standard characteristics of inflection and derivation, and perhaps particularly because the affixes selected as inflectional by this method are the only ones which are used to mark agreement (so are instances of contextual inflection, the type of morphology most clearly at the inflectional end of the cline). However, this proposal has a particular benefit for the discussion of Māori, in that if we adopt it, we find that inflection in Māori is always suffixal, while derivation is always prefixal. This is a very unusual distribution, and a very neat outcome. If any other conclusion about the inflection-derivation split in Māori were postulated, this division between prefixation and suffixation would remain an unexplained oddity. However, this conclusion has some implications which need to be considered.

First, the information in Table 9 shows that across languages we have a range of patterns in the position of the head in compounds, the position of derivational morphology and the position of inflectional morphology. The list of patterns in Table 9 is not exhaustive, but it is often difficult to determine from descriptive grammars precisely what should be considered inflection and what derivation in that particular language, and so it is difficult to decide how the categories are distributed. For instance, it looks from Schadeberg's (1984) description of Swahili morphology as though inflection and derivation may both occur on both sides of the root, but since no overt claims are made about this by Schadeberg, the interpretation may be faulty.

Some of the implications of the data presented in Table 9 will be discussed below. However it is worth saying that its implications are not necessarily clear: it is beyond the scope of this paper to go into the various cross-linguistic patterns to see if there are any correlates of this division and thus whether it might itself be a typological dimension.

Table 9: Distribution of heads in morphological categories

	COMPOSITION	DERIVATION	INFLECTION
Māori	Left	Left	Right
English	Right	Right/(left)	Right
French	Left	Right/(left)	Right
Warrwa (McGregor 1994)	none mentioned	Right	Left/Right
Ura (Crowley 1998)	Left	Left	Left/Right
Mandan (Mixco 1997)	Left/Right	Left	Left/Right

It is, in principle, clear that – independent of our findings in this paper – a particular feature can be inflectional in one language and derivational in another. This could be true even if all the same criteria applied in the languages being compared. Thus there is no contradiction in a conclusion that, for example, plural is inflectional in German (Fleischer 1975: 47) but derivational in Diyari (Austin 1981: 41). It is nevertheless important to state this overtly, since many authorities appear to believe that certain categories will automatically be inflectional or derivational categories. It is difficult to find overt statements to this effect, and likely that authors of less clearcut statements would argue if challenged. Thus Hall (1992: 58) comments that cross-linguistically "many… inflectional features such as tense, plurality and definiteness are prefixed in a great number of languages" but might argue that these are the examples given by the sources he cites, and that this is not a definition. It is in non-specialist dictionaries that such definitions can be found, as in Pearsall (2002) where inflection is defined as "a change in the form of a word… to express a grammatical function or attribute such as tense, mood, person, number, case, and gender".

It ought, equally, to be clear that, if the distinction between inflection and derivation can be defined by a large number of potentially intersecting criteria, what counts as 'inflectional' in one language may not necessarily count as 'inflectional' in another. The oversupply of 'D's in Table 7 mentioned above could be a signal that morphology in Māori is all closer to the derivational end of the continuum than might be expected on the basis of our experience with Indo-European languages.

From this it follows that saying, as we did just above, that a particular category is inflectional in one language but derivational in another (plural is inflectional in German but derivational in Diyari) is not necessarily

informative: we cannot tell whether it means that the language behaves differently in regard to a stated set of criteria, or whether it means that different criteria have to be used to distinguish between inflection and derivation in the two languages (or, indeed, a mixture of the two). This has implications for those engaged in language typology. Distinctions which are claimed to be distinctions in terms of the inflection-derivation dichotomy may not be at all relevant. The case of Māori makes this point particularly clear.

The major point about inflection and derivation that emerges from Table 9 and from the research reported here concerns headedness, and, specifically, the notion that headedness might be determined by the relative position of the head and the non-head, something that was postulated in the Righthand Head Rule. When the Righthand Head Rule was introduced to morphology (Williams 1981), it was argued that compounds and derivatives were always right headed. In the years following that original statement of the rule, it was pointed out many times that the original statement was far too strong: there are many languages, including French and Vietnamese (and, as we have seen, Māori), which have left-headed compounds, for instance. The discussion of the headedness of derivatives has raised rather less controversy. In most cases, derivational affixes are accepted as heads, though most clearly when they determine word-class, something which is not relevant to Māori. The discussion of inflection, however, has not led to any clear outcome. Lieber (1992) argues that inflections are not heads, while Di Sciullo & Williams (1987) argue that they are, but 'relativised' heads – that is, they are heads only in so far as the grammatical category they realise is involved.

What the proposal put forward about Māori here suggests is that if headedness is related to handedness, and derivational endings are typically heads, then inflectional affixes cannot be heads. We have not previously seen evidence which makes this point so clearly, and it has the effect of re-opening the discussion of inflectional morphology and headedness.

Notes

1   We should like to thank Ingo Plag, Liza Tarasova, Natalia Beliaeva and attendees at the IMM15 Conference in Vienna, February 2012, for their discussion and feedback, as well as the anonymous referees for *Te Reo*.

2   Abbreviations used in glosses are: 2 'second person', 3 'third person', DET 'determiner', DO 'direct object', DU 'dual', EQ 'equative', FUT 'future', GEN 'genitive', INTENS 'intensifier', NOM 'nominalisation', ORD 'ordinal',

PASS 'passive', PERS 'person marker', PL 'plural', SG 'singular', SPEC 'specifying', TNS 'tense, aspect'.

3   There are some variants which have a slightly different form: *-inga* or *-nga*.

4   The consonant was originally a base-final consonant in an earlier stage of Austronesian, and was not predictable because it was simply part of the base. In modern Māori there is some variation in the consonant in individual instances and the consonant is not always the same as the one in the passive ending (see below) which has the same origin.

5   Such cases are rare, and, as in this case, controversial; but part of the motivation for the current slogan of *Arohatia te reo* ('Cherish the language') was that the form *arohaina* (the traditional passive form) would be semantically inappropriate.

References

Anderson, Stephen R. 1982. 'Where's morphology?' *Linguistic Inquiry* 13: 571–612.

Austin, Peter. 1981. *A grammar of Diyari, South Australia*. Cambridge: Cambridge University Press.

Bauer, Laurie. 2001. 'Compounding.' In Martin Haspelmath, Ekkehard König, Wulf Oesterreicher and Wolfgang Raible (eds.), *Language Universals and Language Typology*. Berlin and New York: de Gruyter. 695–707.

Bauer, Laurie. 2003. *Introducing linguistic morphology*. 2nd edition. Edinburgh: Edinburgh University Press.

Bauer, Winifred. 1993. *Maori*. London and New York: Routledge.

Bauer, Winifred. 1997. *The Reed reference grammar of Māori*. Auckland: Reed.

Booij, Geert. 1996. 'Inherent versus contextual inflection and the split morphology hypothesis.' *Yearbook of Morphology 1995*: 1–16.

Crowley, Terry. 1998. *Ura*. München and Newcastle: Lincom Europa.

Di Sciullo, Anna-Maria and Edwin Williams. 1987. *On the definition of word*. Cambridge, MA: MIT Press.

Dressler, Wolfgang U. 1989. 'Prototypical differences between inflection and derivation.' *Zeitschrift für Phonetik, Sprachwissenschaft und Kommunikationsforschung* 42: 3–10.

Fleischer, Wolfgang. 1975. *Wortbildung der deutschen Gegenwartssprache*. 4. Auflage. Tübingen: Niemeyer.

Greenberg, Joseph H. 1966. 'Some universals of grammar with particular reference to the order of meaningful elements.' In Joseph H. Greenberg (ed.), *Universals of language*. 2nd edition. Cambridge, MA and London: MIT Press. 73–113.

Hale, Kenneth. 1968. Review of P.W. Hohepa (1967). *Journal of the Polynesian Society* 77: 83–99.

Hall, Christopher J. 1992. *Morphology and mind*. London and New York: Routledge.

Harlow, Ray. 2007. *Māori*. Cambridge: Cambridge University Press.

Haspelmath, Martin. 1996. 'Word-class-changing inflection and morphological theory.' *Yearbook of Morphology 1995*: 43–66.

Haspelmath, Martin. 2002. *Understanding morphology*. London: Arnold.

Krupa, Victor. 1966. *Morpheme and word in Maori*. The Hague: Mouton.

Lieber, Rochelle. 1992. *Deconstructing morphology*. Chicago and London: University of Chicago Press.

Macalister, John. 1999. 'Trends in New Zealand English.' *New Zealand English Journal* 13: 38–49.

McGregor, William. 1994. *Warrwa*. München and Newcastle: Lincom Europa.

Mixco, Mauricio. 1997. *Mandan*. München and Newcastle: Lincom Europa.

Pearsall, Judy. (ed.) 2002. *Concise Oxford English dictionary*. Oxford: Oxford University Press.

Plank, Frans. 1994. 'Inflection and derivation.' In R.E. Asher (ed.), *The encyclopedia of language and linguistics*. Oxford: Pergamon. 1671–8.

Scalise, Sergio. 1988. 'Inflection and derivation.' *Linguistics* 26: 561–581.

Schadeberg, Thilo C. 1984. *A sketch of Swahili morphology*. Dordrecht: Foris.

Waite, Jeffrey. 1989. 'Tough- and pretty-movement in Maori'. *Te Reo* 32: 61–94.

Williams, Edwin. 1981. 'On the notions "lexically related" and "head of a word".' *Linguistic Inquiry* 12: 245–274.

Williams, Herbert W. 1971. *A dictionary of the Maori language*. 7th edition. Wellington: Government Printer.

THE RELATIONSHIP BETWEEN LANGUAGE IDEOLOGY AND LANGUAGE CHANGE IN A SMALL, ISOLATED COMMUNITY: THE CASE OF PALMERSTON ISLAND

Rachel Hendery: *School of Culture, History and Languages, The Australian National University, Canberra 0200, Australia <rachel.hendery@anu.edu.au>*

Abstract

Palmerston Island is a tiny isolated community in the Pacific. Over the past 140 years it has developed a unique linguistic and cultural identity, influenced by England, the Cook Islands, and more recently New Zealand. The islanders strongly identify with England and consider themselves very different from the rest of the Cook Islands, to which Palmerston Island officially belongs. This paper explores the relationship between Palmerston Islanders' conceptions of themselves and their linguistic ideologies. It is shown that the construction of linguistic and social norms is not entirely subconscious: the community is aware of the different origins of lexical items, and the cultural and social affiliations signalled by different linguistic choices. Subconscious co-evolution of culture and language also takes place and appears likely to be responsible for the substrate influences of Cook Island Māori in both realms.

1. Introduction

Linguistic ideologies, defined by Rumsey (1990: 346) as 'shared bodies of commonsense notions about the nature of language in the world', are a key

element in a community's and individuals' self-identity. Speakers' beliefs about language contribute to their linguistic choices and over time, the cumulation of these choices shapes the language itself. This is likely to be particularly true in a linguistic environment that is shaped by language contact, with a large amount of linguistic variation. As noted by Woolard (1998: 9), 'Ideology is variously discovered in linguistic practice itself; in explicit talk about language, that is, metalinguistic or metapragmatic discourse, and in the regimentation of language use through more implicit metapragmatics.' In this paper I will mainly focus on the 'explicit talk about language' I observed in a case-study of a small isolated community with mixed origins: Palmerston Island. There will also be some discussion of the 'implicit metapragmatics' of language use in the community. Most importantly, I will situate this description of Palmerston Island linguistic ideology in the wider context of the islanders' cultural and linguistic identities and affiliations in order to gain a new perspective on development of this unique English variety.

After 140 years of near-total isolation, the inhabitants of Palmerston Island, a tiny atoll in the Cook Islands group, have developed an identity that draws on both English and Polynesian linguistic and cultural traditions. They consider themselves English ethnically, culturally, and linguistically, but also have strong ties to the rest of the Cook Islands, and to New Zealand. As a close-knit community that has historically had between 50 and 100 inhabitants, Palmerston Island has also developed and perpetuated its own unique traditions, linguistic features, and ways of thinking about the world.

The small size and isolation of Palmerston Island makes it an ideal location for examining linguistic ideology and identity, because one is able to explore this at the level of individual community members. A language ideology is not a monolithic view held by speakers of the language. Blommaert (2005: 173–174) points out a 'layered simultaneity' that comes from ideologies operating differently at different levels, for example intrapersonally, in the workplace, in society at large, and over different time periods. I think it is worth pointing out, however, that as well as this polycentricity and stratification in the *system(s)*, there is also a multiplicity of ideologies among speakers. Each individual in a speech community holds his or her own views on language, culture, identity and social organisation (and may, of course, hold different views with regard to the different 'layers' referred to by Blommaert). The diversity of these views, the ways they relate to the ideas and beliefs of others in the community, and the ways in which they are transmitted to children through their families and teachers, are all factors that contribute to the community's construction

of their language and identity (cf. Cameron 1990: 88 for the importance of examining linguistic norms and identity at the level of the individual actor).

The dialect of English spoken on the island is described in detail from a formal linguistic perspective in Ehrhart-Kneher (1996), and Hendery and Ehrhart (2011, in press). The remarkable degree of inter- and intra-speaker linguistic variation on the island is discussed in the latter paper. The link between Palmerston Island's cultural and linguistic traditions provides the key to the existence of such variation, which is found in the islanders' own sense of linguistic identity, as expressed in their views on their speech variety, their history, linguistic norms, prescriptivism, education, and the island's historical and current ties to the region and beyond. In this paper I will describe these views, as expressed by Palmerston Islanders in interviews I conducted in 2009. I will show how these views relate to the island's current linguistic and cultural context, and how they hold clues to the island's history. It will be seen that close attention to community knowledge of and opinions on language, especially at the level of the individual, can result in a more thorough, nuanced understanding of the linguistic and cultural situation than would a heavily abstracted description.

The cultural and social elements that combined in the original settlement of Palmerston Island are not identical to those that are valued and discussed by Palmerston Islanders today. In this paper I will examine how the current linguistic features of the dialect reflect the choices Palmerston Islanders have made historically with regard to the narratives they privilege, the cultures with which they identify, and the ancestors they remember. I will show that these effects are indirect, mediated by the linguistic ideologies that members of the community hold. By examining the roles of linguistic and cultural identity, and language ideology, we can gain a clearer idea of how a newly-formed community develops a shared language, and we can begin to explain the functions of on-going variation in that language.

2. Background

Palmerston Island is approximately 600m across, with a population in August 2009 of 54. It belongs to the Cook Island group and lies about halfway between Rarotonga and Niue, but slightly north of both. It was settled in the mid-1800s by William Marsters, probably originally from Leicestershire or Birmingham, his two (later three) Cook Island wives and a small group of others, including

several more Cook Islanders, a Portuguese man, his wife and child. The inhabitants today are the descendants of this small group and speak a dialect of English; those who grew up on the island are monolingual. In July—August 2009 there were 13 adult women, 13 adult men, and 28 children. Five of the people currently on the island are not originally from Palmerston: these include the teacher (from New Zealand, now married to a local man), two Rarotongan women married to local men, and the two missionaries (from another island in the Cook Islands group). According to Crocombe and Marsters (1987: 222), outside marriage has been rare historically.

The island has been very isolated—there is no regular transport to or from it, and it is 400 km away from the closest other inhabited islands. There is no television signal. In 2009 there was one telephone and one computer with a very slow satellite internet connection. Up until around eight years ago, there were no moorings, so the visitors were few, as it was dangerous to try to anchor if the wind was not exactly right. During the 20th century, there were sometimes many years without any contact with the outside world. Nowadays around 30 yachts visit each year during the August–September 'cruising' season, staying for a few days each. Detailed records of these visitors are kept. Various volunteers and teachers have helped with the school, and records of these are kept at the school. While the church is currently staffed by Cook Island missionaries, previous pastors have all been local Palmerston Islanders.

Older historical influences on the island are also relatively easy to track. Although older records that had been kept in the church were lost in a fire some years ago, the oldest inhabitants were able to recall the few visitors to the island in their youth. These were mainly occasional school board superintendents from England and sometimes Rarotonga, who came once every few years to examine the children on their lessons. Several outsiders spent long periods on the island: Victor Clark was an Englishman shipwrecked on Palmerston Island for a year in the 1950s (Clark 1960); Arthur Helm stayed there for a year in the 1960s (Helm & Percival 1973), and more recently Daniel Curran, a sociology PhD student from the University of Kent, visited for nine months during 2010.

The isolation and small size of the island are advantages for a linguist as they make it possible to (a) interview all Palmerston Islanders, and (b) track all external influences on the language, making Palmerston Island a wonderful laboratory for studying the development of linguistic and cultural identities in small mixed-origin communities.

During four weeks of fieldwork in 2009, I spoke with all of the then-

inhabitants of the island except for one elderly man, who is blind and deaf. While there are certainly common themes in the local views on the island's history, identity and language, there is also much diversity. Each family on the island has its own traditions and stories. The quotations and views recorded in this paper should therefore not be taken to be general consensus, or even representative of more than one family's perspective, unless otherwise stated. As views are so diverse, those reported here are ascribed to individuals, but for privacy reasons, these individuals are not referred to by name, but only with two-letter codes which are not their actual initials.

Palmerston Islanders divide themselves into three groups, named after the three wives of Marsters: *Akakaingara*, *Matavia*, *and Tepou*. The island itself is also divided in three, as are the other islets around the lagoon, each third belonging to one of the families.[1] Membership in one of the three families determines land inheritance, council representation, hunting and gathering rights, marriage possibilities, and to some extent, day-to-day socialisation patterns.

Cutting across the distinction between the three families is another distinction, between the *beachfellas* and the *bush people*, as they are called by Palmerston Islanders. The former live in the cleared sandy main settlement area at the north end of the island. The latter have built houses further south, among the palm trees that cover the rest of the island. In practical terms, the distance between the most distant houses is only a couple of hundred metres – a few minutes' walk. In social terms, however, distinctions are made between the two groups, and each believes the other to have different cultural and linguistic practices.

A further distinction is made between locals and outsiders. The most transient of the outsiders are the 'yachties', but even the women who have married into the community are still seen as *comers* [kʊməs]. The islanders use short wave radios to communicate with each other (much as other communities use mobile telephones), and call signs are assigned only to those who were born on the island. The 'outsider' women, who are married to locals and have been living in the community for 10 or more years, do not have their own call signs but go by their husbands'.

Palmerston Islanders view themselves as distinct from the rest of the Cook Islands, and are proud of their association with England. 'Duke's Day', a public holiday commemorating the day the Duke of Edinburgh visited the island in 1972, is celebrated annually. Local lore has it that he praised the Palmerston Island dialect and said it was excellent English. Moreover,

Palmerston Islanders claim that those who spoke with him understood his English and were themselves more easily understood than English-speaking tourists who were also present. During national celebrations in Rarotonga in 2008, when the representatives from each island wore clothing representative of their heritage, the Palmerston Island contingent dressed as Buckingham Palace guards.

The islands belonging to the Cook Islands are divided into the Northern Group and the Southern Group. While Palmerston officially belongs to the Southern Group, it is the most northern of these, and has more in common geographically with the northern islands, as these, like Palmerston, are true atolls, while all the other southern islands are raised volcanic atolls. It also has the sweeter variety of coconuts that are otherwise found only in the Northern Group. Palmerston Islanders generally feel that Palmerston should either be included in the Northern Group, or not be considered part of the Cook Islands at all. The island is governed by a council, made up of the eldest member of each of the three families and a second member from each, appointed by the eldest.

3. Language ideologies and linguistic history

There are at least two different types of relationship between language ideology and linguistic history. The first is reconciling traditional ideas about language change, the origin of the local language/dialect and its links to other varieties in the region with what linguists can find out about these questions through other means, for example documented history, linguistic reconstruction, comparison or similar objective tools. The other relationship between language ideology and linguistic history is that a community's self-identification and choice of affiliation with other communities can affect its speakers' linguistic choices, and consequently can play a role in determining the evolution of the language. This can occur explicitly, for example through the socialization of children, but it can also be mediated by the effects of identity and affiliation on social networks. If speakers who identify more with the Cook Islands spend more time with visitors or immigrants from the Cook Islands, while speakers who identify more with England spend less time with these, but more time with English visitors, then each group will have more opportunity to draw on the linguistic resources of the variety they have more exposure to, such as borrowing new terminology from it, or accommodating

phonetically to it. This also means that the two groups socialise less with *each other*. Innovations in one group are therefore more likely to be transmitted within the group than across the boundary to the other group.

Moreover, a community's linguistic and cultural identity is likely to inform the stories it tells itself and outsiders about the history of its language. For that reason, even if local lore about the language's history and its relationship to other varieties is different from what conventional linguistic research would tell us, it is still valuable, perhaps essential, if the linguist wants to explore the question of *why* the language has developed as it has.

In this section I will discuss three relationships between Palmerston Island language ideologies/cultural identity and certain characteristics of the linguistic variety. The first is the island's tendency to emphasise its English roots over its Polynesian background, and the remarkably low incidence of Māori substrate features in the dialect. The second is the complete absence of the island's Portuguese influences, both when explicitly talking about the island's linguistic and cultural history, and in the linguistic characteristics of Palmerston Island English. The third is the correspondence of extreme variation in linguistic features with variation in the historical narratives told about the island.

Polynesian substrate

The first relationship between Palmerston Island language ideologies and linguistic features I will discuss is the emphasis on English over Polynesian[2] heritage, both culturally and linguistically. As mentioned in the introduction, many Palmerston Islanders are very proud of their English heritage. They emphasise Marsters as their ancestor, and their stories and songs about the island's history are about him, rarely mentioning his wives by name. My hosts frequently pointed out to me those physical features that they consider English and that they assume they have inherited from Marsters. Public holidays are Duke's Day, Queen Victoria's Birthday, Easter, and Christmas Day, all of which are associated with the English side of the island's history. Some traditions enforced by Marsters during his lifetime—daily evening prayer, morning rows across the lagoon, sweeping the beach and the paths, early rising (2 am)—continued until recently, well beyond Marsters' death; and because they were instigated by Marsters, they are also considered to be English in nature.

There is little explicit mention of the island's Polynesian heritage in my transcripts of conversations about the island's history, culture or language.

In some conversations, a Polynesian connection seemed almost to have been forgotten until I reminded the speaker about Marsters' wives:

RH   I also heard that people used to use a lot more Māori words here: they used to mix the languages a bit. Use some words. Is that true, do you think?

OB   What? How do you mean?

RH   That people used to use more Māori words on Palmerston Island for things, sometimes. And mix the languages a little bit. Use some words from Māori; some words from English.

OB   I don't know.

RH   You don't know, yeah okay.

OB   So far we all English speakers.

RH   But originally, William Marsters' wives... They spoke Māori, didn't they?

OB   Oh yeah.

Underlyingly, however, there are many threads of Cook Island Māori culture that run through everyday Palmerston life. Houses are thatched with woven palm branches, which are also used for making straw hats. The diet is primarily fish, coconut, taro and rice, with staple recipes that are found elsewhere in the Cook Islands and Polynesia (e.g. raw fish, *poke*) As is usual across Polynesia, white clothing is worn to church when communion is celebrated. Women always wear hats to church, and men and women sit on separate sides of the aisle. Two different types of hymns are sung: *Sunday school hymns*, including some that are widely known across the English-speaking world, which are sung in English with a single melodic line, and *traditional hymns*, which are known elsewhere in the Cook Islands as *'imene tuki*, which are always sung in Māori in a chanting style and have complex multiple parts with different words and melodies. Palmerston Islanders memorise the Māori words to these without usually knowing their meanings. The *traditional hymns* are taught and practised in weekly prayer and song meetings known as *'uapo*, which is also a tradition across the Cook Islands.

    Until recently, when they were replaced with modern boats fitted with

outboard motors, canoes were made by hollowing out *tamanu* trees. These are outrigger canoes similar in style to the traditional canoes of other islands in the region.

Land division, hunting and gathering rights, and use of land for building on are all governed by similar principles to those found in other Cook Island communities, as can be seen in Crocombe and Marsters (1987), a thorough description of the rules of land use on Palmerston Island. Although the focal point *from* which descent in the three families is reckoned is the three original women, one's ancestry is calculated *through* the patriline. In other words, all Palmerston Islanders define themselves as belonging to the 'family' of either Akakaingaro, Tepou, or Matavia (the three wives of Marsters), but they trace this ancestry through their male ancesters. Patrilineal descent is a common feature of Polynesian societies, as is the co-existence of this with important female ancestors from which one's lineage might be calculated (cf. Gunson 1987).

Roles of men, women and children on Palmerston Island are also reminiscent of other Polynesian societies. Children are expected to do a large amount of the daily work: sweeping, feeding the animals, cutting up coconuts, making drinks for their parents and for guests, fetching and carrying water, taking messages, and cleaning. Women bake bread most days, prepare the rest of the food except on rare occasions, sweep, wash clothes and dishes – both of which are labour-intensive tasks without running water. Men do most of the fishing and hunting of birds. Children are expected to obey parents and other older family members immediately and without question.

Some of these Polynesian cultural practices (for example the diet, the use of palm leaves) could be argued to be practical necessities due to what is available on the island. Others (division of labour, religious practices, music, particular preparations of food) are not born of necessity and instead might well be retentions from the practices of Marsters' own wives.

Similarly there appears to be a Cook Island Māori substrate in the language, but this sits below the level of consciousness. The lexical borrowings from Cook Island Māori are relatively few and mainly limited to the semantic domains of flora and fauna and religion, but there is also some influence in the phonetic, phonological, morphological and syntactic domains of the language. For example, Palmerston Island English has variation between [v] and [w], for example 'vase' [waz~vaz], 'over' [ɔwə~ɔvə]; and Palmerston [t] or [d] (word-initially) and [s] (word-finally) correspond to standard English interdental fricatives (e.g. Palmerston [tem] or [dem] 'them', [tis] 'teeth').

This is all likely to be due to influence from Cook Island Māori, most dialects of which have few or no fricatives. Simplification of consonant clusters to a single consonant is probably a result of Cook Island Māori's CV syllable structure (e.g. *roun* 'round', *ol* 'old' *mos* 'most', *is* or *it* 'it's/its').

A distinction between singular, dual and plural pronouns is almost always made. Forms for expressing the dual category, as well as the singular/plural distinction where standard English does not express this, are created using the resources of standard English, for example *dem two* 'third person dual' or *dem lot* 'third person plural'. In addition, the inclusive/exclusive distinction may be marked in conjunction with dual number in the pronoun *yami*, 'first person dual inclusive', although it is not used as frequently as in the past. The pronouns of Palmerston Island English therefore mirror the system of Cook Island Māori almost exactly (see Hendery & Ehrhart In Press for a side-by-side comparison of the two paradigms).

Cases in which Palmerston Island word order differs from that of standard English (but aligns in some constructions with that of Cook Island Māori) include fronting and topicalisation constructions, some compounds and some possession constructions. Some of these are illustrated in (1).

1. a. two barrel diesel Korinako's
        'two of Korinako's barrels of diesel'

    b. Too small the table
        'The table is too small'
    c. blood pig
        'pig's blood'

A lack of copula, lack of plural marking on nouns, and optionality of explicit subject (pro-drop, e.g. *It's really fun when hear them speaking* 'It's really fun when you hear them speaking') are features of both Palmerston Island English and Cook Island Māori.

The only one of these linguistic features that was volunteered to me as an example of Palmerston Island English is the use of the pronoun *yami* for the first person dual inclusive. Otherwise they either pass under the conscious radar of the speakers, or are not considered noteworthy. As a comparison, the use of particular tense and aspect forms (double-marking of past tense verbs, for example *passeded*, *blesseded*, formation of present participles with *-en*, for example *fishening*, *singening*), were all pointed out to me numerous times as examples of Palmerston Island English. While Māori-origin lexical

items were sometimes explained to me (*motu* 'island', *'uapo* 'singalong', *tamanu* 'mahogany', *para* 'wahoo', etc), they were never given when I asked explicitly for words that Palmerston Islanders use that are 'different from the way other English speakers talk'. In answer to this question instead I was given terms like *fowl* 'chicken', *bong* 'lid', *basin* 'bowl', *for'ard* 'in front', *yonder* 'over there'—words that the islanders explicitly associated with Marsters and with his (Northern) English dialect.

Similarly, English-focussed cultural practices were often explained to me unelicited—Duke's Day, Christmas celebrations, early rising, the English-only policy—while the Polynesian practices were not discussed. Some of this may be because the islanders are aware of the uniqueness of their heritage through Marsters, but believe some of the Polynesian practices to be ubiquitous (this is certainly the case for wearing white to church—Palmerston Islanders were surprised when I told them this is not a practice everywhere in the world.) It may, however, also be the case that their pride in their English heritage and Marsters' centrality to the story of their origins overshadow other contributions to the language and culture and are therefore foremost in the islanders' minds when linguistic and cultural matters are discussed.

The Portuguese question

According to Helm and Percival (1973: 105–108), the island was originally settled not only by William Marsters with his wives, but also by a Portuguese friend of Marsters, Jean Baptiste Fernandez (or Fernandos) and his wife and child. After Fernandez's death Marsters took Fernandez's wife as his third wife, and adopted Fernandez's children. Helm and Percival report that Fernandez was 'variously described as a "Hindu-Portuguese" from Goa, a "Portuguese sea-cook" and a "native of the Fernando Noronha Islands off the east coast of Brazil"' (106–108). While these descriptions are in quotation marks, Helm and Percival do not give sources for them. Whether Fernandez's native language was Portuguese or an indigenous language of Goa or Brazil, it is likely that his L2 English would have been quite different from the English of the Cook Islanders, and also different from Marsters' L1 English, containing features that could have been transmitted and retained in Palmerston English, just as was the case for Marsters' Northern English dialect or the others' Cook Island L2 English. Fernandez remained for several decades of the early period after settlement, and as his family made up a large percentage of the original population, one might expect that it would have left some traces, possibly in the language and culture, and certainly in the oral histories. The

histories of the island told to me did not, however, mention him at all. Nor does he appear in the traditional song about the island's settlement that the older islanders remember their grandparents (second-generation Palmerston Islanders) singing:

> Oh William Marsters was a brave hero.
> He farewelled to happy Londontown.
> He has been to Birmingham
> where he was,
> in the beautiful country of him.
> He took a joyous voyage,
> and saw pretty countries.
> While on a ship he has passed Tongareva,
> and so called in.
> And when leaving with his [family?][3]
> he came to Palmerston,
> living as the married king for the rest of his day.

There are several other similar songs, all of which focus on William Marsters as the founder of the island, and none of which mention Fernandez at all. It should be noted that all the islanders have the surname Marsters and trace their ancestry to Marsters and one of his wives; none have retained any connection to Fernandez.

The only traces of Fernandez in the materials available to me are in references in the early court records to a woman *Tati Fernandos*, and in John Burland's interview with Ned Marsters in 1959. Ned Marsters tells John Burland the following, but only when directly asked about Fernandez:

NM   He [William Marsters] was a whaler – stayed in America for the gold digging. From the gold digging he came over to Penrhyn and stayed in Penrhyn and get the girl there – his wife – and so they came on shifting down to Samoa – this time he went to Samoa.

JB   I see. And then when he went to Palmerston Island he had not only his wife but the second girl as well?

NM   Yes.

JB   They both went together?

NM   Yes. Both went together. The cousin – his wife and the cousin of his wife. And Fernandez and his wife.

JB   Oh, they all went together?

NM   They all went together.

JB   This is John Fernandez?

NM   John Fernandez and his wife.

JB   And John Fernandez was a Portuguese, was he not?

NM   Yes. A Portuguese half-caste – half white.

This interview is the only reference to Fernandez I have come across directly from Palmerston Islanders.

Similarly absent are linguistic features that are unambiguously attributable to Portuguese influence or the influence of any other language that Fernandez might have spoken. All of the characteristics of Palmerston English can be explained as retentions from William Marsters' Northern English dialect, influences from his wives' Māori varieties, innovations introduced by the wives' L2 English, or recent influences from Cook Island English. There are a few features that exist in both Cook Island Māori and Portuguese, so could have been brought into Palmerston Island English from either (or, perhaps more likely, from both). One of these is pro-drop, the possibility of eliding the subject (or less frequently the object) of a clause. This sort of pro-drop is allowed in both Cook Island Māori and Portuguese, so it is impossible to know for sure which one was the source of the feature in Palmerston Island English. If it were found in the L2 English of all the early Palmerston Island settlers, both Portuguese and Polynesian, it could even be said to derive from both.

Apart from features like this that could just as easily have arisen without the contribution of Portuguese learner English, there are no traces of any linguistic influences in the language today that might be due to Fernandez. The coincidence of this with the absence of Fernandez from the local historical record is striking. An explanation could be that the islanders have identified with Marsters and with English, and actively disassociated themselves from any Portuguese linguistic and cultural history. Because of this the earlier generations may have selected variants that they recalled Marsters himself using, preferring these over alternative 'non-English' variants. Even today there is a certain amount of awareness of which features of the dialect are

especially 'English': for example the pronunciation of the some words that belong to the STRUT lexical set in many English varieties as [ʊ] (the same as FOOT). We can only speculate about the reasons behind this *erasure*, but the results are a common ideological phenomenon: 'in simplifying the sociolinguistic field, [it] renders some persons or activities (or sociolinguistic phenomena) invisible. Facts inconsistent with the ideological scheme either go unnoticed or are explained away' (Irvine and Gal 2000: 38).

In this and the previous section we have seen that the relative influences of English, Cook Island Māori and Portuguese on Palmerston Island English seem to correlate with the islanders' orientation to each of these sources culturally and historically. William Marsters is the major founder figure in their narratives, the main person to whom they trace their ancestry, and because of this they consider themselves English. English is unquestionably the major source of the modern dialect. Their Cook Island Māori heritage, on the other hand, is less foregrounded in their historical narratives and in their deliberate representation of themselves and their language to outsiders. While it has a clear influence on some features of the language, this is not sufficiently pervasive to justify characterising Palmerston Island English as a mixed language. The increase in interaction with New Zealand over the past few decades (through travel, aid, and visitors) is mirrored in some elements of the language that have probably come in from New Zealand English: the use of the tag *eh?*, an extension of *like* to a focussing function, and the formation of future tense with *go* are all features that have appeared in Palmerston Island English in the past twenty years (i.e. are not found in recordings made by Sabine Ehrhart in 1991), and can plausibly be explained as influences from New Zealand English. The Portuguese element in the island's history is almost completely ignored, and similarly has had little or no influence on the language. In this, Palmerston Island demonstrates in microcosm the relationship between language and ethnicity that is also found in larger, longer-established states (cf. Blommaert 2005: 214–217). This too is a kind of ideology: 'we are English and therefore we speak English'/ 'we speak English and therefore we are English'.

Variation

One of the most striking features of Palmerston Island English is the amount of variation that is found, from speaker to speaker, and also in the speech of a single speaker. Variation is found at all linguistic levels: phonetics, phonology, morphology, syntax and lexicon. It is not only a matter of choosing

local variants or standard English variants, but also sometimes of choosing between multiple local variants, for example the use of the terms *moa*, *fowl*, or *chicken*. It appears that some variation is governed by gender (e.g. pro-drop is much more heavily used by men than by women), and some by age (use of singular forms of the verb *be* with plural subjects, e.g. *they is*, *we is*, *you is* is more frequent with younger speakers than with older people). At least some of the variation found in the speech of a single speaker is conditioned by context (i.e. register) and by identity of interlocutor (accommodation). The average use of certain typical Palmerston Island features when aggregated[4] does in fact turn out to be different between the aforementioned *bush people* and the *beachfellas*. However, a large amount of the inter-speaker variation is unaccounted for, and may be the same sort of phenomenon identified in small communities of Gaelic speakers by Dorian (1994), and termed *personal-pattern variation.*

Besides this linguistic variation, variation is also found in the linguistic ideologies, cultural identities, and histories of the island. Some speakers consider themselves, or are considered by others, to be 'more English', as the following quote suggests.

> They don't say it [*shubble* for *shovel*], cos they're Engli- kind of closer to English, like MM and, uh, PT and them there.

Some people seemed to be especially proud of their English heritage, pointing out physical features such as heavy body hair, skin colour, and stature that they associate with the English body type. This also comes up in the earliest interview we have: Burland's interview with Ned Marsters, William Marsters' grandson, in 1959:

> NM   My grandfather, he has the same look as my father, the
> same build as my father, and hairy the same as my father.
> […] Full beard, and this part – well the body is all black
> hair. Oh, yes. Well, the boys, they got all full hair in front.

Other speakers never mentioned England or the English in my hearing at all. For some Palmerston Islanders, their association with New Zealand seems to be more salient than their English connections, as they spontaneously discussed recent New Zealand politics, visits to New Zealand, and New Zealanders they had met, rather than talking about England. These topics may

have been emphasised because of my presence, as a New Zealander myself, but the level of knowledge of those who brought up New Zealand news and politics was far beyond that of most non-New Zealanders. The information seems to be gained from visitors, letters and (rare) phone conversations with family members living in Rarotonga, New Zealand or Australia, and news brought back by Palmerston Islanders who have been off island for medical reasons or to visit family.

Each family has its own version of Palmerston Island's history, and those speakers who shared theirs with me were very careful to avoid claiming that their version was the only correct one, peppering them with hedges like the following:

'To be honest, I don't really know that.'

'What I heard that he... All right. What I heard is...'

'So I don't know how true it is. Really don't know.'

Some of them are certain that the history outlined in Hilyard's biography of the Marsters family (Hilyard 2008) is incorrect ('all lies'), while others accept this version.

As well as these variations in the accounts of the island's history, there is variation in the feelings towards Palmerston Island English. There is widespread understanding that outsiders see it as a substandard dialect:

CC   yeah I've noticed that when I live one and a half year in
      NZ **they used to complain how I talk**. See, I used to work
      as a receptionist in the school an' I u- I always used to say
      'you welcome' never 'you're welcome', and they used to
      tell me off not to say that.

CC   I try a speak **proper English** cos it's goin' be hard in
      school cos our teacher tell us off when we're speaking not
      proper English

MM   But I try to remember them [*standard grammar rules*] and
      try to do it the way she's [*the head teacher*] saying it, cos
      then it's so when I speak to the children it be the way that
      she's wanting.

Some Palmerston Islanders themselves seem to agree with this negative assessment of their dialect:

AP   Palmerston slang, and New Zealand slang and, uh, the other two are **slack slang**.

CC   I don't talk **proper English** to my parents.

JM   They [*the children*] try to say the words correctly. […] Well to me the teachers is more like trying to get them to say **the right words**.

VN   But I think we been morely taught how to speak a bit more **proper** than how we usually talk.

Others are proud of their variety of English and actively encourage it in their children, as is shown in the following exchanges:

OB   Have you hear of a word 'bugger'? You heard the word 'bugger'? Some of us use it in the island. [...] Oh just like when the childrens make something, make the mistake at home or do something wrong, our father use say it, 'You bugger you!' Something like that. I'm not sure how it come, but my grandfather and grandmother speaks it and that's how I pick it up, and is still in me and **I'm proud to speak it**.

LM   To me there's really no proper way o' speaking, it's just the way you speak. If you been understood, then **that's the proper way**.

LM   And a lot of Palmerston people gets that feeling when they go to Rarotonga: The way we speak, they probably think that we are not educated enough how to speak English properly. I guess that is probably **them the one is not educated**. I think for them to be thinking that: about judging other people how they speak.

OB   I speak my language **the old way**. And my children's learning it and we understand ourselves when we speak.

One resident even reported that he discouraged his children from using at least one of the features of Standard English: he tells the children to say [d] instead of the [ð] that the teachers tell them to use, because otherwise 'it make them sound like faggots'.

Many of these assessments of Palmerston Island English take place in a context of comparison to other varieties of English and references to spaces in which Palmerston Islanders come into contact with these varieties: Rarotonga, New Zealand, and even the Palmerston School (where the head teacher is a New Zealander, and the educational materials include books in New Zealand English, American English and British English).   In their discussion of multilingualism, space, and globalization, Blommaert et al. (2005: 198) note that 'the particular environment organizes a particular regime of language, a regime which incapacitates individuals.' While Blommaert et al. are referring to multilingual incapacitation, the same can apply at the level of dialects. Some Palmerston Islanders respond to this sense of incapacitation by concluding that they don't have 'proper' language or the 'right' language, but others push back and define their variety as something to take pride in, and/or deny the existence of a single right way to speak at all.

It is not surprising if an ideology of a single standard English that everyone should aspire to is not universally held on Palmerston Island, because of the timing of original settlement of the island. The belief in the importance of a standard language was an ideology that arose in Victorian England as an offshoot of nationalism and belief in the 'purity' of England's Germanic origins (cf. Milroy 2006). Marsters left England long before this movement was at its peak, and certainly well before the concept of an ideal standard language would have spread into the consciousness of the general public. This may well be part of the reason why we find Palmerston Islanders who are proud of their dialect and adamant that 'there's really no proper way o'speaking'. From the attitudes of a few, however, it is clear that the idea of a prestigious Standard English has taken some hold in the community: most likely via those islanders who have spent time abroad, and through the influence of the school (cf. Collins 1996: 204–206 for a discussion of the way in which such ideologies underpin the typical Western educational experience.)

We see that there is variation in local histories, cultural identity, and language ideology. It would be surprising if there were no relationship between these three types of variation. For one thing, a person's cultural identity informs his or her selection among variants of traditional histories: if a person identifies strongly with the English, it makes sense that he or she emphasises those stories about his or her family's history that involve the English and de-emphasise others. Cultural identity in turn influences language ideology: most people would prefer to believe that their language use reflects the culture they identify with. People who are proud of their cultural identity

are also proud of linguistic features that they believe are uniquely associated with that culture. Furthermore, if one's family has always emphasised historical traditions in which the role of certain cultures is highlighted and that of others is downplayed, it is only natural that one will believe that 'traditional' ways of speaking are also associated with those cultures, and will be proud of these.

All of these connections between history, identity and language would apply whether the community in question is homogeneous or as full of variation as Palmerston Island is. Because there are such strong relationships between history, identity, culture and language, however, variation in the community in any one of these aspects naturally goes hand-in-hand with variation in the other aspects. If one family identifies more with the English, while another sees Palmerston Island as unique and independent, the first is likely to emphasise stories about the community's history in which the English play a role, while the second is likely to downplay these. It would then be quite natural for the first family to prefer and be proud of linguistic features that have clearly English origins, while the second uses more of the innovative lexicon or structural features of the dialect. The first family might be proud of the English features of the dialect, but embarrassed by the 'non-standardness' of Polynesian-inspired substrate features. In this way, diversity of cultural identity can lead to variation in linguistic features and language ideology.

Several studies of other communities have shown links between language attitudes or ideologies and actual use of specific variants. For example Sharma (2003) finds quantitative evidence for a relationship between L2 Indian English speakers' linguistic ideologies and their use of American or Indian phonological variants. Wassink and Dyer (2004) demonstrate that change in phonological variants in two communities has gone hand-in-hand with changing attitudes to these variants and to the people who use them. Milroy (2004) discusses a number of other studies that have shown a relationship between attitudes to language and linguistic change.

This all assumes a great deal of linguistic awareness on behalf of the community members, but I do not think this assumption is necessarily unrealistic in the case of Palmerston Island, given the amount of thought most islanders seem to have given to linguistic matters, and their interest in their cultural and linguistic history. Many community members, for example, showed a keen interest and engagement in my fieldwork methodology. The following examples, from separate conversations, all show an awareness of the problem of accommodation.

TP   Some of them, they know the recorder's there and they might never use the word, you know? They might just talk normal only.

MM   If you saw two Palmerston people speaking it would be slightly different the way they would speak to you in a way.

CC   They would say 'shubble' sometimes, only when people like you are not around their house.

In the following examples, we can see how Palmerston Islanders are aware of the origins of some words, distinguishing loan words from Māori, older English dialect terms, and items unique to Palmerston Island.

AP   Wahoo. *Para*, we call it. *Wahoo* I think is a English name.

AK   A couple of trees out the islets, as we say, *islets*, but we call it *motu* in the Māori, but *islets* in English

OB   *Deye*: I don't know how you pronounce it. But our fathers use it. It's not a Māori word. It's just the word being used on the island.

FN   My cock and pullet.[5]
SM   That's a old English.

AK   The big trees, the *tamanu* – they say the *mahogany* in English. Well, I'm not sure if that's the same tree, but some of them say the *mahogany*.

There is also widespread awareness of linguistic change. The following extracts are all from a conversation with one woman, SD, who is describing the differences between how her grandparents spoke, how she speaks, and how the children speak today. The number of details given show that she has very clear ideas about which elements of the dialect are older, which are newer, and whether they are associated with England, Rarotonga or New Zealand.

But when we got to school, going to school, then there was a different of how we're pronunciation is going to be. I know when we will use the word *p-u-t*. And *b-u-t*. Knowing the phonics, you had two: [ʊ] And a [ʌ]. And I can remember –

I know my mother, my grandmother used to say no, don't say [bʌt], it's [bʊt]. It's [bʊt]. Like 'put', because you say [ʊ]. You say not [bʌt], it is [bʊt] And 'put' is [pʊt]. See these are the difference of their pronunciation.

[...]

I can remember the words that they use. Like when they say, we say, 'You go over there and get something'. They say, 'You go yonder'. And that's how they use, 'You go yonder'. And then sometimes when you s- 'Oh, look yonder. How's the – where's the sun?' Y'said, 'Oh, there's the sun up there.' 'So, well do your work before the sun dips beyond the horizon.' So this is how they say there was, uh, the sun sets. But my grandmother would say, 'Before the sun dip beyond the horizon.' This is how the language my grandmother use. And I think it's a real perfect English.

[...]

So I learn[ed] two different languages, from the school, and from my grandmother.

[...]

So there's two different pronounciation: [hu] and [u] Because our grandparents don't say [hu]. It's [u]. That 'wh', they don't pronounce it. That's right. That's the school uses. So that's how the different languages been change. And so the children nowadays they using what our [u] just like that, that's how they believe 'h' into it.

[...]

Well that's so because people go to New Zealand and stay for a while. Then they come back home. Yeah, well they use the words from New Zealand, they come back home. It goes!

As the community is interested in language, aware of the various sources of influence on their dialect, and, at least lexically and phonetically, is aware of which linguistic choices signal which cultural affiliations, it therefore seems reasonable to associate the diversity of historical perspectives and cultural identities on the island with the diversity in language ideologies and linguistic variation.

Conclusions

Not only does Palmerston Island provide an interesting case study of the close relationship between linguistic and cultural identity, it also shows how various cultural influences can be teased apart when they are compared with similarly intertwined linguistic influences and language ideologies. Palmerston Island is also a reminder that there can be an enormous diversity of language ideologies, cultural identifications and historical understandings even in the smallest communities. This diversity may reflect linguistic diversity, and a causal relationship between these two types of variation cannot be ruled out.

In Palmerston Island, a mixed-origin group of settlers has had to co-construct a new society, with all that entails linguistically, culturally, and socially. If the social and linguistic organisation of Palmerston Island reflect each other, it is because they have evolved together. The cultural and linguistic elements that combined in the original settlement are not, however, the same as those that are most apparent in the community today. There is almost no trace of the Portuguese settler who played an important role in the founding of the community. Why the islanders have made cultural and linguistic choices to pattern themselves after the 'English' model instead of the Portuguese one is not a question that can be answered with 100% certainty. Narratives about William Marsters present him as extremely charismatic and authoritative. The only story passed down about Fernandez presents him in a rather weak light. Perhaps Marsters was a more obvious model for the men to emulate. Moreover, Marsters had three wives, while Fernandez had only one. This suggests that the majority of the children in the first generation would have been Marsters' offspring, not Fernandez's. Even Fernandez's children may have had more experience of Marsters than of their own father, since Fernandez seems to have been away travelling for several years at a time, and Marsters eventually took Fernandez's wife and children as his own. All of these factors mean that Marsters was probably the strongest and the most present father-figure the boys of the first island-born generation had available.

In terms of sheer numbers, the Polynesian influence on the island should be far greater than the English influence (see e.g. Parkvall 2012 for a model in which the settlement demographics of Palmerston Island would predict a Cook Island Māori-based creole). This is where we can clearly see the effect of power. Marsters was the head of the island ('the married king', as the song puts it). It is said that he instructed his wives and children to speak

only English, but the historical record shows this may be an exaggeration, as letters and records from the early 1900s sometimes contain Cook Island Māori and references to Cook Island Māori use, and an English missionary who visited Palmerston Island in 1877 reported that he preached in Rarotongan in Marsters' presence (Gill 1877). The influence of Cook Island Māori on the dialect is relatively minimal, however, with some phonetic influence, some morphological, and a few word order patterns. At least as many of the differences from Standard English can be explained as innovations or retentions from Marsters' own English dialect.

As is seen in the quotations from Palmerston Island residents above, construction of linguistic and social norms is not entirely subconscious: the community is aware of the different origins of lexical items, and the cultural and social affiliations signalled by different linguistic choices. Both school and parents deliberately attempt to regulate the language use of the younger generation – sometimes in conflicting directions.

The complexity of the linguistic landscape, the multiplicity of language ideologies, and the rich network of social and cultural affiliations that persist on the island are only visible because it is possible to study Palmerston Island language and culture at the level of the individual actor. Abstractions away from the individual are necessary when one is writing 'the' grammar of 'a language', or when one compares one community to another, but as this paper has demonstrated, such abstractions are by no means the end of the story.

Notes

1 The term *family* on Palmerston Island usually refers to this wider grouping: that is all descendants of the one female line. For a nuclear family group, the term *household* is more often used.

2 My reference to 'Polynesian' or 'Cook Island Māori' heritage in this section is intended to include Penrhynese heritage. While Penrhynese and other Cook Island Māori varieties are generally considered to be separate languages, Penrhynese is neither different enough, nor well enough described for us to be able to say whether apparent substrate influence in Palmerston English is from Penrhynese specifically or from other varieties of Cook Island Māori.

3 There is some debate about the word *family* in the final verse. Some other suggestions by the locals include *spirit* and *ferret*.

4 To test this I took 50 sentences from each speaker, and counted the frequency of subject pro-drop, bare nouns used for semantically plural referents, verb forms with -*s* used with non-3s subjects, and bare present tense verb forms used with 3s

subjects. For each speaker a percentage use of these features was then calculated by dividing the number of times these features appeared in the speech sample by the number of times it would have been possible to use it in that sample and multiplying by 100. The difference between these percentages for the *bush people* and the *beachfellas* was small but statistically significant (p<0.05).

5   This is the punchline to a dirty joke, but it only works because the words *cock* and *pullet* are still in regular use on the island in the sense of 'chicken'.

References

Blommaert, Jan. 2005. *Discourse: A Critical Introduction*. Cambridge: Cambridge University Press.

Blommaert, Jan, James Collins and Stef Slembrouk. 2005. 'Spaces of multilingualism.' *Language & Communication* 25: 197–216.

Burland, John. 1959. Unpublished transcript of an interview between John Burland, Ned Marsters and Peka Marsters. Wellington, New Zealand. Alexander Turnbull Library. MSX-8809.

Cameron, Deborah. 1990. 'Demythologizing sociolinguistics: Why language does not reflect society.' In John Early Joseph and Talbot J. Taylor (eds) *Ideologies of Language*. New York: Taylor & Francis. 79–96.

Clark, Victor. 1960. *On the Wind of a Dream: the Saga of Solace*. London: Hutchinson.

Crocombe, Ron G. and William F. E. Marsters. 1987. 'Land tenure in a test tube: the case of Palmerston Atoll.' In Ron G. Crocombe (ed.) *Land Tenure in the Atolls: Cook Islands, Kiribati, Marshall Islands, Tokelau, Tuvalu*. Suva: Institute of Pacific Studies, University of the South Pacific. 202–233.

Dorian, Nancy C. 1994. 'Varieties of variation in a very small place: Social homogeneity, prestige norms, and linguistic variation.' *Language* 70, 4: 631–696.

Ehrhart-Kneher, Sabine. 1996. 'Palmerston English.' In Stephen A. Wurm, Peter Mühlhäusler and Darrell T. Tryon (eds.) *Atlas of Languages of Intercultural Communication in the Pacific, Asia and the Americas, Vol II*. The Hague: Walter de Gruyter. 523–532.

Gill, William Wyatt. 1877. Unpublished letter to the London Missionary Society regarding visit to outstations July 3–Aug 14. LMS archives, School of Oriental and African Studies, London. CWM South Seas Journals Box 11. 1871–1881.

Gunson, Neil 1987. 'Sacred women chiefs and female 'headmen' in Polynesian history.' *The Journal of Pacific History* 22: 139–172.

Helm, Arthur Stanley and W. H. Percival.1973. *Sisters in the Sun: the Story of Suwarrow and Palmerston Atolls*. London: Hale.

Hendery, Rachel and Sabine Ehrhart. 2011. 'Palmerston Island English.' In Bernd Kortmann and Kerstin Lunkenheimer (eds.) *The Electronic World Atlas of Varieties of English* [eWAVE]. Leipzig: Max Planck Institute for Evolutionary Anthropology. http://www.ewave-atlas.org/, accessed on 2012-10-02.

Hendery, Rachel and Sabine Ehrhart. In press. 'Palmerston Island English.' In Bernd Kortmann and Kerstin Lunkenheimer (eds.) *The World Atlas of Varieties of English*. Berlin/New York: Mouton de Gruyter.

Hilyard, Maureen Marama. 2008. *The Marsters of Walcote, Leicestershire: The Family of William Richard Marsters of Palmerston, Cook Islands*. Self-published.

Irvine, Judith T. and Susan Gal. 2000. 'Language ideology and linguistic differentiation.' In P. Kroskrity (ed.) *Regimes of Language*. Santa Fe: SAR Press. 35–83.

Milroy, James. 2006. 'Historical description and ideology of the standard language.' In Laura Wright (ed.) *The Development of Standard English, 1300–1800: Theories, Descriptions, Conflicts*. Cambridge: Cambridge University Press. 11–28.

Milroy, Lesley. 2004. 'Language ideologies and linguistic change.' In Carmen Fought (ed.) *Sociolinguistic Variation: Critical Reflections*. Oxford: Oxford University Press. 161–177.

Parkvall, Mikhael. 2012. 'Simulating the genesis of Mauritian'. Paper presented at the *Ninth Creolistics Workshop* in Aarhus, Denmark, April 2012.

Rumsey, Alan. 1990. 'Wording, meaning, and linguistic ideology'. *American Anthropologist*. 92: 346–361.

Sharma, Devyani. 2003. 'Structural and social constraints on non-native varieties of English.' Unpublished PhD dissertation. Stanford: Stanford University.

Wassink, Alicia Beckford and Judy Dyer. 2004. 'Language ideology and the transmission of phonological change changing indexicality in two situations of language contact.' *Journal of English Linguistics* 32: 3–30.

Woolard, Kathryn A. 1998. 'Language ideology as a field of inquiry'. In Bambi B Schieffelin, Kathryn A. Woolard, and Paul V. Kroskrity (eds.) *Language Ideologies: Practice and Theory*. Oxford: Oxford University Press. 3–48.

TOWARDS A CORPUS OF EARLY WRITTEN NEW ZEALAND ENGLISH – NEWS FROM *EREWHON*?

Marianne Hundt: Englisches Seminar, Universität Zürich, Plattenstrasse 47, CH-8043 Zürich, Switzerland, <m.hundt@es.uzh.ch>

Abstract

The history of New Zealand English is well attested. Previous studies focussed on the development of the New Zealand accent (Gordon et al. 2004) and are based on spoken data. Written data would enable linguists to study the emergence of standard New Zealand English (NZE) and differential change in this variety vis à vis British and American English. The present article discusses the requirements that such a diachronic corpus of written NZE should meet and presents a case study on the use of the progressive. The data from the Corpus of Early New Zealand English (CENZE) show that the frequency with which the progressive is currently used in NZE is a very recent development that is unlikely to be attributable to influence from Irish English (IrE) during the colonial period.

1. Introduction

Erewhon is the title of a novel by Samuel Butler that was published in 1872. Reading the title of this novel backwards provides a clue for the literary genre. Moreover, on the basis of Butler's biography and the descriptions of landscapes, this utopia has been localised in the south island of New Zealand. However, *Erewhon* turns out not to be the sought-after Atlantis but a place that

allows Butler to project some shortcomings of Victorian England. The hero of the novel, a farmer called Higgs, therefore leaves *Erewhon* disillusioned in a hot air balloon. The connection between Butler's novel and the potential corpus of Early New Zealand English (NZE) is that it forms part of a collection of early New Zealand texts that were digitized at the *New Zealand Electronic Text Centre* at Victoria University of Wellington (VUW). I learnt about their efforts to digitize texts during a stay as a visiting professor at VUW and contacted the head of the electronic Text Centre, Alison Stevenson, who made their texts available to me. Additional material was obtained from the National Library (namely the *Proceedings der Royal Philosophical Society of New Zealand*) and from the world-wide-web (mainly newspaper texts and an early letter collection).

Obviously, a collection of digitized texts is not automatically a corpus. As Biber et al. (1998:4) point out, a corpus is a "[…] large and principled collection of natural texts." How one might get from a collection of digitized texts to a representative corpus of early New Zealand English and why such a corpus might be useful for linguists will be the topic of this article.

Apart from the utopian (or rather dystopian) tenor, Butler's novel also makes use of satirical elements, so I might be permitted to quote Fillmore's (1992: 35) caricature of two extremist approaches to the study of language at some length:

ARMCHAIR linguist:
He sits in a deep soft comfortable armchair, with his eyes closed and his hands clasped behind his head. Once in a while he opens his eyes, sits up abruptly shouting, "Wow, what a neat fact!", grabs his pencil, and writes something down. Then he paces around for a few hours in the excitement of having come still closer to knowing what language is really like. (There isn't anybody exactly like this, but there are some approximations.)

CORPUS linguist:
He has all of the primary facts he needs, in the form of a corpus, of approximately one zillion running words, and he sees his job as that of deriving secondary facts from his primary facts. At the moment he is busy determining the relative frequencies of the eleven parts of speech as the first word of a sentence versus the second word of a sentence. (There isn't anybody exactly like this, but there are some approximations.)

The lesson to be learnt from this satirical description is that a corpus should always be the answer to a linguistic query, that is a means to an end and not an end in itself. So what are potential research questions that a corpus of early New Zealand texts might enable us to answer? One area of research might be differential language change, i.e. the development of NZE vis à vis other national varieties of English such as British (BrE) and American English (AmE).[1] In this case, our corpus of early New Zealand texts would have to be compiled in a way that would make a comparison with existing historical corpora possible. Obviously, NZE grammar is not categorically distinct from BrE or AmE. What makes NZE grammar distinct is mostly a question of preference for certain grammatical options available in global English. At this level, NZE grammar may actually be rather 'exotic'. One example of the currently exotic state of NZE vis à vis varieties such as Australian English (AusE) or AmE is in the use of the progressive form, e.g. *John **is texting** a message to his girlfriend with his new mobile*, which is used much more frequently in NZE than in other native varieties of English (see Collins 2009 and Hundt and Vogel 2011). In this paper, I will therefore present a case study on the use of the progressive in what constitutes the nucleus of a corpus of early New Zealand texts.

In part two of this paper, I will briefly comment on previous research on the diachronic and regional developments of the progressive in English. In section three, I will focus on the steps involved in moving from a text database to a corpus, as well as the challenges and limitations that such a project involves. Part four will present results from a study on progressive constructions in early NZE as well as historical BrE and AmE texts.

2. The progressive – historical and regional developments

The origins of this grammatical construction can be traced back to Old English times, but even in Shakespeare's writing it had not become obligatory (see Polonius' question *What do you read my lord?* (Act II, Scene ii) which is not a question about Hamlet's reading habits). It is only during the nineteenth century that the progressive becomes more frequently used (Strang 1982, Smitterberg 2005, Kranich 2008). The progressive is still spreading in the twentieth century (Mair and Hundt 1995, Smith 2005, Leech et al. 2009), but there is relatively little regional difference between BrE and AmE (Leech et al. 2009: 122). NZE turns out to be quite exotic because New Zealanders use

the progressive much more frequently than people in the UK or the US. Hundt (1998: 75) provides empirical evidence of a regional difference between northern and southern hemisphere varieties; more recently, Collins (2009) has used a subset of the ICE corpora to show that usage in NZE is actually significantly different from AmE and BrE but also from AusE (see Table 1).

Table 1: Progressives across four Englishes (approx. 120,000 words per variety; from Collins 2009: 116)[2]

	NZ	AUS	US	GB
speech	57.7%	71.8%	76%	69.5%
writing	42.3%	28.2%	24%	30.5%
N	894	753	626	660

A particularly interesting finding is that New Zealanders use the progressive – a construction that is typical of colloquial, spoken English – much more frequently in written language than Americans, Britons or Australians. Our study on progressives in student writing shows that New Zealanders actually use the progressive with a similar frequency to some people who have learnt English as a second (ESL) or foreign (EFL) language (see Figure 1).[3]
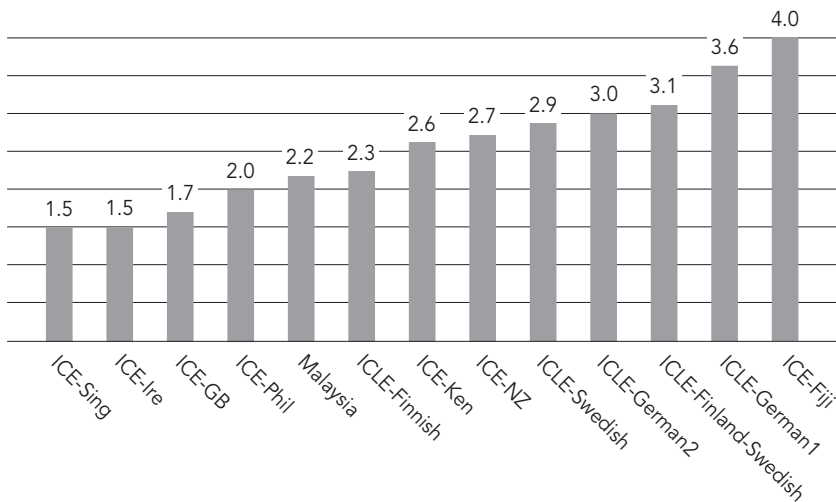

Figure 1: Normalized frequencies (per 1,000 words) of progressives across ENL, ESL and EFL corpora (student writing) – from Hundt and Vogel (2011: 154)

A look at Figure 1 shows that students in New Zealand use the progressive almost as frequently as Swedish-speaking students in Finland and more frequently than students in Kenya. This regional difference is also confirmed for printed academic texts (see Hundt and Vogel, 2011: 155). Incidentally, students from Ireland use the progressive with a lower frequency than that attested in essays collected for the British component of the ICE corpus.

It is important to note in this context that the similarities between NZE, on the one hand, and ESL and EFL usage, on the other hand, mainly concern the frequency with which the progressive is used. Less obvious are structural similarities in the use of the construction, such as the combination of a progressive with stative verbs like *be* or *love*; these are considered typical of non-native speaker usage, but are also occasionally found in BrE or AmE, for instance.[4] The following examples illustrate instances in second-language varieties of English in Kenya, Singapore, Fiji or the Philippines where writers of BrE or AmE are more like to use a simple form (emphasis added throughout):

1. This essay *will be discussing* six factors why women have to work for empowerment. (ICE-Fiji w1a-015)

2. It spread due to movement of laborers. It *is being used* now in Zambia as a language of education. (ICE-Ken w1a-003)

3. Whereas in the 2nd article, it says that the economy *is fast rising* ever since the Ramos Administration started. (ICE-Phil w1a-011)

4. However, according to Hume, there is not guarantee that just because nature *has been uniformly functioning* in the past, it will continue to do so always. (ICE-Sing w1a-014)

Example (3) is particularly interesting because the present progressive is used in a context in which we would expect to see the present perfect or a present perfect progressive in BrE. Interestingly, it bears a striking resemblance to a chance finding from NZE: While holidaying in New Zealand, I came across a notice in the shared bathroom facilities on a camping ground which asked parents to accompany young children to showers and toilets. The reason given was "*We are experiencing too many accidents of late.*" From the perspective of a speaker of BrE, this sentence is unusual because it combines a present progressive with an adverbial that normally combines with the present perfect (most likely a simple present perfect).[5] However, there was no obvious non-

native influence on this sign and, furthermore, the native speakers of NZE that were asked to comment on the sentence did not find this usage unusual. It is possible that some native speakers of English may extend the progressive to contexts of perfective marking because the past progressive is also used in a similar way, namely as a marker of recent past: "*Tom, you were just telling me that in all you had nine students going down there*" (COCA:CNN_Morning, 1997; quoted from Bergs and Pfaff 2009). Moreover, Fraser Gupta (2006: 104f.) found that the progressive is occasionally used by inner-circle speakers (mainly in the US) following expressions such as *This is the first time I ... .* In other words, the extended use of the progressive in New Zealand English (both in terms of its frequency and some of its functions) fits with Gachelin's (1997: 43) claim that the extended use of the progressive in New Englishes may eventually lead to long-term change in the English language as a whole: "Its generalization […] may herald what will be World English usage in the next century."

The grammaticalisation and spread of the progressive construction from Old English onwards are well documented (see, e.g., Denison 1993: 371–410). The historical details are not relevant to our discussion here. There is one aspect, however, that is worthwhile mentioning, namely the possible influence of language contact with speakers of Gaelic (see e.g. Keller 1925 or Filppula and Klemola 2012). Gaelic has a periphrastic construction combining the verb *be*, a preposition and a verbal noun that is used to refer to ongoing events, including the possibility of combining with some stative verbs like *living* (Ronan, 2001: 50). In other words, possible influence from Gaelic would open up a wider functional range of the progressive construction in a contact variety of English such as IrE. Language contact and/or dialect contact might thus also have played a role in the spread of the progressive construction in NZE. There is language contact with non-native speakers of English in the colonial and post-colonial context. But it is also possible that contact between speakers of different regional varieties of English might be the reason why the progressive is used so frequently in present-day NZE. The most likely sources of regional dialectal influence in the development of NZE would be Scottish and Irish English (see Bauer 1994, 1997 or Gordon et al. 2004).[6] McCafferty and Moreno (2010, ms.) have investigated, among other things, the use of the progressive in a diachronic corpus of IrE letters that provides valuable comparative data.

A diachronic corpus of early New Zealand writing would allow us to verify whether the frequent use of the present progressive in current NZE is the result

of language contact with (a) non-native speakers of English or (b) speakers of other regional dialects of English like IrE. We would also be able to show whether the progressive was used more frequently in New Zealand than in BrE from the early colonial days or whether the prolific use of the construction in NZE is a more recent development. The two aspects are connected in so far as an early (colonial or immediately post-colonial) dating of the phenomenon would speak for influence from regional dialects whereas recent spread is more likely to have been supported by contact with speakers of English as a second language.

The question is what a diachronic corpus should look like that might allow us to test these hypotheses. Holmes (1994: 27) described the ideal scenario for a study of recent change in New Zealand English. The same requirements also apply to the earlier periods of NZE in comparison with its 'parent' variety, BrE, or other relevant corpora of English as a first language:

> The ideal situation […] would appear to be to use two corpora constructed on parallel principles at […] different points in time. Assuming that any variation identified can be reasonably attributed to language change over time, rather than to, say, topic differences or stylistic differences between the corpora […]. Unfortunately, no such parallel corpora exist for New Zealand English.

In the following, I will discuss how we might build such a parallel corpus from existing digitized texts.

3. From electronic text collection to corpus

3.1 Existing diachronic corpora

As pointed out in the introduction, a corpus is not simply a collection of texts but one that has been based on sampling principles. For a corpus of early written New Zealand English, sampling with criteria that will make the corpus comparable with existing diachronic corpora of reference varieties like British, American and Australian English is advisable because this will minimise 'cost'. Suitable diachronic corpora that sample these varieties are COOEE (a *Corpus of OZ Early English*) and ARCHER (*A Representative Corpus of Historical English Registers*).[7]

COOEE contains texts from the years 1788–1900, including both speech-like texts and private letters alongside more formal text types such as official announcements by the government. However, the corpus is not publicly

available and can therefore not be used for comparative research. But it provides important methodological input for the compilation of a corpus of early New Zealand writing, as we will see.

ARCHER is a diachronic corpus of British and American texts from the middle of the seventeenth to the end of the twentieth century. The corpus is divided into sub-corpora of fifty-year periods. It provides comparative data on the two varieties for a number of speech-like and written registers (drama, fiction, medical, scientific and legal writing, newspapers, journals and diaries, private letters, sermons). Individual samples consist of approximately 2,000 words (sometimes comprising more than one text, for instance in the category 'newspapers'). The target for each text category, register and sub-period is a total of 10 samples (i.e. approximately 20,000 words). ARCHER is currently the best available corpus for comparative studies on differential change in varieties of English in the late Modern period. It thus provides a very suitable sampling frame for a corpus of early New Zealand writing.

3.2 A corpus of Early New Zealand writing: challenges

Even with a suitable, ready-made sampling frame, the compilation of a corpus of early New Zealand texts is far from straightforward. The main challenges are to (a) determine the criteria for including a text in the corpus and (b) to cope with the available spread of text types, and (c) to evaluate whether the diachronic cuts that the ARCHER sampling frame provides also provide helpful sub-samples for a diachronic corpus of early New Zealand writing. These questions will be addressed in the following sections.

3.2.1 Who qualifies as a New Zealander?

Corpus compilation in the colonial and early post-colonial context faces the problem that it is not easy to determine when an immigrant becomes a New Zealander and thus eligible to be considered as an author whose texts should be included in the corpus. Even the compilers of the spoken corpus of New Zealand English collected in the 1990s asked themselves this question:

> Who should be allowed to contribute to the corpus? [...] It is a particularly vexatious problem for colonial societies where large sections of the community are immigrants. At what point does an immigrant become a New Zealander? (Holmes, Vine and Johnson 1998: 23)

Bauer (1991, unpaginated) speculates on the (socio-)linguistic processes that may have affected the language of immigrants, and that grammar, in

particular, is likely to have remained relatively stable even after a lengthy stay in the colony:

> Britons (or Australians or Americans ...) arriving in New Zealand may consciously or unconsciously adapt their speech to use particular vocabulary items, but they are unlikely to be even subconsciously aware of the statistical trends in the usage of particular grammatical patterns. We must therefore predict that they are unlikely to make appropriate changes to these aspects of their speech, even after lengthy residence. Now, it might be that this supposition is false, and that they do adapt appropriately after sufficient length of time.

This is a speculation and the only way of settling the matter would be longitudinal data on the development of the grammar of individual immigrants, i.e. the kind of evidence that we are unlikely ever to be able to collect for previous periods. Conscious or unconscious adoption of grammatical features is not that unlikely to occur, though.[8] Rissanen (1984: 418f.) argues that the language of people who migrated to America even after having received their education in Britain is a good source for the study of Early American English: "The people producing these texts [from the 1640s, M.H.] had spent their youth and acquired their education in England, but they had lived in America for a number of years […]." Likewise, anyone who migrated to New Zealand as a child or young adult would be a good informant for an emerging variety of NZE.

In addition to migration to New Zealand, New Zealand-born authors might also leave the country and spend time in another English-speaking country and thus adopt (grammatical) features from a different regional variety of English. The criteria that were applied in the compilation of the spoken and written corpora of New Zealand English in the 1990s considered both possibilities (immigration to and temporal emigration from New Zealand): only speakers who had been resident in New Zealand at age 10, who had spent less than ten years outside of the country, and who had returned at least a year before the text to be included in the corpus was produced. For several reasons, such strict selection criteria are difficult if not impossible to apply in the collection of a corpus of early written New Zealand texts. Why this is the case can be illustrated with the biographies of some authors that were included in the text database compiled by the *New Zealand Electronic Text Centre.* I will briefly summarize some biographical facts and then comment on their relevance to corpus compilation.

- Hon. James Coutts Crawford (1817–1889) was born in Scotland, arrived in New Zealand (via Australia) in 1939 at the age of 22 but returned to England twice between 1841 and 1857. He died 1889 in London.

- Walter Buller (1838–1906) was born in Hokianga as the son of a missionary and is thus, by birth, a true New Zealander; but he travelled to Europe, too, in 1870. He gave a paper to the *Philosophical Society* before his journey, though, which makes this particular text a clear candidate for inclusion in the corpus.

- Edwin Fairburn (1827–1911), New Zealand-born and the son of early immigrants, like Crawford and Buller, also travelled to Europe. But even though we know that he went to Germany and Austria we do not have information on how long he stayed there.

- Richard Treacy Henry (1845–1929) was born in Ireland. Aged 6, he migrated to Australia in 1852 with his parents and thus spent his formative years in the southern hemisphere (if not in New Zealand itself). In 1874, aged 29, he moved from Australia to New Zealand. His biography is typical of some migrants in so far as they did not necessarily arrive directly from the British Isles but sometimes via Australia (see Gordon et al., 2004: 44f.), one fact that has been taken to explain the close historical connection between the two varieties.

For the early colonial period, biographies such as those of Buller and Fairburn are quite rare since most migrants arrived in New Zealand as young adults. Most of the authors included in the text database of the *New Zealand Electronic Text Centre* did not spend their lives exclusively in the colony, even after they had arrived there. This is a complication that Fritz (2007: 65f.) also faced in the compilation of COOEE; he concluded that

> AusE developed […] from the dialects and sociolects the immigrants spoke and wrote. Therefore all English texts in early Australia are valid sources. None is inherently better than the other.

But there are also clear criteria for excluding certain authors. Samuel Butler (1835–1902), the author of *Erewhon*, is one of them. He was born in England and only spent five years of his life in New Zealand (between 1859 and 1864), working on a sheep farm near Christchurch; he published a couple of articles in a local newspaper, among them one entitled 'Darwin Among the Machines'

(1863). *Erewhon*, however, was only published on his return to England. It is probably the descriptions of landscapes in the novel that are so obviously related to his stay in New Zealand that lead to the novel being included in the database of the *New Zealand Electronic Text Centre* (but in a later edition, namely from 1927). The fact that Butler spent five years of his life in New Zealand is not enough to qualify him as an author of emerging New Zealand English in the colonial period.

Another potentially problematic case is Katherine Mansfield (1888–1923). She was born in New Zealand, spent her formative years in the country and also received most of her schooling in Wellington. Between 1902 and 1906 she attended school in London and returned to New Zealand for a short while afterwards; she died of tuberculosis in France aged only 35. The texts that were included in the corpus of early New Zealand writing were written in Europe, but the fact that she was born in New Zealand and spent her childhood and early youth there make her a New Zealand author. Fritz would have excluded her as an eligible source as he only included texts that were produced in Australia, New Zealand or on Norfolk Island in COOEE (2007: 66).

There are a few additional complications that do not allow us to apply the same strict criteria in the collection of corpora of early colonial and post-colonial writing as we would apply in the compilation of a corpus of current English. First of all, the names of authors for individual texts are not necessarily known (e.g. in the case of newspaper articles that are published without the author's name). But even if the name of the author is known we do not necessarily have any biographical background information. A lot of potential contributors to a corpus of early NZE were simply not well-known or important enough to be included in biographical sources. If we were to include only those authors where biographical information is available this might even skew the data included in the corpus by giving preference to well-known informants who are likely to be of a relatively high social background.

Second, because a lot of the material comes from published sources, we can never rule out some editorial influence and thus the editor's linguistic background as an additional layer in the text. This is an aspect that Bauer (1991, unpaginated) has also pointed out as a potential source of non-authentic language use even for corpora of written post-colonial NZE: "[...] it is impossible to avoid speakers who are not technically speakers of New Zealand English [...] the problem is likely to be greatest in the print media [...]." In a paper that investigates personal letters from an edited collection (Hundt, forthcoming) I was able to demonstrate that such editorial influences

are more likely to affect aspects of orthography but largely seem to leave morpho-syntactic variables unaffected.

3.2.2 Availability of text types

The readily available digital early New Zealand texts obviously do not perfectly match the sampling frame of the ARCHER corpus. And even when there are texts for a particular text category, there is not necessarily enough material to fill a sample (ten times 2,000 words) that would match the ARCHER framework. At other times, the available material allows for sampling at 30-year intervals within or across two ARCHER sub-periods. An overview of the number of words in a first version of the CENZE corpus is given in Table 2.

Table 2: Availability of early New Zealand texts according to the ARCHER sampling frame

	1800–49	1850–99	1900–49	1950–99
Drama	—	—	—	—
Fiction	—	✓	✓	—
Medical	—	—	—	—
Scientific	—	✓	✓	✓
Legal	—	—	—	—
Newspapers	✓	✓	✓	✓
journals & diaries	—	—	—	—
private letters	✓	✓	✓	—
Sermons	?	?	?	?

New Zealand only became a crown colony in 1840, and it is therefore not surprising that, with the exception of letters from emigrants and newspapers, no material is available for the first half of the nineteenth century. The large gaps in the second half of the twentieth century is due to copyright restrictions: the *New Zealand Electronic Text Centre* mostly digitized texts that are not subject to copyright restrictions.[9]

The sampling for private letters beyond those by early settlers in the 1840s is somewhat problematic, too, since the letter collections included in the material of the *New Zealand Electronic Text Centre* are from two authors only (one for the 1860s, the other for the 1920s). As pointed out previously,

some of the earliest texts are letters by emigrants to New Zealand that were published soon after New Zealand had become a crown colony to advertise the new colony to potential settlers in GB; this material was digitized by the University of Auckland and is publicly available on the internet (for a more detailed discussion of these data, see Hundt forthcoming).[10]

The text category 'sermons' (religious writing) turned out to be problematic for a different reason. ARCHER samples mostly sermons (i.e. persuasive texts) for this register. The material digitized by the *New Zealand Electronic Text Centre* consists of texts that describe the mythology of the Maori, exclusively, and are therefore not suitable as a parallel source of texts for the CENZE.

Historical newspapers were not digitized by the *New Zealand Electronic Text Centre* but by the national Library of New Zealand (for the years 1839–1920).[11] Over eight million individual articles can be downloaded from the library's webpage. However, the texts were OCRed[12] but not manually post-edited. This means that for each article to be included in the corpus, the facsimile of the original print version has to be consulted and the texts have to be corrected manually before they can be included in the corpus. Finally, narrative prose texts had to be supplemented by additional material beyond that available from the *New Zealand Electronic Text Centre.*

3.2.3 Which sub-periods and how many?

In terms of diachrony, the ARCHER sampling frame uses 50-year periods. Individual samples are spread more or less evenly[13] across this time span resulting in a continuous coverage of the material. However, it might be easier to demonstrate diachronic change if the sampling keeps to more narrowly defined sampling points. In previous research on recent grammatical change, sampling points at approximately 30-year intervals have proven useful as this roughly corresponds to the distance between two generations of speakers (see Leech et al., 2009 and Hundt and Leech, 2012). For registers that are well attested early on, it might even make sense to sample at 20-year intervals. One problem, as we will see, is that not enough material is available to fill the text categories of the ARCHER sampling frame in whatever chronological grid is adopted: with nine registers and ten samples of about 2,000 words each, every diachronic sample would require 180.000 words worth of text – a somewhat ambitious goal.

3.3 A first Corpus of Early New Zealand English (CENZE)

With all the limitations discussed in this section, what will a corpus of early New Zealand texts look like? Table 3 shows that we are still far from the goal of 180.000 words per diachronic sample.

Table 3: Number of words in CENZE corpus according to the ARCHER sampling frame (registers rather than diachronic cuts)

	1800–49	1850–99		1900–49		1950–99	TOTAL
drama	—	—		—		—	—
fiction	—	20.969		20.855		—	41.824
medical	—	—		—		—	—
scientific writing	—	*1870s* 20.266		*1900s* 14.390	*1930s* 20.776	*1960s* 20.429	75.861
legal texts	—	—		—		—	—
newspapers	*1840s* 20.180	*1860s* 20.437	*1880s* 20.372	*1920s* 21.215	*1940s* 20.401	—	102.606
journals & diaries	—	—		—		—	—
private letters	*1840s* 20.364	*1860s* 20.790		*1920s* 20.709		—	61.863
sermons	?	?		?		?	—
Total	40.544	102.835		118.346		20.429	282.154

The registers with the best diachronic coverage are personal letters and newspaper texts. Newspaper texts are available from 1839 onwards – initially from the *New Zealand Gazette and Wellington Spectator*, only. For the very early colonial years, we will probably not be able to move beyond what is currently available: the settlers had other immediate concerns than to write novels or scientific treatises soon after their arrival in the new country. And even though they were likely to have gone to church, archiving early sermons in those days was not a priority, either. Later gaps in the coverage of the ARCHER registers (e.g. fiction from the first half of the twentieth century) are more likely to be filled. In some cases, the representativeness of the texts is not particularly good (see the problems discussed in relation to private letters in the 1860s and 1920s discussed in section 3.2.2). All in all, the first CENZE is a bit of a patchwork affair. Nonetheless, it can fruitfully be used to monitor

the development of some grammatical patterns. In this paper, I will use the progressive as a case study.[14]

4. Case study: The progressive in CENZE

In order to allow for comparability with previous ARCHER-based studies (Hundt 2004a, 2004b), the same criteria were applied for the definition of the linguistic variable. Using WordSmith Tools, I searched for combinations of the auxiliary *be* with a present participle (allowing for material to occur between auxiliary and participle). In a second step, all non-progressives were manually removed from the concordances, including instances where the participle has adjectival rather than verbal function (e.g. *This news is shocking* or *His countenance was repulsive and forbidding*) and examples with participles that function as an apposition rather than as part of the verb phrase (e.g. *He was at home, repairing the roof*) (see Hundt 2004a: 56). Similarly, patterns where *be* was a copula followed by a gerund were excluded manually from the concordance (e.g. *Consequently what is called keeping the length of arc constant is really allowing it to become slightly longer than the desired length*, […] ARCHER 1925angu.s7b). Instances with two participles (e.g. *A deadly bark beetle is attacking and killing many hickories*, ARCHER 1932FeltS7a) were only counted once. As in Hundt (2004a), instances of *going to* as a future time expression were excluded from the datasets.

Table 4 gives the results, both in terms of absolute frequencies and normalized (per 10,000 words). Normalization is necessary to enable comparison across the differently sized sub-corpora and to facilitate comparison with previous research. The data from the CENZE corpus have been supplemented with searches in the written part of the Wellington Corpus of NZE for the second half of the twentieth century to obtain data for the last sub-period sampled in ARCHER.

Not surprisingly, the progressive occurs with different normalized frequencies in different registers. It is most frequent in private letters (a text type that was found to have relatively high frequency of other colloquial patterns in previous studies, see e.g. Smitterberg, 2005: 77f.). As far as diachronic developments are concerned, however, the letters data might not be a reliable indicator because the material from the 1860s and 1920s are not representative samples (one author only in each of the sub-periods).

Influence from IrE in the letters is unlikely if we compare the results

Table 4: Progressives in CENZE (normalized frequencies per 10,000 words in brackets; figures in square brackets give normalized frequencies from the Wellington corpus)

	1800–49	1850–99		1900–49		1950–99
fiction	—	54 (25.8)		82 (29.3)		[53]
scientific writing	—	*1870s* 8 (3.9)		*1900s* 23 (16.0)	*1930s* 4 (1.9)	*1960s* 8 (3.9)
newspapers	1840s 18 (8.9)	1860s 35 (17.1)	1880s 43 (21.1)	1920s 57 (26.9)	1940s 41 (20.1)	[43.2]
private letters	1840s 88 (43.2)	1860s 45 (21.6)		1920s 106 (51.2)		—
Total	106 (26.1)	185 (18.0)		313 (26.4)		—

from CENZE with those from McCafferty and Moreno (2010, ms.). In the 1830s letters in *CORIECOR* (Corpus of Irish English Correspondence), the progressive occurs with a frequency of only 41.8 occurrences per 10,000 words. Moreover, MacCafferty and Moreno include instances of *be going to* (Kevin McCafferty, p.c.) which were excluded from my counts. In other words, the progressive is used more frequently in the early New Zealand letters in the 1840s than in a contemporaneous collection of IrE letters. It is also quite frequent in newspapers and fictional writing. The register with the lowest occurrence of progressives is the most formal one represented in CENZE, namely scientific writing.

The two data points available from fictional writing do give evidence of an increase of progressives across time. In the fiction sample from ICE-NZ that Collins (2009: 116) analysed, progressives are used with a frequency of 122 per 10,000 words and thus significantly more often than in the first half of the twentieth century. More reliably, the newspaper evidence shows that progressives become more frequent in New Zealand English between the early colonial days and the first half of the twentieth century, even though there is a decrease between the 1920s and 1940s.

How does the development of the progressive in NZE compare with its spread in BrE and AmE? We will look at the two text types with the best diachronic coverage in CENZE, science and newspaper reportage.
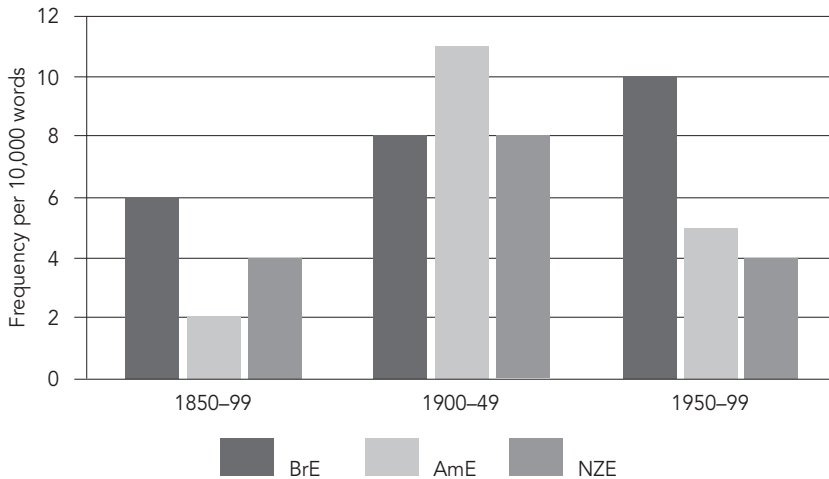
Figure 2: Progressives in the science sub-corpus – ARCHER vs. CENZE

In scientific writing, the text frequency of progressives is rather low, so the figures have to be interpreted rather cautiously. Nevertheless, we see an increase over time in the BrE part of ARCHER on the one hand, whereas in AmE and NZE, the peak in the first half of the twentieth century is followed by a decline. With the overall text frequency of progressives in scientific writing being so low, there is an obvious risk of individual samples having a skewing effect on the results. This seems to be the case for the New Zealand texts from the beginning of the twentieth century: most progressives are found in only two out of the ten samples. One of these samples is from a text written by the Irishman Richard Treacy Henry, whose parents had migrated to New Zealand via Australia (see 3.2.1). The author of the other text with a higher-than-average frequency of progressives is an Englishman who arrived in New Zealand aged 26. In other words, only one of the two authors has an Irish background, thus making language contact with a variety of IrE as the sole reason for a frequent use of the progressive rather unlikely. The significantly higher frequency of progressives in New Zealand academic writing that Collins (2009: 116) and Hundt and Vogel (2011: 154f.) observe must thus be a recent development. Further evidence from this comes from a comparison of the ARCHER and CENZE data with evidence from the ICE corpora: ARCHER samples scientific writing from 1975 (BrE) and 1954–1997 (AmE), CENZE from the 1960s; the ICE corpora, on the other hand, are comprised

of material that was collected from the 1990s onwards. This diachronic bias does not seem to play a role for BrE, with the (natural) science sub-samples in ARCHER and ICE-GB yielding comparable normalized frequencies of progressives at 10 and 7 occurrences per 10,000 words, respectively. The difference between the CENZE and ICE-NZ data, on the other hand, shows that the progressive has increased significantly in New Zealand academic writing towards the end of the twentieth century: the science texts in CENZE yield 4 progressives per 10,000 words, whereas those in ICE-NZ yield 31 progressives per 10,000 words.

Let us now turn to the diachronic development in a text type where progressives are used more frequently: newspaper texts. Figure 3 plots the diachronic developments in ARCHER and CENZE. Even though the sub-periods in CENZE are different from those in ARCHER, the overall diachronic trend becomes clear: Early New Zealand newspapers have a comparable relative frequency of progressives as we find in the newspaper texts included in ARCHER.
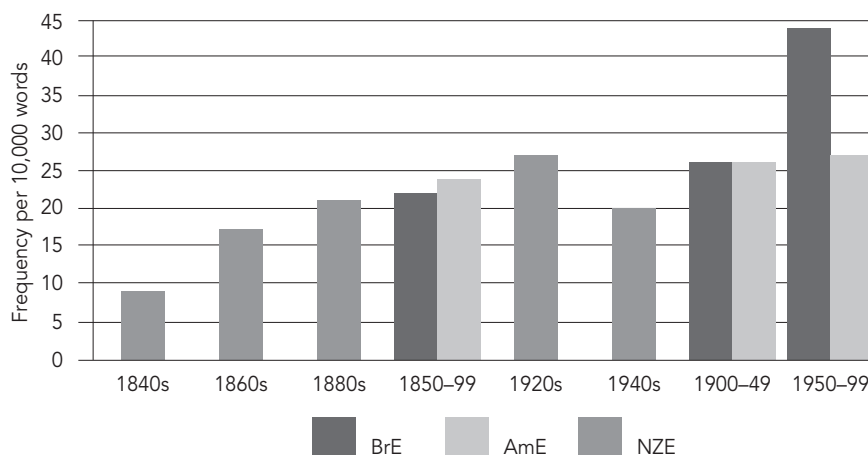

Figure 3: Progressives in newspaper writing – ARCHER and CENZE

Figure 4 compares late twentieth-century newspaper data from ARCHER with evidence from the corresponding sections in the Brown family of corpora and the Wellington Corpus of Written New Zealand English, which was compiled from texts published in the late 1980s (see Hundt, 1998: 75f.). These results, together with those from CENZE, again suggest that the frequent use of the

progressive in NZE is more likely to be due to recent change rather than an earlier predilection of New Zealanders to use the progressive.
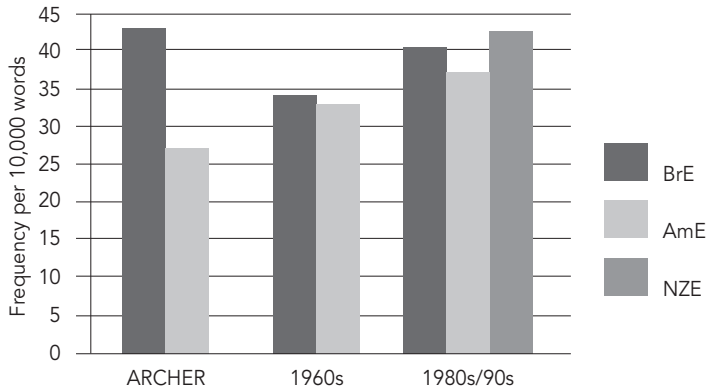

Figure 4: Progressives in 20th-century newspaper writing: ARCHER and the Brown family

Figure 4 also shows that we have to be cautious when we compare results from different corpora. ARCHER samples only reportage whereas the Brown family includes reportage, editorials and reviews. A sub-sample from FLOB comparable in size and composition to the ARCHER texts (national newspapers rather than provincial; a cross-section of different kinds of news) yields 42.2 progressives per 10,000 words and thus a slightly higher normalized frequency than the overall sample (40.4 progressives/10,000 words), which includes editorials and reviews. The topic also appears to play a role, with society news containing more progressives than political news or sports reportage. Furthermore, a sub-sample of 1990s British provincial newspaper reportage yields a much higher normalized frequency of progressives at 57.6 per 10,000 words. Thus, the composition of samples is particularly important for small diachronic corpora that comprise only about 20,000 words per register and period. The press section of the Brown-type corpora with a total of 88 samples and approximately 176,000 words may thus produce somewhat more robust results than the newspaper texts in ARCHER.

To sum up, the more frequent use of the progressive in current NZE is a recent development that is likely to date to the second half of the twentieth century. Dialect contact with IrE during the early days of the development of NZE is an unlikely source of the more frequent use of the progressive in New

Zealand today, both in terms of the diachronic developments as well as the evidence from individual authors in this small-scale study.

5. Conclusion and outlook

Despite the availability of digitized texts from the early colonial period and later stages in the history of New Zealand, compiling a corpus of early New Zealand writing is not as easy and straightforward a task as one would hope. The question is whether our brief visit to the *Erewhon* of historical corpus linguistics in New Zealand has discouraged us to the extent that we simply want to board that hot air balloon and leave. Contrary to the fears of the inhabitants of Butler's *Erewhon*, technical evolution has not lead to the development of machines that think and act for themselves. The Cyborgs of computational linguistics are still not even a remote possibility on our horizon. There is still a lot of manual labour involved in the compilation of historical corpora. The germ of a historical corpus of early New Zealand texts described in this article could be developed into a more representative corpus with additional data. For the category of letters, this would probably mean the inclusion of handwritten documents that are hopefully to be found in some archives. But even though the texts included in my embryonic corpus of early New Zealand writing do not yet amount to a representative sample of the emerging written variety in colonial and post-colonial New Zealand, what is available so far can be used to test hypotheses on relatively frequent grammatical patterns, such as the progressive. The case study has also shown that results from relatively small sub-samples have to be treated with particular caution, especially if findings from different corpora are compared. The evidence from CENZE suggests that the progressive was not used significantly more frequently in early New Zealand writing than in comparable texts from Britain and the US. Instead, data from the Brown-family of corpora and components of the International Corpus of English indicate that New Zealanders seem to have moved ahead of other ENL speakers and writers in the use of the progressive quite recently.

Notes

1 See Hundt (2009a) on differential change in BrE and AmE.

2 The differences are prove significant at p ≤ 0.001 in a chi-square test. Note that Hundt (1998: 75), using newspaper data from the Brown family of corpora, only, did not find a significant difference between Australian and New Zealand English. Both are ahead of British and American English in the growing use of the progressive in that study, indicating that text type is an important factor to consider in the study of progressives.

3 Note that not all ESL varieties use the progressive more frequently than it occurs in BrE: SingE, for instance, has an even lower incidence of progressives.

4 For its use in an advertising campaign, based on a Timberlake song, see http://en.wikipedia.org/wiki/McDonald's_advertising (accessed 15th February 2010).

5 For similar use of the progressive in IrE, see McCafferty and Moreno (2010, unpaginated).

6 Note that Gordon et al. (2004) investigate the development of the New Zealand accent rather than developments in the grammar of the variety.

7 For COOEE, see Fritz (2007). Background information on ARCHER can be found in Biber, Finegan and Atkinson (1994). For information on the different versions of ARCHER, see Yáñez Bouza (2011). The comparative data used in this paper come from material to be included in the forthcoming version of the corpus (ARCHER 3.2), which provides broader coverage of AmE than previous versions of the corpus did. Information on COOEE and ARCHER is also available from http://www.helsinki.fi/varieng/CoRD.

8 For second dialect acquisition (accent), see Tagliamonte and Molfenter (2007). Some principles described in this article might also apply to the acquisition of grammatical preferences in a new or evolving dialect.

9 Narrative prose from the second half of the twentieth century was not digitized by the *New Zealand Electronic Text Centre* for reasons of copyright. This gap in the corpus could be filled relatively easily because these texts are available either in print or as samples in existing corpora, such as the Wellington Corpus of Written New Zealand English or the New Zealand component of the ICE corpus.

10 The letters were digitized by the Early New Zealand Books project at the University of Auckland, New Zealand. They can be found at http://www.enzb.auckland.ac.nz/document//1843_-_Letters_from_Settlers_and_Labouring_Emigrants (last accessed 17.01.2011).

11 http://paperspast.natlib.govt.nz/cgi-bin/paperspast

12 OCR stands for 'optical character recognition' and thus is shorthand for 'automatic digitization of text'.

13 Occasionally, sampling for an individual register diverges from this sampling principle: scientific British texts in the twentieth century, for example, stem from the years 1925 and 1975 (for the two sub-periods) only.

14 Hundt and Szmrecsanyi (2012) use the same corpus to investigate animacy as a determinant of grammatical variation in NZE vis à vis BrE and AmE. In

Hundt (forthcoming), I focus on a broader range of potentially non-standard constructions in early New Zealand letters (the focus there is on the 1840s material, only). Hundt (in preparation), finally, investigates the use of relativizers in restrictive vs. non-restrictive relative clauses in the science part of the corpus.

References

Bauer, Laurie. 1991. 'Who speaks New Zealand English?' *ICE Newsletter* 11. (unpaginated)

Bauer, Laurie. 1994. 'English in New Zealand.' In Robert W. Burchfield (ed). *The Cambridge History of the English Language. Volume V: English in Britain and Overseas. Origins and Developments*. Cambridge: Cambridge University Press. 382–429.

Bauer, Laurie. 1997. 'Attempting to trace Scottish English influence on New Zealand English.' In Edgar W. Schneider (ed). *Englishes Around the World: Studies in Honour of Manfred Görlach. Volume 2. Caribbean, Africa, Australasia*. Amsterdam and Philadelphia: Benjamins. 257–272.

Bergs, Alexander and Maike Pfaff. 2009. 'I was just reading this article. Is the perfect of the recent past on its way out?' Paper presented at the SEU Symposium *Current Change in the English Verb Phrase*, 14 July 2009.

Biber, Douglas, Edward Finegan and Dwight Atkinson. 1994. 'ARCHER and its challenges: Compiling and exploring a representative corpus of historical English registers.' In Udo Fries, Gunnel Tottie and Peter Schneider (eds). *Creating and Using English Language Corpora: Papers from the Fourteenth International Conference on English Language Research and Computerized Corpora. Zürich 1993*. Amsterdam: Rodopi. 1–13.

Biber, Douglas, Susand Conrad and Randi Reppen. 1998. *Corpus Linguistics. Investigating Language Structure and Use*. Cambridge: Cambridge University Press.

Collins, Peter. 2009. 'The progressive in English.' In Pam Peters, Peter Collins and Adam Smith (eds). *Comparative Studies in Australian and New Zealand English Grammar*. Amsterdam: Benjamins. 115–123.

Fillmore, Charles J. 1992. '"Corpus linguistics" or "computer-aided armchair linguistics"?' In Jan Svartvik (ed). *Directions in Corpus Linguistics. Proceedings of the Nobel Symposium 82, Stockholm 4–8 August 1991*. Berlin and New York: Mouton de Gruyter. 35–60.

Filppula, Markku and Juhani Klemola. 2012. 'Celtic and Celtic Englishes.' Bergs and Laurel J. Brinton (eds). *English Historical Linguistics. An International Handbook. (HSK 34-2)*. Berlin: Mouton de Gruyter. 1687–1703.

Fritz, Clemens W.A. 2007. *From English in Australia to Australian English. 1788–1900*. Frankfurt: Peter Lang.

Gachelin, J.M. 1997. 'The progressive and habitual aspects in non-standard Englishes.' In Edgar W. Schneider (ed). *Englishes Around the World 1. General Studies, British Isles, North America. Studies in Honor of Manfred Görlach*. Amsterdam and Philadelphia: Benjamins. 33–46.

Gordon, Elizabeth, Lyle Campbell, Jennifer Hay, Margaret Maclagan, Andrea Sudbury and Peter Trudgill. 2004. *New Zealand English. Its Origins and Evolution*. Cambridge: Cambridge University Press.

Holmes, Janet, Bernadette Vine and Gary Johnson. 1998. *Guide to the Wellington Corpus of Spoken New Zealand English*. Wellington: Victoria University.

Holmes, Janet. 1994. 'Inferring language change from computer corpora: Some methodological problems.' *ICAME Journal* 18: 27–40.

Hundt, Marianne. 1998. *New Zealand English Grammar. Fact or Fiction?* Amsterdam and Philadelphia: John Benjamins.

Hundt, Marianne. 2004a. 'The passival and the progressive passive – a case study of layering in the English aspect and voice systems.' In Hans Lindquist and Christian Mair (eds). *Corpus Approaches to Grammaticalisation in English*. Amsterdam and Philadelphia: Benjamins. 79–120.

Hundt, Marianne. 2004b. 'Animacy, agency and the spread of the progressive in eighteenth- and nineteenth-century English.' *English Language and Linguistics* 8(1): 47–69.

Hundt, Marianne. 2009a. 'Colonial lag, colonial innovation, or simply language change?' In Günter Rohdenburg and Julia Schlüter (eds). *One Language, Two Grammars: Morphosyntactic Differences between British and American English*. Cambridge: CUP. 13–37.

Hundt, Marianne. 2009b. 'Global feature – local norms? A case study on the progressive passive.' In Lucia Siebers and Thomas Hoffmann (eds). *World Englishes – Problems, Properties and Prospects*. Amsterdam and Philadelphia: Benjamins. 287–308.

Hundt, Marianne. 2012. 'Varieties of English: Australian/New Zealand English.' In Alexander Bergs and Laurel J. Brinton (eds). *English Historical Linguistics. An International Handbook.* (HSK 34-2). Berlin: Mouton de Gruyter. 1995–2012.

Hundt, Marianne. Forthcoming. 2013. 'Heterogeneity vs. Homogeneity.' To appear in Anita Auer, Dominic Watts and Daniel Schreier (eds). *Letter Writing and Language Change*. Cambridge: Cambridge University Press.

Hundt, Marianne. In preparation. Relatives in scientific English: Variation across time and space. To appear in Winnie Cheng and Franca Poppi (eds). *The Three Waves of Globalization. Proceedings from the 2011 CLAVIER conference*. Amsterdam and Philadelphia: Benjamins.

Hundt, Marianne and Katrin Vogel. 2011. 'Overuse of the progressive in ESL and learner Englishes – fact or fiction?' In Joybrato Mukherjee and Marianne Hundt (eds). *Second-language varieties and learner Englishes*. Amsterdam and Philadelphia: John Benjamins. 145–166.

Hundt, Marianne and Geoffrey Leech. 2012. 'Small is beautiful – on the value of standard reference corpora for observing recent grammatical change.' In T. Nevalainen & E. Traugott (eds.). *The Oxford Handbook of the History of English*. Oxford: Oxford University Press, 175–188.

Hundt, Marianne and Benedikt Szmrecsanyi. 2012. 'Animacy in early New Zealand English.' *English World-Wide* 33(3): 241–263.

Keller, W. 1925. 'Keltisches im englischen Verbum.' In *Anglica: Untersuchungen zur Englischen Philologie, Alois Brandl zum Siebzigsten Geburtsgage überreicht*, In *Sprache und Kulturgeschichte*. (Palaestra, 147). Leipzig: Mayer and Müller. 55–66.

Kranich, Svenja. 2008. *The Progressive in Modern English. A Corpus-Based Study of Grammaticalization and Related Changes*. PhD Dissertation, Freie Universität Berlin.

Leech, Geoffrey, Marianne Hundt, Christian Mair and Nicholas Smith. 2009. *Changes in Contemporary English: A Grammatical Study*. Cambridge: CUP. (Chapter 6, 'The Progressive' by Nicholas Smith).

Mair, Christian and Marianne Hundt. 1995. 'Why is the progressive becoming more frequent in English? – A corpus-based investigation of language change in progress.' *Zeitschrift für Anglistik und Amerikanistik* 43 (2): 111–122.

McCafferty, Kevin and Carolina P. Amador Moreno. 2010. '*A Corpus of Irish English Correspondence* (CORIECOR). A tool for studying the history and evolution of Irish English'. Unpublished ms.

Ronan, Patricia. 2001. 'Observations on the progressive in Hiberno-English.' In John M. Kirk and Dónall P. Ó Baolill (eds.). *Language Links. The Languages of Scotland and Ireland.* Belfast: Cló Ólscoil na Banriona. 43–58.

Rissanen, Matti. 1984. 'The choice of relative pronouns in seventeenth-century American English.' In Jacek Fisiak (ed). *Historical Syntax*. Berlin: Mouton de Gruyter. 419–35.

Smith, Nicholas. 2005. *A Corpus-Based Investigation of Recent Change in the Use of the Progressive in British English*. Unpublished PhD thesis, Lancaster University.

Smitterberg, Erik. 2005. *The Progressive in 19th-century English: A Process of Integration.* Amsterdam: Rodopi.

Strang, Barbara M.H. 1982. 'Some aspects of the history of the *be + ing* construction.' In J. Anderson (ed) *Language Form and Linguistic Variation: Papers Dedicated to Angus McIntosh*. (Current Issues in Linguistic Theory 15). Amsterdam: John Benjamins. 427–74.

Tagliamonte, Sali A. and Sonja Molfenter. 2007. '"How'd you get that accent?" Acquiring a second dialect of the same language.' *Language in Society* 36: 649–675.

Yáñez Bouza, Nuria. 2011. 'ARCHER past and present (1990–2010).' *ICAME Journal* 35: 205–36.

REVIEW

Gounder, Farzana 2011. *Indentured Identities: Resistance and Accommodation in Plantation-era Fiji.* Amsterdam: John Benjamins. Pp. xviii + 345. ISBN 978-90-272-2655-6

Anne Feryok: *Linguistics, Department of English, University of Otago, PO Box 56 Dunedin 9054. <anne.feryok@otago.ac.nz>*

It's got maps and even a couple of old photographs. This was the first thing I noticed on opening *Indentured Identities: Resistance and Accommodation in Plantation-era Fiji,* by Farzana Gounder, a recent PhD graduate from Massey University. The book explores the life narratives of indentured Indians who worked on the plantations of Fiji between 1879 and 1916. The preface titled 'Shards of memories' makes the ties between the author and her research explicit, and sets the tone for this account that is both beautiful and harrowing, both compelling and scholarly, and both personal and political. If you are in the habit of skipping through the front matter: don't. The roman numeral pages, from its dedication to the list of glosses, tell the story of this book.

The book proper begins by introducing us to the seven Girmityas, the Indian laborers indentured to work in Fiji, whose life narratives are the subject of the book. Note is made of the extent to which background, ethnicity, gender, marital status, recruitment and indenture of these Girmityas are typical, and of those voices that are unheard. The research is then introduced through a brief general overview of the linguistic and narrative approaches, and of the chapters to come. The overview of the chapters shows that the book is, in fact, roughly divided in two: the first half is largely the research background and

the second half is the analysis or restorying of the narratives. I say roughly and largely because the division is not neat: the author uses the narratives themselves to explicate her approach to them.

In Chapter 2, one of the narratives is used to provide background knowledge about Girmityas. Chapter 3 explains the context of the narratives as data, which were originally recorded for a Hindi-language Fijian radio station beginning in 1979, using extracts from the narratives to demonstrate how the interviews were co-constructed. In Chapter 4, the focus on methodological considerations continues in the admirably thorough and critical account of the transcription, transliteration and translation of the narratives, again illustrated with narrative extracts. The use of narrative extracts to illustrate the processes of the research continues in Chapters 5 and 6. Unfortunately these two chapters lack the rigor that one might expect given that of Chapter 3, and lack the breadth that one might expect given their length. Chapter 5 uses Labov's (Labov & Waletzky, 1967/2003) structural approach with one of the narratives to show how the narrative was constructed. This chapter has several references to literature that goes beyond the structural approach, namely Ochs and Capps' (2001) dimensional approach, but these differences are unacknowledged. This is unfortunate because the contrasts between dramatic and everyday narratives and the tensions between narrative as activity and text of the dimensional approach might have usefully informed the discussion of habitual narratives and possibly issues raised in Chapter 6 as well. Chapter 6 uses Bamberg's (1997) positioning analysis with the same narrative to show how it was reconstructed, that is, how identity and agency were performed through the narratives. I found myself wishing for a more thorough discussion of the issues this chapter raises about positioning, which are presented in less than two pages. Gounder then unites the two types of analysis; again, I wished for a more thorough discussion of what appears to be a novel and important contribution to narrative analysis.

Chapters 7 through 12 are an impressively detailed analysis of the other six narratives (one of them is in two parts in two chapters). We are carefully guided through each narrative in terms of structural and positioning analysis, so that they are genuinely woven together. The analysed narratives are broken up into segments, which are attractively set off from the text with table-like formatting. They are cleverly organized so that those who want to wallow in the linguistic detail of narrowly transcribed, transliterated and glossed Fijian Hindi text can do so by focusing on the left hand side of the tables, while those who wish to go directly to the English translation may focus on the right

side. There is surprisingly good flow between the text and the narratives in the tables, making these chapters far more readable than might be expected with such a detailed analysis. Nonetheless the brief introduction and overview of the structure of each narrative is essential reading. So, too, is the summary and discussion that concludes each of these chapters, which identify themes not only relevant to the individual narratives, but to the history of Girmit. Although these vary in length, in general I found that they did not say quite as much as I was hoping for. Given the overall length of each chapter, however, brevity may have been compulsory.

The last two chapters do much more than merely bring the book to a conclusion. Chapter 13 sets out a model for narrativization that some will think needs to be explained, supported, and argued for more thoroughly and possibly even presented much earlier. It may well be the case that this model arose through the research process, but that does not mean that temporal location is best mapped onto the corresponding spatial location. It would have enhanced Chapter 6. This could have allowed a single concluding chapter that focused on the themes that arose through the analysis, some of which involve the cultural and historical situatedness of identity and agency, and are interesting enough to deserve a fuller treatment than they receive. As it is, Chapter 14 is a rather mechanical thesis-style conclusion to a book that is anything but mechanical.

This book achieves a difficult balance between the need for technical detail and readability. Much of its readability is because it is a history that many of us may be completely unaware of, but that resonates with other histories we do know about. It is, therefore, a quietly powerful book, one that makes us think not only about language and narrative, but also about the very real people whose lives were lived and relived through language and narrative. This is a praiseworthy accomplishment.

References

Bamberg, Michael. 1997. 'Positioning between structure and performance.' *Journal of Narrative and Life History* 7: 334–342.

Ochs, Elinor and Capps, Linda. 2001. *Living narrative*. Cambridge: Cambridge University Press.

Labov, William and Joshua Waltezky. 1967/2003. 'Narrative analysis: Oral versions of personal experience.' In Christina Bratt Paulston & G. Richard Tucker (eds), *Sociolinguistics: The essential readings*. Malden MA: Blackwell. 74–104.

NEW ZEALAND POSTGRADUATE THESES / RESEARCH PORTFOLIOS IN LINGUISTICS COMPLETED IN 2010–11*

Paul Warren: *School of Linguistics and Applied Language Studies, Victoria University of Wellington <paul.warren@vuw.ac.nz>*

Allan, Stu. 2009. *Passive be damned: the construction that wouldn't be beaten.* MA Thesis, University of Canterbury.

Bram, Barli. 2011. *Major Total Conversion in English: The Question of Directionality.* PhD Thesis, Victoria University of Wellington.

Chapman, Brett. 2011. *Attitudinal Factors Affecting Language Use and Policy in Papua New Guinea with an Emphasis on Tok Pisin*. MA thesis, University of Auckland.

Connor, Cherie. 2010. *A diachronic exploration of the contribution of the harvesting of the marine environment to a distinctive New Zealand English lexicon, 1796–2005*. PhD Thesis, Victoria University of Wellington.

Dimock, Laura. 2010. *A grammar of Nahavaq (Malakula, Vanuatu).* PhD Thesis, Victoria University of Wellington.

Drager, Katie. 2009. *A sociophonetic ethnography of Selwyn Girl's High.* PhD Thesis, University of Canterbury, 2009.

Duhamel, Marie. 2010. *The noun phrase of Atchin, a language of Malakula, Vanuatu*. MA Thesis, University of Auckland.

Featherstone-Santosuosso, Giordana A. 2011. *Lau of North Malaita, Solomon Islands: a language description*. MA Thesis, University of Auckland.

* This list includes some 2008 & 2009 theses omitted from the list published in *Te Reo* 53.

Fletcher, Jeanette. 2011. *The role of discourse in establishing an enabling context for organizational knowledge creation: An ethnographic study.* PhD Thesis, Victoria University of Wellington.

Fu, Xitao. 2011 *Metonymy in Mind, Language, and Communication.* PhD Thesis, Victoria University of Wellington.

Gounder, Farzana. 2011. *Restorying indenture: the first Fiji Hindi speakers narrate Girmit.* PhD Thesis, Massey University.

King, Brian. 2011. *Linguistic Negotiations of Sexual Agency in Sexuality Education.* PhD Thesis, Victoria University of Wellington.

Kingsley, Leilarna. 2010. *Language policy in multilingual workplaces: management, practices and beliefs in banks in Luxembourg.* PhD Thesis, Victoria University of Wellington.

Macdonald, Daryl. 2010. *A grammar sketch of Kwaraque.* MA Thesis, University of Waikato.

Murata, Kazuyo. 2011. *Relational Practice in Meeting Discourse in New Zealand and Japan: A Cross-Cultural Study.* PhD Thesis, Victoria University of Wellington.

Protheroe, Pamela. 2010. *The effect of illustrations on the ability of children to draw inferences while reading narrative texts.* PhD Thesis, Victoria University of Wellington.

Stubbe, Maria. 2010. *"Was that my misunderstanding?" Managing miscommunication and problematic talk at work.* PhD Thesis, Victoria University of Wellington.

Takau, Lana G. 2010. *English-based tongue with Oceanic flavour: a comparison of pronouns and agreement marking in Bislama and Raga.* MA Thesis, University of Canterbury.

Tay, Dennis Zhiming. 2011. *Metaphor in psychotherapy: description and applications*, PhD Thesis, University of Otago.

Walker, Abby. 2008. *Phonetic detail and grammaticality judgements.* MA Thesis, University of Canterbury.

Wilson, Nicholas. 2011. *Leadership as communicative practice: The discursive construction of leadership and team identity in a New Zealand rugby team.* PhD Thesis, Victoria University of Wellington.

Zare, Samad. 2011. *Home and away: blogging emotions a Persian virtual dowreh*. PhD Thesis, Massey University.

BOOKS PUBLISHED BY NEW ZEALAND LINGUISTS IN 2010–11

Paul Warren: *School of Linguistics and Applied Language Studies, Victoria University of Wellington <paul.warren@vuw.ac.nz>*

Angouri, Jo & Meredith Marra. 2011(eds.). *Constructing Identities at Work.* Houndmills: Palgrave MacMillan, 247pp.

Bardsley Dianne. 2011 (2nd edn.) *New Zealand Oxford Primary School Dictionary and Thesaurus.* Melbourne: Oxford University Press, 369pp.

Bardsley Dianne. 2011 (4th edn.) *New Zealand Oxford Junior Dictionary.* Melbourne: Oxford University Press, 268pp.

Bauer Laurie, Dianne Bardsley, Janet Holmes, & Paul Warren. 2011. *Q & Eh - Questions and answers on language with a Kiwi twist.* Auckland: Random House, 330pp.

Carstairs-McCarthy, Andrew. 2010. *The evolution of morphology*. Oxford: Oxford University Press.

Comrie, Bernard, Ray Fabri, Elizabeth Hume, Manwel Mifsud, Thomas Stolz & Martine Vanhove. 2009. *Introducing Maltese Linguistics*. Amsterdam: John Benjamins.

Culicover, Peter & Elizabeth Hume. 2010. *Basics of Language for Language Learners*. Columbus, OH: OSU Press.

Goldsmith, John, Elizabeth Hume and Leo Wetzels. 2011. *Tones and Features*. *Phonetic and Phonological Perspectives*. Mouton de Gruyter.

Gounder, Farzana. 2011. *Indentured Identities: Resistance and Accommodation in Plantation-era Fiji.* Amsterdam: John Benjamins, 345pp.

Holmes Janet, & Meredith Marra. 2010 (eds.), *Femininity, Feminism and Gendered Discourse.* Newcastle upon Tyne: Cambridge Scholars Publishing, 250pp.

Holmes Janet, Meredith Marra, & Bernadette Vine. 2011. *Leadership, Discourse, and Ethnicity.* New York: Oxford University Press, 192pp.

Kearns, Kate. 2011 (2nd edn.). *Semantics*. New York: Palgrave Macmillan.

Kuiper, K. & W.S. Allan. 2010 (3rd edn.). *Introduction to English language: Word, sound and sentence*. Basingstoke: Palgrave Macmillan. xxiv, 364p.

Kuiper, K. 2009. *Formulaic genres.* Basingstoke: Palgrave Macmillan. xiv, 223p.

Kuiper, K. 2010 (ed.). *Yearbook of phraseology* v1. Berlin: Mouton de Gruyter, x, 213.

Kuiper, K. 2011 (ed.). *Teaching linguistics: Reflections on practice*. London: Continuum. xii, 256.

Kuiper, K. 2011 (ed.). *Yearbook of phraseology* v.2. Berlin: Mouton de Gruyter, x, 194.

McKee David, Rachel McKee, Sara Pivac Alexander, Micky Vale, & Lynette Pivac. 2011 *The Online Dictionary of New Zealand Sign Language.* Online, Victoria University of Wellington: Deaf Studies Research Unit.

McKee Rachel, & Janet Davis. 2010 (eds.). *Signed Language Interpreting in Multilingual Contexts. Studies in Interpretation Series, Volume 6.* Washington DC: Gallaudet University Press, 272pp.

Napier, Jemina, Rachel McKee, & Della Goswell. 2010 (2nd edn.). *Sign Language Interpreting: Theory and Practice in Australia and New Zealand.* Sydney: Federation Press.

Song, Jae Jung. 2011 (ed.). *The Oxford Handbook of Linguistic Typology*. Oxford: Oxford University Press.

Torii Shizuja. 2010. *Wa/ga-Subjects in Japanese and Subdivisions of Tense: An In-depth Study of the Syntax and Semantics of Tense and Subjects in Japanese.* Saarbrücken: Lambert Academic Publishing.

van Oostendorp, M., C.J. Ewen, E.V. Hume, & K. Rice. 2011 (eds.). *Companion to Phonology, Volume I: General Issues and Segmental Phonology*. Cambridge, Boston: Wiley-Blackwell.

van Oostendorp, M., C.J. Ewen, E.V. Hume, & K. Rice. 2011 (eds.). *Companion to Phonology, Volume II: Suprasegmental and Prosodic Phonology*. Cambridge, Boston: Wiley-Blackwell.

van Oostendorp, M., C.J. Ewen, E.V. Hume, & K. Rice. 2011 (eds.). *Companion to Phonology, Volume III: Phonological Processes*. Cambridge, Boston: Wiley-Blackwell.

van Oostendorp, M., C.J. Ewen, E.V. Hume, & K. Rice. 2011 (eds.). *Companion to Phonology, Volume IV: Phonological Interfaces*. Cambridge, Boston: Wiley-Blackwell.

van Oostendorp, M., C.J. Ewen, E.V. Hume, & K. Rice. 2011 (eds.). *Companion to Phonology, Volume V: Phonology Across Languages*. Cambridge, Boston: Wiley-Blackwell.

Wallace Derek, & Janet Hughes. 2010. *Fit to Print: the writing and editing style guide for Aotearoa New Zealand.* Wellington: Dunmore Publishing, 214pp.

Wallace Derek. 2011. *Governing the future.* Champaign, IL: Common Ground Publishing, 182pp.

THE EDITOR'S NOTICEBOARD

Paul Warren: *School of Linguistics and Applied Language Studies, Victoria University of Wellington (PO Box 600, Wellington 6140) <paul.warren@vuw.ac.nz>*

My thanks to the authors of the papers submitted for this volume of *te Reo* and to the reviewers for their careful reading of the manuscripts and for the positive contribution that they have made to the journal. This was a particularly interesting year for manuscript submission and review as the management of the journal has now made its transition into the digital age, with on-line submission and review using Open Journal Systems at http://www.nzlingsoc.org/ojs/index.php/TeReo. While there have been a few teething problems using the system, as editor I have found it extremely helpful, and my expectation is that on-line submissions will rapidly improve turn-round times and enable authors to track the progress of their manuscript. In addition, of course, we are reducing the amount of paper that is travelling up and down the country and overseas.

This year there have been some changes to the Editorial Board. Sandy Chung, Kate Kearns and John Lynch have stood down, while Janet Fletcher, Diane Massam and Cynthia White have joined the rest of the board. On behalf of the authors, readers and editors of *te Reo*, I extend thanks to Sandy, Kate and John for their service to the journal, and I welcome Janet, Diane and Cynthia to their new roles. The Editorial Board members are listed in the front matter of the journal, as well as on the website given above.

This is my last volume as editor – I have enjoyed looking after the journal as it has moved through its early 50s, and it has been a good experience getting

to know authors and reviewers who I might not otherwise have encountered. It is time for the journal to pass into new hands and to be housed in a different institution. My colleagues at Victoria University will attest to the fact that an editor frequently resorts to calling on his/her local colleagues for last minute reviews or even submissions, and it is reasonable that such responsibilities should be shared around. Martin Paviour-Smith at Massey University will be taking over the editorship starting with volume 56 in 2013.

Nga mihi nui
Paul