OF WEKA AND WAIATA:
FAMILIARITY WITH BORROWINGS FROM TE REO MAORI [1]

John Macalister: *Victoria University of Wellington, P. O. Box 600, Wellington,* <john.macalister@vuw.ac.nz>

Abstract

Maori words are a distinctive feature of New Zealand English. They account for roughly six of every thousand running words in both spoken and written New Zealand English, and are increasing as a proportion of the language in use. The fact however that an increasing number of Maori words are being used does not necessarily mean that they are familiar to speakers of New Zealand English. Earlier estimates of the number of Maori words (other than proper nouns) familiar to the average speaker of New Zealand English have been in the range of 40 to 50. In order to investigate this question, a survey was designed and implemented with 674 senior secondary students in the greater Wellington area. The survey found differences between male and female respondents in their familiarity with Maori words, and between Maori and non-Maori, with those differences being most pronounced with words referring to material and social cultural aspects of Maoridom. The survey also provided support for earlier claims that future growth in the Maori component of the New Zealand English lexicon is likely to come from the social cultural domain, and suggested that previous estimates of the size of an average New Zealand English speaker's Maori word vocabulary have been too conservative.

1. Introduction

Words of Maori origin have long been recognised as a distinctive feature of New Zealand English and have been described as 'the most unmistakably New Zealand part of New Zealand English' (Deverson 1991: 18). Recent corpus-based assessments of the Maori word presence in both spoken and written New Zealand English have found that such words account for roughly six words per thousand in contemporary New Zealand English running text (Kennedy and Yamazaki, 1999; Macalister, 1999). Approximately two-thirds of these words are proper nouns (including place and personal names, and the type *Maori*), with the remaining third being contributed by the names of flora and fauna, terms referring to the material culture (i.e. visible and tangible objects such as *poi* or *waka*), and terms referring to the social culture of Maoridom (i.e. words such as *aroha, kaumatua* and *kaitiakitanga,* signifying social relationships, concepts and the intangible). Neither the proportion of Maori words in New Zealand English nor their distribution in the semantic categories just mentioned has remained constant, however. In a corpus-based study of the Maori word presence in New Zealand English from 1850 to 2000 (Macalister 2006), the overall trend was for an increase in the proportion of words of Maori origin found in written New Zealand English during the period, and a steep rise in the contribution of the social culture category to this presence in recent years.

It does not necessarily follow however that an increased Maori word presence in the more recent files in the corpus reflects greater familiarity with such borrowings among users of New Zealand English; or, to put it another way, it does not follow that these borrowings have become established in the New Zealand English lexicon. There is relatively little known about the size of an average New Zealander's Maori loanword vocabulary. Deverson (1984) proposed a figure of 40-50 words other than place names as the size of an average New Zealander's Maori vocabulary, which estimate was repeated by Gordon and Deverson (1998) and endorsed by Bellett (1995). This figure refers to the size of an individual's passive vocabulary, rather than an indication of Maori loanwords in active use. This paper reports on a survey that was carried out to investigate the extent to which Maori words in use in the year 2000 files of the corpus corresponded with real-life word knowledge. The paper begins with a discussion of the survey design, followed by a description of the respondent population, the results of the survey, and a discussion of the findings by a number of variables.

2. Survey Questionnaire

The presence of a word of Maori origin in the corpus compiled for this study does not necessarily mean that it is, or was, familiar to readers of the time. Glossing practices and means of marking a word as non-English recognise this. Current newspaper editorial policy in New Zealand tends to allow only a relatively small number of Maori cultural words as well-known additions to the New Zealand English lexicon (see Appendix A). Observation and intuition appear to play a considerable role in determining which words are regarded as likely to be familiar. This is unsurprising as there is no definitive guide as to which Maori words are widely understood by speakers of New Zealand English, and which are not. The best-known attempt to measure familiarity with borrowings from te reo Maori prior to this was Bellett's survey (1995), but there were two apparent weaknesses in its design and implementation:

- the selection of the words tested was to some extent subjective; her research was 'an attempt to identify commonly known [Maori] words' with words being selected because of their 'common appearance within the media or for [their] likely comprehensibility within New Zealand English' or because they 'were words that I felt should have been known' (1995: 79).
- the population sample in most age bands was too small to be representative of the wider population.

These issues were addressed in this survey by basing the selection of questionnaire items on the corpus findings of words in use, and by restricting the survey to one age band.

The survey was designed to be representative of Maori words, other than proper nouns, found in the three files for 2000. The type distribution among the three non-proper noun categories is shown in Table 1.

Table 1: Non-Proper Noun Categories in 2000 Files

	FLORA & FAUNA	MATERIAL CULTURE	SOCIAL CULTURE	TOTAL
No. of Types	51	38	89	178
% of Types	28.7	21.3	50.0	100.0%

On this basis, a 50-item questionnaire was designed containing 14 flora and fauna types, 11 material culture types, and 25 from the largest category, social culture. The corpus was drawn from three sources (newspapers, parliamentary debate and the New Zealand *School Journal*), and it is fair to say that the *School Journals* were the biggest contributor to flora and fauna, the newspaper file to material culture, and *Hansard* to the social culture category. Only three items were common to all three files, those items being *haka, tapu* and *taonga*. A further eighteen items had a presence in two of the three files, and of the remaining 29 items, all except eight were selected because of their frequency. Six types, however, only had one token apiece, and two were chosen despite having no presence in the 2000 files. The inclusion of those two (*taihoa, raupo*) was justified on the grounds of their frequency in earlier files. The effect, therefore, was to include a comprehensive range of Maori words that, on the corpus analysis, ranged from being widespread and frequent, to being low frequency, possibly restricted in use, and conceivably dated.

In designing this survey, early consideration needed to be given to the type of word knowledge being tested. Joe, Nation and Newton (unpublished) have discussed "three major factors which affect the difficulty of test items – the recognition\recall distinction, the receptive\productive distinction, and the imprecise\precise distinction". The primary distinction is between recognising (from choices) and recalling, or providing, the correct form or meaning of the test item. Recognition and recall can be applied to either receptive (i.e. listening and reading) or productive (i.e. speaking and writing) knowledge, and such knowledge can be either of a precise or an imprecise nature. Using these distinctions, Joe et al proposed different test formats for the eight permutations of the three factors affecting difficulty. Knowledge of the words of Maori origin found in the three files for 2000 was felt to require receptive recognition, quite often of an imprecise nature. It would generally, for example, be sufficient to know that a pohutukawa is a tree rather than an ice cream van or a fish; further, it would not usually be necessary to apply detailed knowledge, such as that a pohutukawa generally has red flowers, that it flowers in the summer, or that it is occasionally referred to as the New Zealand Christmas tree. In such cases, Joe, Nation and Newton suggest employing an easy multi-choice test where 'the distractors or choices are not related to each other … and may even be a different part of speech' (ibid: 6).

This is similar to the lowest level of difficulty used by Nagy, Herman and Anderson (1985), where the 'distractors were chosen to be as dissimilar from the target word meaning as possible, even in terms of the implied part

of speech' (ibid: 239). These researchers also had an intermediate level of difficulty, with distractors 'mostly in the same part of speech, but otherwise fairly diverse semantically.' It was towards this level that the questionnaire items tended. Otherwise, any respondent aware that the great majority of borrowings from te reo Maori into New Zealand English are nouns could have guessed the meaning of unfamiliar words, if the distractors for any particular test item had been dissimilar parts of speech.

Pioneers in the field of vocabulary familiarity scores, the Americans Edgar Dale and Joseph O'Rourke, tested 44,000 different items using a three-choice multiple-choice test, partly because it was time-efficient (Dale and O'Rourke, 1979: iv). Nagy, Herman and Anderson (1985) used a six choice format, which included a 'don't know' option. For the current survey, where time-efficiency was not a major consideration, a four-choice format was chosen, with a default option of 'don't know' available through not answering.

Once the questionnaire items had been chosen, a principled method of designing the questionnaire was then considered, and the four choices were identified as being the correct meaning and any three of the four following distractors:

- the meaning of a similar-sounding Maori word type
- the meaning of a similar-sounding English word type
- an item in a related semantic field
- a randomly selected meaning

To illustrate the application of this method, let us consider *toetoe*. The correct answer is *a type of grass*, with a distractor from a related semantic field being *a type of seaweed*; *a type of poison* is another distractor, from the relatively similar-sounding Maori type *tutu*, and the final choice is *a type of clothing*, chosen for its echo of the ballet costume, the *tutu*.

Where a word selected as a questionnaire item was glossed in the corpus, this gloss was taken to supply the meaning of the word as used in New Zealand English, and on a handful of occasions further support for the meaning was sought from the *Dictionary of New Zealand English*. The survey did not attempt to capture the richness of meaning of these borrowings when used in their original language, te reo Maori. The meaning of a word was limited by evidence of its use in New Zealand English.

Furthermore, as the survey was designed to gauge familiarity with words borrowed from te reo Maori, choices for distractors were necessarily limited

to meanings of words that are likely to have been borrowed. Thus, 'ice cream van' as an option for pohutukawa would not have been possible, as there is no likelihood of a loanword signifying such having been borrowed. This consideration was also a factor in tending towards Nagy, Herman and Anderson's intermediate-level model rather than Joe, Nation and Newton's easy multi-choice format (see above). In addition, one further principle applied was that the distractors in any one item should be of similar length (Dale and O'Rourke, ibid: vi). This resulted in, for example, the expansion of 'fish' to 'freshwater fish' when the other three distractors for the test item followed the adjective-noun pattern. Note, however, that the level of difficulty was not necessarily complicated as there was no distinction being called for between fresh- and saltwater fish.

One further important step was also taken during the questionnaire design. The questionnaire was given to a Maori-speaker to ensure that the correct answers were acceptable, and that the distractors were not alternatives that would appear correct to a speaker of te reo Maori. Although the survey was focussing on the meanings of the words as used in New Zealand English, this step was taken in recognition of the fact that at least some of the respondents would be speakers of te reo Maori.

More than the required 50 items were constructed; these were then modified and selected from after a series of observation studies, which operate as follows:

> Doing observation studies involves getting someone to fill in the questionnaire while the researcher watches them and takes notes; the respondents are encouraged to think out loud and talk about how they are going about answering the questions; and the researcher asks them questions if they seem to have a problem and when they have finished. (Statistics New Zealand 1995: 50)

During this process, respondents were asked to say whether they regarded each type as being easy, hard (which included unknown), or somewhere in-between. Subsequent examination of the choices made in the questionnaire suggested that the survey was successfully discriminating between familiar and unfamiliar words, as those judged by the respondents to be easy were generally answered correctly, and those they judged as hard were almost always answered incorrectly, with those regarded as somewhere in-between being answered correctly approximately half the time.

Approaches were made to the principals of fifteen secondary schools in the Greater Wellington area and permission sought to conduct the survey among sixth and/or seventh form students. Eight schools agreed to participate, six declined, and one did not respond. Two versions of the questionnaire were distributed to the participating schools; the only difference between the two versions was in the item order. This was done to counteract the effect on familiarity scores of respondents not having time to complete the questionnaire, and thus leaving the same set of items unanswered.

In all, 674 responses were received. The questionnaire was coded for computer input and analysis, using the software package *SPSS 9.0*. For each item any of five responses was possible; in addition to the correct answer and the distractors, a fifth option, labelled 'do not know', was included to encompass:

- those who did not know and did not hazard a guess
- those who did not answer the question because of (a) lack of time or (b) failure to turn the page
- those who selected more than one answer to a question

While these possibilities of interpretation are acknowledged, it was assumed that in general an unanswered question was a result of exercising the 'do not know' option. This was explicitly signalled as a valid response in the instructions.

In the remainder of this paper the results of the survey are described, beginning with a description of the respondent population. This is followed by sections examining the results by semantic domain – flora and fauna, material culture, social culture. The final section discusses the significance of the results.

3. The Respondents

The survey targeted just one age band. Six hundred and seventy four senior secondary students from eight schools in the greater Wellington region completed the questionnaire during February-March 2002. It was assumed that all respondents, being senior secondary students, had adequate English language skills to understand the distractors. Information about the composition of the respondent population by gender and by ethnicity is provided in Tables 2, 3 and 4.

Table 2: Composition by gender

	NUMBER	PERCENTAGE
Male	196	29.1
Female	477	70.8
Not given	1	0.1
Total	674	100.0

Table 3: Composition by ethnicity

	NUMBER	PERCENTAGE
Pakeha	378	56.1
NZ Maori	72	10.7
Pacific Island Nation	69	10.2
Other	129	19.1
Multiple Responses	26	3.9
Total	674	100.0

Table 4: Composition by gender x ethnicity

	PAKEHA	NZ MAORI	PACIFIC ISLAND NATION	OTHER	MULTIPLE RESPONSES	TOTAL FOR GENDER
Male	116	23	23	26	8	196
	59.2%	*11.7%*	*11.7%*	*13.3%*	*4.1%*	*29.1%*
Female	262	48	46	103	18	477
	54.9%	*10.1%*	*9.6%*	*21.6%*	*3.8%*	*70.8%*
Not given		1				1
		100%				*0.1%*
Total for Ethnicity	378	72	69	129	26	674
	56.1%	*10.7%*	*10.2%*	*19.1%*	*3.9%*	*100.0%*

As can be seen, the population sample was predominantly female and Pakeha. In addition, 78.2% were born in New Zealand, with, as would be expected, the majority of those who identified as Other being born elsewhere (Table 5).

Table 5: Composition by ethnicity x place of birth

	NZ-BORN	NON-NZ-BORN	NO RESPONSE	TOTAL FOR ETHNICITY
Pakeha	352	26		378
	93.1%	*6.9%*		*56.1%*
NZ Maori	70	2		72
	97.2%	*2.8%*		*10.7%*
Pacific Island Nation	41	27	1	69
	59.4%	*39.1%*	*1.4%*	*10.2%*
Other	45	84		129
	34.9%	*65.1%*		*19.1%*
Multiple Responses	19	3	4	26
	73.1%	*11.5%*	*15.4%*	*3.9%*
Total for Place of Birth	527	142	5	674
	78.2%	*21.1%*	*0.7%*	*100.0%*

Of the 674 respondents, 213 (31.6%) said that they had studied te reo Maori, but only 41 (6.1%) claimed to be able to speak the language. Over 85% of those who claimed to speak the Maori language identified as Maori, as shown in Table 6, although, as Table 7 shows, Pakeha formed the largest single group of those who had studied the language.

Table 6: Composition by ethnicity x speaking te reo Maori

	SPEAK MAORI	NOT SPEAK MAORI	NO RESPONSE	TOTAL FOR ETHNICITY
Pakeha	2	374	2	378
	4.9%	*60.1%*	*18.2%*	*56.1%*
NZ Maori	35	33	4	72
	85.4%	*5.3%*	*36.4%*	*10.7%*
Pacific Island Nation	3	66		69
	7.3%	*10.6%*		*10.2%*
Other		129		129
		20.7%		*19.1%*
Multiple Responses	1	20	5	26
	2.4%	*3.2%*	*45.5%*	*3.9%*
Speaking Variable Total	41	622	11	674
	6.1%	*92.3%*	*1.6%*	*100.0%*

Table 7: Composition by ethnicity x studied te reo Maori

ETHNICITY	STUDIED MAORI	NOT STUDIED MAORI	NO RESPONSE	TOTAL FOR ETHNICITY
Pakeha	90	287	1	378
	23.8%	*75.9%*	*0.3%*	*56.1%*
NZ Maori	64	8		72
	88.9%	*11.1%*		*10.7%*
Pacific Island Nation	22	47		69
	31.9%	*68.1%*		*10.2%*
Other	27	102		129
	20.9%	*79.1%*		*19.1%*
Multiple Responses	10	12	4	26
	38.5%	*46.2%*	*15.4%*	*3.9%*
Total for Studied Variable	213	456	5	674
	31.6%	*67.7%*	*0.7%*	*100.0%*

However, it is also worth noting that while almost 90% of Maori had studied the language, less than a quarter of Pakeha had. When examined by gender, a slightly greater percentage of male respondents was found to have studied te reo Maori, as shown in Table 8.

Table 8: Composition by gender x studied te reo Maori

	STUDIED MAORI	NOT STUDIED MAORI	NO RESPONSE	TOTAL FOR GENDER
Male	68	128		196
	34.7%	*65.3%*		*29.1%*
Female	144	328	5	477
	30.2%	*68.8%*	*1.0%*	*70.8%*
Not given	1			1
	100.0%			*0.1%*
Total for Studied Variable	213	456	5	674
	31.6%	*67.7%*	*0.7%*	*100.0%*

Finally, the survey incorporates a response to Greenberg's suggestion that one of the three indices of assimilation into a borrowing language is sociological.[2] 'Currency in various socio-economic strata and occupational groups,' he

argues (1971: 188), 'is a basic sociological factor'. If it can be shown that a loanword is familiar throughout society, that is an argument in favour of its having been assimilated into the lexicon. The inclusion of a socio-economic variable in the survey acknowledges this suggestion. The schools that agreed to assist in this survey were drawn from across the socio-economic spectrum, as measured by a school's decile ranking. Three decile bands were established for this survey, being:

Band 1: Deciles 1-3 (the bottom socio-economic band)
Band 2: Deciles 4-7 (the mid socio-economic band)
Band 3: Deciles 8-10 (the top socio-economic band)

Their contribution to the composition of the sample is given in Table 9.

Table 9: Composition by decile band

	NUMBER	PERCENTAGE
Decile 1-3	170	25.2
Decile 4-7	89	13.2
Decile 8-10	415	61.6
Total	674	100.0

The contribution of the mid-band was small, and consisted solely of female respondents, making it unrepresentative of the band as a whole and therefore unsuitable for comparative purposes. Decile band comparisons made in the following discussion will, therefore, be between top and bottom bands only. As can be seen in Table 10 (overleaf), the top and bottom bands show significant differences in Pakeha and Pacific Island Nation representation, when analysed by ethnicity.

Before moving to a discussion of the results of the survey, two questions remain to be asked. The first is: to what extent can this respondent population be taken as representative of New Zealand English speakers? Common sense suggests that senior secondary school students who will be leaving school within the year to enter the work-force or to seek further education represent a base-line of adult word knowledge. Continued exposure to language in use is likely to lead to increased vocabulary size. For example, newspaper readership increases with age (Newspaper Publishers' Association 2002), so

Table 10: Composition by decile x ethnicity

	PAKEHA	NZ MAORI	PACIFIC ISLAND NATION	OTHER	MULTIPLE` RESPONSES	TOTAL FOR DECILE
Decile 1-3	70	22	42	27	9	170
	41.2%	12.9%	24.7%	15.9%	5.3%	25.2%
Decile 4-7	42	4	5	33	5	89
	47.2%	4.5%	5.6%	37.1%	5.6%	13.2%
Decile 8-10	266	46	22	69	12	415
	64.1%	11.1%	5.3%	16.6%	2.9%	61.6%
Total for Ethnicity	378	72	69	129	26	674
	56.1%	10.7%	10.2%	19.1%	3.9%	100.0%

not only are more people likely to have exposure to the Maori words contained in that genre, but over time there will be increased exposure to those words, potentially enhancing their familiarity to readers. Familiarity with words of Maori origin among senior secondary students may, therefore, be seen as a minimum among adult speakers of New Zealand English. However, common sense also suggests that, as this respondent population is drawn from one urban region, there may be slight differences between its familiarity with words of Maori origin and that of comparable groups in other regions or in rural areas. Respondents in the Waikato region, for instance, may be more familiar with *raupatu* than respondents in any other region, because of Tainui's settlement with the Crown of its land confiscation grievances, and *raupo* may be more familiar to those who live in an area where the plant flourishes than to city-dwellers. That said, the size of the respondent population and the nature of the variables monitored suggest that the findings are broadly generalisable to the population as a whole. Claims about familiarity with words of Maori origin among any particular group are not based on responses from a handful of individuals, who may or may not be representative of the group they are intended to represent. It would, for example, be unhelpful to make claims about familiarity with Maori words among female New Zealanders if all respondents were Pakeha from high-decile schools.

The second question is: to what extent can this respondent population be expected to have had exposure to the items in the questionnaire? The

items were drawn from three sources. New Zealand-born respondents and a proportion of those born elsewhere are likely to be familiar with the language used in the *School Journals,* as they are distributed to every school in New Zealand. Among 15-19 year olds, 39% read a newspaper daily, and 70% read a newspaper a week (Newspaper Publishers' Association 2002).[3] It is unlikely that any of the respondents read *Hansard*, but this would be true of the New Zealand population as a whole. The Maori words used in *Hansard*, however, as with those found in the newspapers and the *School Journals*, are not restricted to that medium. Respondents are likely to be familiar with Maori words from a range of situations, including encountering them in the classroom. It seems reasonable to suggest, therefore, that the questionnaire items are potentially familiar to the respondents. The extent to which they did indeed prove familiar is discussed in the following sections of this paper.

4. Flora and Fauna Results

The 50-item questionnaire contained fourteen items drawn from the flora and fauna category. The percentage of people doing the test who could show that they knew the meaning of these words is given in Table 11, as are the percentages by a number of variables.

Table 11: Flora and Fauna Results (% correct)

TYPE	% CORRECT	MALE %	FEMALE %	PAKEHA %	MAORI %	SPEAK %	NOT SPEAK %`	STUDY %	NOT STUDY %
pohutukawa	94.7	91.3	96.2	97.4	95.8	97.6	94.4	98.1	93.0
pipi	88.7	87.8	89.3	94.4	87.5	90.2	88.6	91.1	87.7
weta	85.2	82.1	86.4	92.9	86.1	87.8	84.7	86.4	84.6
kea	84.3	80.4	86.2	92.1	80.6	78.0	84.6	84.0	84.4
pukeko	81.2	78.1	82.6	89.2	79.2	80.5	81.0	83.1	80.3
kowhai	78.8	64.3	84.9	84.9	77.8	82.9	78.1	81.7	77.4
kina	76.4	74.0	77.4	78.3	94.4	95.1	75.2	85.9	72.1
hoki	72.3	69.4	73.4	77.5	73.6	75.6	71.5	72.8	71.9
toetoe	53.0	41.8	57.7	64.6	47.2	43.9	54.2	50.7	53.9
raupo	33.8	32.1	34.6	36.2	37.5	46.3	33.3	38.0	32.0
kotuku	29.4	29.6	29.1	26.5	43.1	63.4	26.8	36.6	25.9
piwakawaka	10.2	10.7	10.1	8.7	20.8	36.6	8.5	10.8	10.1
tieke	8.3	7.7	8.6	7.7	13.9	22.0	7.4	10.3	7.2
akeake	7.4	10.2	6.3	7.9	2.8	4.9	7.7	6.1	8.1

It can be seen that female respondents had higher familiarity scores than males on the ten most familiar of these items, and that Pakeha respondents had higher familiarity scores than Maori on the nine most familiar of these items, apart from *kina*, for which the comparison between Pakeha and Maori was 78.3% : 94.4%.

Being Pakeha and being New Zealand-born were the best indicators of familiarity with the nine highest ranking of these items. Those who were speakers of te reo Maori were much more likely to be familiar with the less well-known borrowings, particularly *kotuku*, where the comparison between speakers and non-speakers of Maori was 63.4% : 26.8%.

5. Material Culture Results

Eleven types drawn from the material culture category were included in the survey questionnaire. The percentage of people doing the test who could show that they knew the meaning of these words is given in Table 12, as well as the results by a number of variables.

Table 12: Material Culture Results (% correct)

TYPE	% CORRECT	MALE %	FEMALE %	PAKEHA %	MAORI %	PACIFIC ISLAND NATION %	SPEAK %	NOT SPEAK %	STUDY %	NOT STUDY %
waka	87.1	80.1	89.9	90.7	93.1	85.5	92.7	86.5	90.6	85.5
poi	84.6	79.6	86.8	88.6	90.3	81.2	90.2	83.9	87.8	83.1
hangi	77.4	78.6	77.1	80.2	87.5	81.2	85.4	76.7	82.2	75.2
pounamu	73.0	65.3	76.3	75.1	84.7	78.3	87.8	71.9	80.3	70.0
whare kai	62.8	51.0	67.5	59.3	81.9	85.5	85.4	61.1	71.8	58.8
pa	62.3	65.8	60.8	69.6	79.2	46.4	87.8	60.9	71.4	58.1
maunga	61.0	51.0	65.2	59.0	83.3	73.9	85.4	59.2	67.6	58.1
koru	52.4	38.8	58.1	56.3	69.4	36.2	73.2	50.5	53.1	52.0
taiaha	51.9	59.2	49.1	46.6	79.2	65.2	85.4	49.4	62.4	46.9
nohoanga	28.2	20.9	31.2	21.2	47.2	53.6	61.0	25.4	31.5	26.3
paepae	21.5	14.3	24.5	15.9	48.6	29.0	65.9	18.0	28.2	18.2

Female respondents had higher familiarity scores on all items except *pa* and *taiaha*. Male respondents also scored marginally better on *hangi*, but the difference was not significant (78.6% : 77.1%).

Being Maori and being a speaker of te reo Maori were the best indicators of familiarity with borrowings in this category, although those who identified as Pacific Island Nation scored better than Maori for *whare kai* (85.5% : 81.9%) and *nohoanga* (53.6% : 47.2%). Both these scores, however, were markedly better than those of Pakeha respondents (59.3% and 21.2% respectively).

6. Social Culture Results

The largest of the three semantic domains represented in the survey was that of social culture, which contributed 25 types, or 50% of the total. The results for these items are provided in Table 13.

Table 13: Social Culture Results (% correct)

TYPE	% CORRECT	MALE %	FEMALE %	PAKEHA %	MAORI %	PACIFIC ISLAND NATION	SPEAK %	NOT SPEAK %	STUDY %	NOT STUDY %
haka	96.1	92.3	97.7	97.4	98.6	89.9	97.6	96.0	97.7	95.4
aroha	82.6	69.4	88.3	84.9	86.1	89.9	90.2	82.2	88.3	80.3
te reo	81.2	73.5	84.3	85.2	91.7	87.0	97.6	80.2	87.3	78.5
tapu	79.1	71.9	82.0	82.8	91.7	81.2	92.7	78.0	87.3	75.2
hui	73.4	73.0	73.8	77.2	86.1	68.1	90.2	72.5	71.8	74.3
whakapapa	62.9	54.6	66.5	59.5	87.5	72.5	95.1	60.5	73.2	58.3
taonga	61.7	52.0	65.8	60.6	86.1	68.1	90.2	59.8	70.0	58.1
mihi	61.3	52.6	65.0	58.2	95.8	71.0	95.1	58.8	72.8	55.9
korero	58.8	48.5	63.1	53.4	88.9	79.7	95.1	55.9	74.2	51.5
mokopuna	56.7	51.0	59.1	52.6	88.9	71.0	92.7	53.7	69.0	50.7
kaumatua	55.5	48.5	58.5	54.0	69.4	71.0	73.2	54.5	64.3	52.0
karakia	53.6	47.4	56.0	47.1	86.1	72.5	90.2	50.5	67.1	46.9
mana whenua	47.5	37.8	51.6	49.2	50.0	53.6	53.7	46.8	47.9	47.6
utu	42.6	57.7	36.5	41.0	66.7	44.9	73.2	40.4	58.7	35.3
hapu	42.1	33.7	45.7	40.7	69.4	43.5	92.7	38.6	49.8	38.4
wairua	41.7	32.1	45.7	38.6	56.9	43.5	70.7	39.4	46.0	39.7
kura kaupapa	40.7	35.7	42.8	35.4	68.1	55.1	78.0	37.8	55.4	33.8
taihoa	40.1	35.7	41.9	37.0	75.0	39.1	80.5	37.1	53.5	33.6
tumuaki	36.2	27.0	40.0	36.5	52.8	30.4	58.5	34.2	46.0	31.6
hikoi	35.2	33.7	35.8	31.0	63.9	42.0	73.2	32.2	41.3	32.0
taha Maori	34.1	29.1	36.3	31.5	62.5	27.5	61.0	32.2	43.7	29.6
kaitiaki	30.9	23.0	34.0	28.6	56.9	26.1	70.7	28.0	38.0	27.6
raupatu	27.3	19.9	30.4	29.6	29.2	21.7	34.1	27.0	25.4	28.3
rangatiratanga	21.1	14.8	23.7	21.2	34.7	15.9	41.5	19.6	24.4	19.7
rahui	11.9	16.8	9.6	11.4	13.9	7.2	17.1	11.4	9.4	12.9

Female respondents had higher familiarity scores than males on all items except *utu* and *rahui,* the least familiar item in this domain.

Being Maori and being a speaker of te reo Maori were the best indicators of familiarity with borrowings in this category, although those who identified as Pacific Island Nation scored better than Maori for *aroha*, *kaumatua* and *mana whenua*. Furthermore, Pacific Island Nation respondents tended to display greater familiarity with the types in this category than did Pakeha.

Among speakers of te reo Maori the words that were least likely to be familiar were those that belonged to the realm of politics: *mana whenua* (53.7%), *raupatu* (34.1%), *rangatiratanga* (41.5%), and *rahui* (17.1%).[4] All the same, they were more familiar with these types than non-speakers of te reo Maori.

7. Discussion of Findings

The survey was not presented as a test. There was, therefore, no compulsion upon respondents to try and 'get it right'. This is important to bear in mind when discussing the results. The assumption is that when respondents selected an option they believed it to be the correct answer, because the default option of not answering was available to signal 'I do not know' and not answering bore no penalty. There remains, of course, the possibility that a proportion of those who selected any given option did so through random selection. It is not, however, possible to draw inferences about respondents' thinking and thus it is necessary to assume the accuracy of the answers given. In this context, it may be useful to recall that during the pilot stage the questionnaire was found to discriminate successfully between 'easy' and 'hard' words. With those thoughts in mind, it is possible to propose for discussion purposes five bands along a continuum of familiarity, as set out below. These bands were chosen on the assumption that the percentage of respondents correctly recognising a given meaning was a reasonable indication of the word's familiarity to a speaker of New Zealand English. Thus, if 85% of respondents correctly identified a word's meaning, this was taken to indicate that the word would be familiar to most speakers of New Zealand English.

Degree of familiarity	Description
80-100%	likely to be familiar to most speakers of New Zealand English
60-79%	likely to be familiar to a majority of speakers of New Zealand English
40-59%	likely to be familiar to around half the speakers of New Zealand English, and to become better known
20-39%	likely to be familiar to a minority of speakers of New Zealand English
0-19%	likely to be known to a relatively few speakers of New Zealand English, and generally to be regarded as an unknown word

Words in the two highest bands can, therefore, be said to have become reasonably well-established in the New Zealand English lexicon, whereas those in the two lowest bands are not well-known and could reasonably be expected to be glossed in an English language text. In the middle band, although they may still require glossing, the words are in a transition phase. They may be moving into the lexicon, or falling out of it. It seems likely, however, that if around half the population is familiar with a borrowing, and that borrowing continues to be used in written New Zealand English, it will grow in familiarity and be increasingly well-established. Such an outcome may, however, depend not only on the frequency of the word's occurrence, but also on the frequency with which a language user encounters the word's referent. Encounters with a kiwi may be rare, but *kiwi* is a relatively frequent word. *Raupo* may be an infrequent word, but may well be familiar to those who live where it grows. *Tieke* is not only an infrequent word but, being an endangered species restricted to off-shore islands, is most unlikely to be frequently encountered. The interplay of these two factors may have some predictive value when considering the likely familiarity of a loanword to a speaker of New Zealand English.

Before discussing the results by semantic domain and by variable, it is necessary to emphasise that the above interpretative guide is intended as a continuum. The division of the continuum into five bands is a convenience, for discussion purposes. There is unlikely to be a real difference in familiarity between a type with a 61% score, and a type with a 59% score.

7.1 The Semantic Domains

The questionnaire findings can be analysed in terms of the five bands of the continuum described above, and tentative conclusions drawn. Table 14 assigns the proportion of correctly answered items in each of the three domains to the appropriate band.

Table 14: Correct Answers (% bands) x Semantic Domains

	FLORA & FAUNA	MATERIAL CULTURE	SOCIAL CULTURE
80-100%	36%	18%	12%
60-79%	21%	45%	20%
40-59%	7%	18%	40%
20-39%	14%	18%	24%
0-19%	21%	0	4%

The largest proportion of well-known borrowings is to be found in the flora and fauna and the material culture categories, but the category with the greatest potential to contribute to growth in the Maori word presence in the lexicon would appear to be social culture, where fully 40% of the loanwords are in the middle band. This is in line with earlier predictions about the likely source of growth in types of words of Maori origin within the New Zealand English lexicon (Ryan 1977; Macalister 1999: 48). Indeed, it is probable that a speaker of New Zealand English will be familiar with a greater number of social than material culture borrowings already, given that the former contributed more than double the number of items to the questionnaire as the latter.

The category with the greatest proportion of poorly known borrowings was flora and fauna. This would appear to confirm a suspicion that much of New Zealand's plant- and wildlife is unfamiliar, both by name and by appearance, to a majority of New Zealanders, and the result is thus not unexpected. Kennedy and Yamazaki (1999) have pointed out the disparity between *DNZE* and two corpora in the presence of flora and fauna types, and the weight given to this apparently specialist vocabulary has also been noted by Ryan (1977), and is well-exemplified by Andersen (1946).

Overall, 46% of the types in the 50-item questionnaire were found in the two highest bands. Given that the 50 items were drawn from a database of 178 types found in the 2000 files, and assuming that the 128 types not included in

the survey were of comparable difficulty, it seems probable that Gordon and Deverson's estimate (1998) that non-Maori speakers of New Zealand English are familiar and comfortable with 40 to 50 borrowings from te reo Maori, other than place names, is now an under-estimate.[5]

7.2 The Variables

Information on a number of variables was gathered during the implementation of the survey. The relevance of these variables is discussed below.

Gender

There is a consistent difference between male and female respondents. Females show greater familiarity with borrowings from te reo Maori than do males, despite the fact that male respondents are more likely to have studied the language (refer Table 8). The exception is a small group of words – *pa, taiaha, utu*.

Given that a greater proportion of female respondents selected Other under ethnicity (Table 4), and that those who identified as Other were generally not born in New Zealand (Table 5), the difference between male and female respondents is made more striking. This finding is contrary to Bellett's survey which found 'no significant differences between genders' (1995: 88).

Ethnicity

Those who identify as New Zealand Maori are likely to have greatest familiarity with borrowings in the material and social culture domains, but Pakeha are likely to have greatest familiarity with flora and fauna borrowings. Pacific Island Nation respondents display greater familiarity with social culture borrowings than do Pakeha. Twelve social culture items were correctly answered by 60-100% of Pacific Island Nation respondents; the corresponding figure for Pakeha respondents was six items.

Place of Birth

Those born outside New Zealand included new migrants, refugees and exchange students, as well as those who arrived in the country as infants. No information was gathered as to the length of time non-New Zealand born respondents had been in the country. However, it is no surprise that being New Zealand-born inclines one to familiarity with borrowings from te reo Maori. Only two types were correctly recognised by more than 80% of non-New Zealand-born respondents – *haka* and *pohutukawa*.

Speaking Te Reo

Again, it is no surprise that speakers of te reo Maori show greater familiarity with borrowings than do non-speakers. The only occasion when non-speakers out-performed speakers was with a handful of flora and fauna types – *kea, pukeko, toetoe* – although the difference with *pukeko* was minimal. It is possible that the use of *toetoe* rather than the frequent but erroneous *toitoi*[6] may have mislead some Maori-speaking respondents, but this is unverifiable. *Toetoe* was used in the survey as that was the form found in the files for 2000. Similarly, the choice of *piwakawaka* rather than one of the many variant forms may also have militated against some Maori-speakers. Again, however, this was the form present in the source data.

Studying Te Reo

Almost one-third of respondents had studied te reo Maori, and this does appear to have affected the familiarity scores. Those who have studied the language are familiar with a greater number of borrowings, and have a better recognition rate even for items that do not appear to be familiar.

Decile Ranking

Top-band respondents tended to show greater familiarity with flora and fauna terms than did bottom-band respondents, apart from a handful of kai moana terms – *pipi, kina, hoki*. In the other two semantic domains, however, bottom-band respondents show greater familiarity overall, and out-perform top-band respondents on all except a few items. The type on which the top-band group do most significantly better is *koru*.

Generally, however, there is no strong evidence that socio-economic factors are a determinant of familiarity with borrowings. Rather, borrowings tend to be similarly known across the socio-economic spectrum. The differences that do exist may be the result of differences in the ethnic composition of the decile bands (see Table 10). This is in line with the finding in Bauer and Bauer's study of playground language that the forms 'closely associated with the Maori population … are associated most strongly with low decile' (2000: 60).

Overall, therefore, the survey supports earlier research (Thompson 1990; Kennedy 2001; Macalister 2001) that suggests the presence of Maori words in New Zealand English is more likely to be a feature of language use by those who identify as Maori than by non-Maori. This feature is undoubtedly linked to the fact that Maori are more likely than non-Maori to have studied and to

speak te reo Maori (Tables 6, 7). However, on the evidence of this survey, it is the presence of Maori words from the cultural domains rather than from flora and fauna that is the distinctive lexical feature of English language use by those who identify as Maori.

8. Conclusion

The survey was intended as a tool to investigate the link between the Maori words in use in New Zealand English in 2000 and real-life word knowledge. The survey results suggest that a reasonably sized body of loanwords are established in New Zealand English – in the sense that the correct meaning of a word of Maori origin can be recognised even when presented devoid of context. Further, the survey suggests that the standard estimate of the Maori loanword vocabulary of an average New Zealander as 40-50 words other than place names is conservative. Nevertheless, there is a convincing case that many loanwords need to be supported by glosses and/or context clues if they are to be successfully understood. Such practices are, of course, likely to assist in their establishment.

The survey also suggests that further expansion of the Maori word component of the New Zealand English lexicon is likely, coming in particular from the social culture category. This is likely not only because of the number of responses in this category in the mid-band of the familiarity continuum, but because of the nature of the respondent population. Their loanword lexicon is likely to increase with age, and with increased exposure to loanword use. One could argue, therefore, that the survey presents a snapshot of the loanword vocabulary with which school leavers enter adult life.

Finally, it is probable that, given the size of the respondent population, the findings of the survey, and in particular the effect of the identified variables, can be transferred to the senior secondary student population as a whole. The chief caveat would be that different levels of familiarity with borrowings from te reo Maori may exist between rural and urban communities. This claim will be one of the foci of a planned further implementation of the survey[7], proposed for 2007, which will investigate regional differences, as well as evidence of any on-going changes, in the level of familiarity with words of Maori origin in New Zealand English.

Appendix A: Maori words that would not be glossed

The following table is based upon information supplied by three of the four newspapers included in the corpus. This information was supplied in 2000.

	NEW ZEALAND HERALD	*WANGANUI CHRONICLE*	*THE DOMINION*
aroha	x		
haka	x		x
hangi	x		x
hongi	x		x
hui	x		x
iwi	x	x	x
kai	x		
kaumatua	x		
mana	x		x
pa	x		
Pakeha	x		
poi	x		x
tangi	x		x
tangata whenua	x		x
taonga	x		
utu	x		
waka	x		x
whanau	x		x
hapu		x	
karakia		x	
karanga		x	
powhiri		x	
waiata		x	
marae			x
mere			x
taniwha			x
tapu			x
tiki			x

Notes

1 A summary of this research has previously appeared in *English in Aotearoa* 52 (May 2004): 69-73.
2 The two other proposed indices were linguistic and psychological. The survey questionnaire was not an appropriate tool to address either of those indices. The linguistic index, for example, includes phonology, and the psychological index would require speakers making judgements about the extent to which borrowings were 'native'.
3 A note of caution should be sounded here, as these are industry findings, and it is not known what 'reading a newspaper' entails.
4 Politics here is loosely defined and includes land confiscation (*raupatu*) and decision-making, such as the decision to impose a ban (*rahui*).
5 Given that 46% x 178 = 81.88 , a more accurate assessment might be familiarity with 70-80 Maori words.
6 So-described in the *Dictionary of New Zealand English*.
7 Because of this, the questionnaire has not been appended to this article.

References

Andersen, Johannes C. 1946. 'Maori words incorporated into the English language.' *Journal of the Polynesian Society* 55: 141–162.

Bauer, Laurie and Winifred Bauer. 2000. 'The influence of the Maori population in NZ dialect areas.' *Te Reo* 43: 39–61.

Bellett, Donella. 1995. Hakas, Hangis and Kiwis: Maori lexical influence on New Zealand English. *Te Reo* 38: 73–104.

Dale, Edgar and Joseph O'Rourke. 1979. *The Living Word Vocabulary: Supplement to Existing 1976 Edition* [Chicago?]: Field Enterprises Educational Corp.

Deverson, Tony. 1984. '"Home Loans": Maori input into current New Zealand English.' *English in New Zealand* 33: 4–10.

Deverson, Tony. 1991. 'New Zealand English lexis: the Maori dimension.' *English Today* 26: 18–25

Gordon, Elizabeth and Tony Deverson. 1998. *New Zealand English and English in New Zealand* Auckland: New House.

Greenberg, Joseph H. 1971. *Language, Culture and Communication: Essays by Joseph H. Greenberg*. Stanford: Stanford University Press.

Joe, Angela, Paul Nation and Jonathon Newton. (unpublished) 'Choosing the right vocabulary test item.'

Kennedy, Graeme. 2001. 'Lexical borrowing from Maori in New Zealand English.' In Bruce Moore (ed) *Who's Centric Now?: The Present State of post-Colonial Englishes*. Melbourne: Oxford University Press. 59–81

Kennedy, Graeme and Shunji Yamazaki. 1999. 'The influence of Maori on the New Zealand English lexicon.' In John M. Kirk (ed) *Corpora Galore: Analyses and Techniques in Describing English*. Amsterdam: Rodopi. 33–44.

Macalister, John.1999. 'Trends in New Zealand English: Some observations on the presence of Maori words in the lexicon.' *New Zealand English Journal* 13: 38–49
Macalister, John. 2001. 'Writing Maori English: voices in *Pounamu, Pounamu*.' *Kotare* 4, 1: 46–54.
Macalister, John. 2006. The Maori Presence in the New Zealand English Lexicon. *English World-Wide* 27, 1: 1–24.
Nagy, William E., Patricia A. Herman and Richard C. Anderson. 1985. 'Learning words from context.' *Reading Research Quarterly* 20: 233–253.
Newspaper Publishers' Association. 2002. *Information About Newspapers 2002/3*. Wellington: Newspapers Publishers' Association.
Ryan, J. S. 1977. 'The Tolerance of Maori in New Zealand English.' *Orbis* 26, 2: 341–370.
Statistics New Zealand. 1995. *A Guide to Good Survey Design*. Wellington: Statistics New Zealand.
Thompson, Wendy. 1990. 'Attitudes to Maori and the use of Maori lexical items in English.' *Wellington Working Papers in Linguistics* 1: 37–46.